MULTISCALE MODEL. SIMUL.
Vol. 23, No. 1, pp. 25–130

STABLE NEARLY SELF-SIMILAR BLOWUP OF THE 2D BOUSSINESQ AND 3D EULER EQUATIONS WITH SMOOTH DATA II: RIGOROUS NUMERICS*

JIAJIE CHEN† AND THOMAS Y. HOU‡

Abstract. This is Part II of our paper in which we prove finite time blowup of the two-dimensional Boussinesq and three-dimensional axisymmetric Euler equations with smooth initial data of finite energy and boundary. In Part I of our paper [Chen and Hou, preprint, arXiv:2210.07191, 2022], we establish an analytic framework to prove the nonlinear stability of an approximate self-similar blowup profile using a combination of weighted L^∞ and weighted $C^{1/2}$ energy estimates. We reduce proving nonlinear stability to verifying several inequalities for the constants in the energy estimate which depend on the approximate steady state and the weights in the energy functional only. In Part II of our paper, we construct approximate space-time solutions with rigorous error control, which are used to obtain sharp stability estimates of the linearized operator in Part I. We also obtain sharp estimates of the velocity in the regular case using numerical integration with computer assistance. These results enable us to verify that the constants in the energy estimate obtained in Part I [Chen and Hou, preprint, arXiv:2210.07191, 2022] indeed satisfy the inequalities for nonlinear stability. The nonlinear stability further implies the finite time singularity of the axisymmetric three-dimensional Euler equations with smooth initial data and boundary.

Key words. 3D Euler singularity, approximate space-time solution, computer-assisted proof, numerical integral

MSC codes. 35Q31, 65G50

DOI. 10.1137/23M1580395

1. Introduction.

The three-dimensional (3D) incompressible Euler equations are one of the most fundamental nonlinear partial differential equations that govern the motion of the ideal inviscid fluid flow. They are closely related to the incompressible Navier–Stokes equations. Due to the presence of nonlinear vortex stretching, the global regularity of the 3D incompressible Euler equations with smooth initial data and finite energy has been one of the longstanding open questions in nonlinear partial differential equations. Let $\mathbf{u}$ be the divergence free velocity field and we define $\boldsymbol{\omega} = \nabla \times \mathbf{u}$ as the *vorticity vector*. The 3D Euler equations governing the vorticity $\boldsymbol{\omega}$ are given by

$$\boldsymbol{\omega}_t + \mathbf{u} \cdot \nabla \boldsymbol{\omega} = \boldsymbol{\omega} \cdot \nabla \mathbf{u}, \tag{1.1}$$

where $\mathbf{u}$ is related to $\boldsymbol{\omega}$ via the *Biot–Savart law*. The velocity gradient $\nabla \mathbf{u}$ formally has the same scaling as vorticity $\boldsymbol{\omega}$. Thus the vortex stretching term, $\boldsymbol{\omega} \cdot \nabla \mathbf{u}$, has a nonlocal quadratic nonlinearity in terms of vorticity. Although many experts tend to believe that the 3D Euler equations would form a finite time singularity from smooth initial data, the nonlocal nature of the vortex stretching term could lead to dynamic depletion of nonlinearity, thus preventing a finite time blowup (see, e.g., [20, 23, 36]).

*Received by the editors June 20, 2023; accepted for publication April 18, 2024; published electronically January 6, 2025.
https://doi.org/10.1137/23M1580395
Funding: The research was in part supported by NSF grants DMS-1907977 and DMS-2205590. We would like to acknowledge the generous support from Mr. K. C. Choi through the Choi Family Gift Fund and the Choi Family Postdoc Gift Fund.
†Courant Institute, NYU, New York, NY 10012, USA (jiajie.chen@cims.nyu.edu).
‡Applied and Computational Mathematics, Caltech, Pasadena, CA 91125, USA (hou@cms.caltech.edu).

The interested readers may consult the excellent surveys [19, 30, 34, 38, 43] and the references therein.

Our work is inspired by the computation of Luo and Hou [41, 42], in which they presented some convincing numerical evidence that the 3D axisymmetric Euler equations with smooth initial data and boundary develop a potential finite time singularity. In Part I of our paper [13], we establish an analytic framework and obtain the essential stability estimates to prove finite time singularity of the 2D Boussinesq and 3D axisymmetric Euler equations with smooth initial data and boundary. The main results of this paper are stated by the two informal theorems below. The more precise and stronger statement of Theorem 1 can be found in Theorem 3 in section 2.

THEOREM 1. *Let θ, $\mathbf{u}$, and ω be the density, velocity, and vorticity in the 2D Boussinesq equations* (2.3)–(2.5)*, respectively. There is a family of smooth initial data (θ_0, ω_0) with $\theta_0(x, y)$ being even and $\omega_0(x, y)$ being odd in x, such that the solution of the 2D Boussinesq equations develops a singularity in finite time $T < +\infty$. The velocity field $\mathbf{u}_0$ has finite energy. The blowup solution $(\theta(t), \omega(t))$ is nearly self-similar in the sense that $(\theta(t), \omega(t))$ with suitable dynamic rescaling is close to an approximate blowup profile $(\bar{\theta}, \bar{\omega})$ up to the blowup time. Moreover, the blowup is stable for initial data (θ_0, ω_0) close to $(\bar{\theta}, \bar{\omega})$ in some weighted L^∞ and $C^{1/2}$ norm.*

THEOREM 2. *Consider the 3D axisymmetric Euler equations in the cylinder $r, z \in [0, 1] \times \mathbb{T}$. Let u^θ and ω^θ be the angular velocity and angular vorticity, respectively. The solution of the 3D Euler equations* (2.1)–(2.2) *develops a nearly self-similar blowup (in the sense described in Theorem 1) in finite time for some smooth initial data ω_0^θ, u_0^θ supported away from the symmetry axis $r = 0$. The initial velocity field has finite energy, and u_0^θ and ω_0^θ are odd and periodic in z. The blowup is stable for initial data $(u_0^\theta, \omega_0^\theta)$ that are close to the approximate blowup profile $(\bar{u}^\theta, \bar{\omega}^\theta)$ after proper rescaling subject to some constraint on the initial support size.*

We first review some main ideas in our stability analysis of the linearized operator presented in Part I [13]. We use the 2D Boussinesq system as an example. Let $\bar{\omega}$, $\bar{\theta}$ be an approximate steady state of the dynamic rescaling formulation. We denote $W = (\omega, \theta_x, \theta_y)$ and decompose $W = \overline{W} + \widetilde{W}$ with $\overline{W} = (\bar{\omega}, \bar{\theta}_x, \bar{\theta}_y)$. We further denote by $\mathcal{L}$ the linearized operator around $\overline{W}$ that governs the perturbation $\widetilde{W}$ in the dynamic rescaling formulation (see section 2):

$$\widetilde{W}_t = \mathcal{L}(\widetilde{W}). \tag{1.2}$$

We decompose the linearized operator $\mathcal{L}$ into a leading order operator $\mathcal{L}_0$ plus a finite rank perturbation operator $\mathcal{K}$, i.e., $\mathcal{L} = \mathcal{L}_0 + \mathcal{K}$. The leading order operator $\mathcal{L}_0$ is constructed in such a way that we can obtain sharp stability estimates using weighted estimates and sharp functional inequalities.

In Part I [13], we have performed the weighted energy estimates using a combination of weighted L^∞ and $C^{1/2}$ norm. In our analysis, we decompose $\widetilde{W} = \widetilde{W}_1 + \widetilde{W}_2$, where $\widetilde{W}_1$ is the main part of the perturbation, which is essentially governed by the leading order operator $\mathcal{L}_0$ with a weak coupling to $\widetilde{W}_2$ through nonlinear interaction. The perturbation $\widetilde{W}_2$ captures the contribution from the finite rank operator. The key is to show that the energy estimate of the main part $\widetilde{W}_1$ satisfies the inequalities stated in our stability Lemma 2.1 (see section 2). For this purpose, we need to obtain relatively sharp energy estimates for the leading order operator $\mathcal{L}_0$ by subtracting a finite rank operator $\mathcal{K}$. Without subtracting the finite rank operator, we would not be able to obtain the linear and nonlinear stability of the approximate self-similar profile.

The constants in the weighted energy estimates obtained in Part I [13] depend on the approximate self-similar profile that we constructed numerically in section 7 of Part I [13] and the singular weights we use. In this paper and in the supplementary material (supplement.pdf [local/web 1.43MB]), we will provide sharp and rigorous upper bounds for these constants by estimating the higher order derivatives and then using interpolation estimates from numerical analysis. We also obtain sharp estimates of the regular part of the velocity, which is more regular than the vorticity, by bounding various integrals using numerical integration with computer assistance. These sharp estimates of the constants enable us to prove that the inequalities in our stability lemma hold for our approximate self-similar profile. Thus we can complete the stability analysis of the approximate self-similar profile and complete our blowup analysis for the 2D Boussinesq and 3D Euler equations. See section 2.2 for more discussion of the main steps in our blowup analysis.

We use the following toy model to illustrate the main ideas of our stability analysis by considering $\mathcal{K}$ as a rank-one operator $\mathcal{K}(\widetilde{W}) = a(x)P(\widetilde{W})$ for some operator P satisfying (i) $P(\widetilde{W})$ is constant in space; (ii) $\|P(\widetilde{W})\| \le c\|\widetilde{W}\|$. Given initial data $\widetilde{W}_0$, we decompose (1.2) as follows:

$$\begin{aligned} &\partial_t \widetilde{W}_1(t) = \mathcal{L}_0 \widetilde{W}_1, \quad \widetilde{W}_1(0) = \widetilde{W}_0, \\ &\partial_t \widetilde{W}_2(t) = \mathcal{L}\widetilde{W}_2 + a(x)P(\widetilde{W}_1(t)), \quad \widetilde{W}_2(0) = 0. \end{aligned} \tag{1.3}$$

It is easy to see that $\widetilde{W} = \widetilde{W}_1 + \widetilde{W}_2$ solves (1.2) with initial data $\widetilde{W}_0$ since $\mathcal{L} = \mathcal{L}_0 + a(x)P$. By construction, the leading operator $\mathcal{L}_0$ has the desired structure that enables us to obtain sharp stability estimates. The second part $\widetilde{W}_2$ is driven by the rank-one forcing term $a(x)P(\widetilde{W}_1(t))$. Using Duhamel's principle and the fact that $P(\widetilde{W}_1(t))$ is constant in space, we get

$$\widetilde{W}_2(t) = \int_0^t P(\widetilde{W}_1(s)) e^{\mathcal{L}(t-s)} a(x) ds. \tag{1.4}$$

If $\widetilde{W}_1$ is linearly stable in some $L^\infty(\varphi)$ space, by checking the decay of $e^{\mathcal{L}(t)}a(x)$ in the energy space for large t, we can obtain the stability estimate of $\widetilde{W}_2$. Note that $e^{\mathcal{L}(t)}a(x)$ is equivalent to solving the linear evolution equation $v_t = \mathcal{L}(v)$ with initial data $v_0 = a(x)$. We can solve this initial value problem by constructing a space-time solution with rigorous error control.

We remark that our stability analysis is performed mainly for $\widetilde{W}_1$ since $\widetilde{W}_2$ is driven by $\widetilde{W}_1$. The approximation errors in constructing the space-time approximation to $\widetilde{W}_2$ can be controlled by the decay estimate of $\widetilde{W}_1$. Moreover, the region where we need to modify the linearized operator by a finite rank operator is mainly located in a small sector near the boundary where we have the smallest amount of damping. The total rank is less than 50. In our construction of an approximate solution to $\widetilde{W}_2$, we need to solve the linear PDE (1.2) in space-time with a number of initial data, which can be implemented in full parallel.

There has been a lot of effort in studying 3D Euler singularities. The most exciting recent development is Elgindi's breakthrough result in which he proved the finite time singularity of the axisymmetric Euler equation with no swirl for C^α initial vorticity [24] (see also [25]). In [12], we established the finite time blowup of the 2D Boussinesq and the 3D axisymmetric Euler equations with $C^{1,\alpha}$ velocity, large swirl, and boundary in a setting similar to the Hou–Luo scenario [41, 42]. See also [9] for further developments. Earlier efforts include the Constantin–Lax–Majda (CLM)

model [21], the De Gregorio model [22], the generalized CLM (gCLM) model [50], and the Hou–Li model [35]. See also [5, 6, 7, 8, 15, 21, 26, 28] for the De Gregorio model and for the gCLM model with various parameters. Inspired by their work on the vortex sheet singularity [4], Caflisch and Siegel have studied complex singularity for the 3D Euler equation; see [3, 54] and also [51] for the complex singularities for the 2D Euler equation.

In [17], the authors proved the blowup of the Hou–Luo model proposed in [42]. In [16], Chen, Hou, and Huang proved the asymptotically self-similar blowup of the Hou–Luo model by extending the method of analysis established for the finite time blowup of the De Gregorio model by the same authors in [15]. In [18, 31, 32, 33, 39], the authors proposed several simplified models to study the Hou–Luo blowup scenario [41, 42] and established finite time blowup of these models. In [27, 29], Elgindi and Jeong proved finite time blowup for the 2D Boussinesq and 3D axisymmetric Euler equations in a domain with a corner using $\mathring{C}^{0,\alpha}$ data.

The rest of the paper is organized as follows. In section 2, we review the analytic framework that we established in Part I [13] and state the key lemmas which we use to prove the finite time blowup of the 2D Boussinesq and 3D Euler equations with smooth initial data. In section 3, we discuss the construction of the approximate space-time solution to the linearized operator $\mathcal{L}$. This is crucial to obtain sharp estimates of the perturbed operator $\mathcal{L}-\mathcal{K}$ in the stability analysis. In section 4, we show how to estimate the L^∞ and Hölder norms of the regular part of the velocity. Some technical estimates and derivations are deferred to the appendix.

2. Review of the analytic framework from Part I [13]. In this section, we will review some main ingredients in our analytic framework to establish the stability analysis that we presented in Part I [13]. We will mainly focus on the 2D Boussinesq equations since the difference between the 3D Euler and 2D Boussinesq equations is asymptotically small. As in our previous works [12, 15, 16], we will use the dynamic rescaling formulation for the 2D Boussinesq equations to study the linear stability for the linearized operator around the approximate steady state of the dynamic rescaling equations. Passing from linear stability to nonlinear stability is relatively easier by treating the nonlinear terms and the residual error as small perturbations to the linear damping terms.

Denote by ω^θ, u^θ, and ϕ^θ the angular vorticity, angular velocity, and angular stream function, respectively. The 3D axisymmetric Euler equations are given below,

$$\begin{aligned} &\partial_t(ru^\theta)+u^r(ru^\theta)_r+u^z(ru^\theta)_z=0,\\ &\partial_t\left(\frac{\omega^\theta}{r}\right)+u^r\left(\frac{\omega^\theta}{r}\right)_r+u^z\left(\frac{\omega^\theta}{r}\right)_z=\frac{1}{r^4}\partial_z((ru^\theta)^2), \end{aligned} \tag{2.1}$$

where the radial velocity u^r and the axial velocity u^θ are given by the Biot–Savart law:

$$-\left(\partial_{rr}+\frac{1}{r}\partial_r+\partial_{zz}\right)\phi^\theta+\frac{1}{r^2}\phi^\theta=\omega^\theta,\quad u^r=-\phi^\theta_z,\quad u^z=\phi^\theta_r+\frac{1}{r}\phi^\theta, \tag{2.2}$$

with the no-flow boundary condition $\phi^\theta(1,z)=0$ on the solid boundary $r=1$ and a periodic boundary condition in z. For 3D Euler blowup that occurs at the boundary $r=1$, we know that the scaling properties of the axisymmetric Euler equations are asymptotically the same as those of the 2D Boussinesq equations [43]. Thus, we also study the 2D Boussinesq equations on the upper half space:

$$\omega_t + \mathbf{u}\cdot\nabla\omega = \theta_x, \tag{2.3}$$

$$\theta_t + \mathbf{u}\cdot\nabla\theta = 0, \tag{2.4}$$

where the velocity field $\mathbf{u} = (u,v)^T : \mathbb{R}^2_+ \times [0,T) \to \mathbb{R}^2_+$ is determined via the Biot–Savart law

$$-\Delta\phi = \omega, \quad u = -\phi_y, \quad v = \phi_x, \tag{2.5}$$

where ϕ is the stream function with the no-flow boundary condition $\phi(x,0) = 0$ at $y = 0$. By making the change of variables $\tilde{\theta} \triangleq (ru^\theta)^2, \tilde{\omega} = \omega^\theta / r$, we can see that $\tilde{\theta}$ and $\tilde{\omega}$ satisfy the 2D Boussinesq equations up to the leading order for $r \geq r_0 > 0$.

2.1. Dynamic rescaling formulation. Following [12, 15, 16], we consider the dynamic rescaling formulation of the 2D Boussinesq equations. Let $\omega(x,t), \theta(x,t)$, $\mathbf{u}(x,t)$ be the solutions of (2.3)–(2.5). Then it is easy to show that

$$\begin{aligned} &\tilde{\omega}(x,\tau) = C_\omega(\tau)\omega(C_l(\tau)x, t(\tau)), \quad \tilde{\theta}(x,\tau) = C_\theta(\tau)\theta(C_l(\tau)x, t(\tau)), \\ &\tilde{\mathbf{u}}(x,\tau) = C_\omega(\tau)C_l(\tau)^{-1}\mathbf{u}(C_l(\tau)x, t(\tau)) \end{aligned} \tag{2.6}$$

are the solutions to the dynamic rescaling equations

$$\tilde{\omega}_\tau(x,\tau) + (c_l(\tau)\mathbf{x} + \tilde{\mathbf{u}})\cdot\nabla\tilde{\omega} = c_\omega(\tau)\tilde{\omega} + \tilde{\theta}_x, \qquad \tilde{\theta}_\tau(x,\tau) + (c_l(\tau)\mathbf{x} + \tilde{\mathbf{u}})\cdot\nabla\tilde{\theta} = c_\theta\tilde{\theta}, \tag{2.7}$$

where $\tilde{\mathbf{u}} = (\tilde{u}, \tilde{v})^T = \nabla^\perp(-\Delta)^{-1}\tilde{\omega}$, $\mathbf{x} = (x,y)^T$,

$$C_\omega(\tau) = \exp\left(\int_0^\tau c_\omega(s)d\tau\right), \; C_l(\tau) = \exp\left(\int_0^\tau -c_l(s)ds\right), \; C_\theta = \exp\left(\int_0^\tau c_\theta(s)d\tau\right), \tag{2.8}$$

$t(\tau) = \int_0^\tau C_\omega(\tau)d\tau$ and the rescaling parameters $c_l(\tau), c_\theta(\tau), c_\omega(\tau)$ satisfy [12]

$$c_\theta(\tau) = c_l(\tau) + 2c_\omega(\tau). \tag{2.9}$$

To simplify our presentation, we still use t to denote the rescaled time in (2.7) and simplify $\tilde{\omega}, \tilde{\theta}$ as ω, θ

$$\omega_t + (c_l x + \mathbf{u})\cdot\nabla\omega = \theta_x + c_\omega\omega, \quad \theta_t + (c_l x + \mathbf{u})\cdot\nabla\theta = c_\theta\theta. \tag{2.10}$$

Following [16], we impose the following normalization conditions on c_ω, c_l:

$$c_l = 2\frac{\theta_{xx}(0)}{\omega_x(0)}, \quad c_\omega = \frac{1}{2}c_l + u_x(0), \quad c_\theta = c_l + 2c_\omega. \tag{2.11}$$

For smooth data, these two normalization conditions play the role of enforcing

$$\theta_{xx}(t,0) = \theta_{xx}(0,0), \quad \omega_x(t,0) = \omega_x(0,0) \tag{2.12}$$

for all time.

We remark that the dynamic rescaling formulation was introduced in [40, 45] to study the self-similar blowup of the nonlinear Schrödinger equations. This formulation is closely related to the modulation technique in the literature and has been developed by Merle, Raphael, Martel, Zaag, and others (see, e.g., [1, 2, 37, 44, 46, 47, 48]). Moreover, it is related to the method of modulation equations developed by Soffer and Weinstein [55, 56, 57]. Recently, this method has been applied to study singularity formation in incompressible fluids [12, 24] and related models [6, 7, 8, 15]. The more precise statement of our Theorem 1 is stated as follows.

THEOREM 3. *Let $(\bar\theta, \bar\omega, \bar{\mathbf{u}}, \bar c_l, \bar c_\omega)$ be the approximate self-similar profile constructed in section 7 of Part* I [13] *and $E_* = 5\cdot 10^{-6}$. For initial data $\theta_0(x,y)$ even in x and $\omega_0(x,y)$ odd in x of* (2.10) *satisfying $E(\omega_0-\bar\omega, \theta_{0,x}-\bar\theta_x, \theta_{0,y}-\bar\theta_y) < E_*$, we have*

$$
\begin{aligned}
&||\omega-\bar\omega||_{L^\infty},\ ||\theta_x-\bar\theta_x||_{L^\infty},\ ||\theta_y-\bar\theta_y||_\infty < 200E_*,\\
&|u_x(t,0)-\bar u_x(0)|,\ |\bar c_\omega - c_\omega| < 100E_*
\end{aligned}
\tag{2.13}
$$

for all time. In particular, we can choose smooth initial data $\omega_0, \theta_0 \in C_c^\infty$ in this class with finite energy $||\mathbf{u}_0||_{L^2} < +\infty$ such that the solution to the physical equations (2.3)–(2.5) *with these initial data blows up in finite time T.*

The energy E is quite complicated, and we refer to section 2.3 in Part I [13] for its formula.

Nearly self-similar blowup and the blowup time. Based on the main theorem, Theorem 3, the vorticity in the physical space ((2.3), (2.4)) ω_{phy} has the following form:

$$
\omega_{phy}(x, t(\tau)) = C_\omega^{-1}(\tau)\omega_{ss}(C_l(\tau)^{-1}x, \tau), \quad ||\omega_{ss}(\tau)-\bar\omega||_{L^\infty} \ll 1,
$$

where ω_{ss} is the self-similar variable ($\tilde\omega$ in (2.6)). We can generalize the rescaling parameters C_ω, C_l, C_θ (2.8) to $C_\omega(\tau) = C_\omega(0)\exp(\int_0^\tau c_\omega(s)ds), C_l(\tau) = C_l(0)\exp(\int_0^\tau -c_l(s)ds)$, $C_\theta = C_\omega^2 C_l^{-1}, t(\tau) = \int_0^\tau C_\omega(\tau)d\tau$. Using the estimates (2.13) and $c_l(\tau)\equiv \bar c_l$ (2.11), (2.12), we obtain

$$
\begin{aligned}
&||\omega_{phy}(\tau)||_{L^\infty} \approx C_\omega^{-1}(\tau)||\bar\omega||_{L^\infty}, \quad C_\omega(\tau)\approx C_\omega(0)e^{\bar c_\omega \tau}, \quad C_l(\tau) = C_l(0)e^{-\bar c_l\tau},\\
&T = t(\infty) \approx C_\omega(0)|\bar c_\omega|^{-1} = \frac{||\bar\omega||_{L^\infty}}{||\omega_{phy}(0)||_{L^\infty}|\bar c_\omega|}, \quad T - t(\tau) \approx C_\omega(\tau)|\bar c_\omega|^{-1},\\
&C_l(\tau)^{-1} \approx C(T-t(\tau))^{\bar c_l/\bar c_\omega}, \quad \omega_{phy}(\tau) \approx (T-t(\tau))^{-1}|\bar c_\omega|^{-1}\bar\omega(Cx(T-t(\tau))^{\bar c_l/\bar c_\omega})
\end{aligned}
$$

with $\bar c_l/\bar c_\omega \approx -2.92 < 0$, for some $C>0$ depending on $\bar c_l, \bar c_\omega, C_l(0), C_\omega(0)$. The notation $\approx$ means that the relation holds approximately. The exact relation can be inferred from (2.13), (2.6), (2.8). The blowup time is approximately inversely proportional to $||\omega_{phy}(0)||_{L^\infty}$. Since we only prove that $\omega_{ss}(\tau)$ is sufficiently close to the approximate profile $\bar\omega$ and do not prove the convergence of $\omega_{ss}(\tau)$ as $\tau\to\infty$, Theorem 3 does not imply an *asymptotically* self-similar blowup.

2.2. The main steps in the proof of Theorem 3. We will follow the framework in [12, 15, 16] to establish finite time blowup by proving the nonlinear stability of an approximate steady state to (2.10). We divide the proof of Theorem 3 into proving the following lemmas. The energy norm below is defined in section 5 in Part I [13] for energy estimates, and the requirement of smallness is incorporated into the conditions (2.17), e.g., the term $a_{ij,3}$, in Lemma 2.5.

The upper bar notation is reserved for the approximate steady state, e.g., $\bar\omega, \bar\theta$. Given the approximate steady state $\bar\omega, \bar\theta, \bar c_l, \bar c_\omega$, we denote by $\overline{\mathcal{F}}_i$ and $\bar F_\omega, \bar F_\theta$ the residual error

$$
\begin{aligned}
&\bar F_\omega = -(\bar c_l x + \bar{\mathbf{u}})\cdot\nabla\bar\omega + \bar\theta_x + \bar c_\omega\bar\omega, \quad \bar F_\theta = -(\bar c_l x+\bar{\mathbf{u}})\cdot\nabla\bar\theta + \bar c_\theta\bar\theta,\\
&\overline{\mathcal{F}}_1 \triangleq \bar F_\omega, \quad \overline{\mathcal{F}}_2 \triangleq \partial_x\bar F_\theta, \quad \overline{\mathcal{F}}_3 \triangleq \partial_y\bar F_\theta, \quad \bar c_\theta = \bar c_l + 2\bar c_\omega.
\end{aligned}
\tag{2.14}
$$

We have the following nonlinear stability lemma for an L^∞-based energy estimate, which is proved in Appendix A.1 of Part I [13].

LEMMA 2.1. *Suppose that* $f_i(x,z,t):\mathbb{R}^2_{++}\times\mathbb{R}^2_{++}\times[0,T]\to\mathbb{R}, 1\le i\le n$, *satisfies*

$$\partial_t f_i + v_i(x,z)\cdot\nabla_{x,z} f_i = -a_{ii}(x,z,t)f_i + B_i(x,z,t) + N_i(x,z,t) + \bar{\varepsilon}_i, \tag{2.15}$$

where $v_i(x,z,t)$ *are some vector fields Lipschitz in* x,z *with* $v_i|_{x_1=0}=0, v_i|_{z_1=0}=0$. *For some* $\mu_i>0$, *we define the energy*

$$E(t)=\max_{1\le i\le n}(\mu_i||f_i||_{L^\infty}).$$

Suppose that B_i, N_i, *and* $\bar{\varepsilon}_i$ *satisfy the following estimate:*

$$\begin{aligned}&\mu_i(|B_i(x,z,t)|+|N_i(x,z,t)|+|\bar{e}_i|)\\ &\quad\le\sum_{j\ne i}(|a_{ij}(x,z,t)|E(t)+|a_{ij,2}(x,z,t)|E^2(t)+|a_{ij,3}(x,z,t)|).\end{aligned} \tag{2.16}$$

If there exists some $E_*,\varepsilon_0,M>0$ *such that*

$$\begin{aligned}&a_{ii}(x,z,t)E_*-\sum_{j\ne i}(|a_{ij}|E_*+|a_{ij,2}|E_*^2+|a_{ij,3}(x,z,t)|)>\varepsilon_0,\\ &\sum_{j\ne i}(|a_{ij}|E_*+|a_{ij,2}|E_*^2+|a_{ij,3}(x,z,t)|)<M\end{aligned} \tag{2.17}$$

for all x,z, *and* $t\in[0,T]$. *Then for* $E(0)<E_*$, *we have* $E(t)<E_*$ *for* $t\in[0,T]$.

LEMMA 2.2. *There exists a nontrivial approximate steady state* $(\bar{\omega},\bar{\theta},\bar{c}_l,\bar{c}_\omega)$ *to* (2.10), (2.11) *with* $\bar{\omega},\bar{\theta}\in C^{4,1}$ *and residual errors* $\bar{\mathcal{F}}_i, i=1,2,3$ (2.14) *sufficiently small in some energy norm.*

The construction of an approximate self-similar profile with a small residual error stated in Lemma 2.2 is provided in section 7 of Part I [13] and the properties of $(\bar{\omega},\bar{\theta},\bar{c}_l,\bar{c}_\omega)$ are described in section 2.4 of Part I [13]. We will estimate the local part of the residual error in Appendix C.4. We linearize (2.10) around $(\bar{\omega},\bar{\theta},\bar{c}_l,\bar{c}_\omega)$ and perform an energy estimate of the perturbation $W=(\omega,\theta_x,\theta_y)$ in section 5 in Part I [13]. In our estimates, we need to control a number of nonlocal terms.

LEMMA 2.3. *Let* ω *be odd in* x_1. *Denote* $\delta(f,x,z)=f(x)-f(z)$. *There exists finite rank approximations* $\hat{\mathbf{u}},\widehat{\nabla\mathbf{u}}$ *for* $\mathbf{u}(\omega),\nabla\mathbf{u}(\omega)$ *with rank less than* 50 *such that we have the following weighted* L^∞ *and directional Hölder estimate for* $f=u,v,\partial_l u,\partial_l v, x,z\in\mathbb{R}^{++}_2, i=1,2,\gamma_i>0$:

$$\begin{aligned}&|\rho_f(f-\hat{f})(x)|\le C_{f,\infty}(x,\varphi,\psi_1,\gamma)\max\Big(||\omega\varphi||_\infty, s_f\max_{j=1,2}\gamma_j[\omega\psi_1]_{C^{1/2}_{x_j}(\mathbb{R}^+_2)}\Big),\\ &\frac{|\delta(\psi_f(f-\hat{f}),x,z)|}{|x-z|^{1/2}}\le C_{f,i}(x,z,\varphi,\psi_1,\gamma)\max\Big(||\omega\varphi||_\infty, s_f\max_{j=1,2}\gamma_j[\omega\psi_1]_{C^{1/2}_{x_j}(\mathbb{R}^+_2)}\Big)\end{aligned} \tag{2.18}$$

with $x_{3-i}=z_{3-i}$, *where* $s_f=0$ *for* $f=u,v$, $s_f=1$ *for* $f=\partial_l u,\partial_l v$, *the functions* $C(x),C(x,z)$ *depend on* γ, *the weights, and the approximations, the singular weights* $\varphi=\varphi_1,\varphi_{g,1},\varphi_{elli},\psi_{\partial u}=\psi_1,\psi_u$ *are defined in* (A.2), *and the weight* ρ_{10} *for* $\mathbf{u}$ *and the weight for* ρ_{ij} *for* $\nabla\mathbf{u}$ *with* $i+j=2$ *are given in* (A.2). *In the estimate of* $f=u,v$, *we do not need the Hölder seminorm and we set* $s_f=0$. *Moreover,* $C(x),C(x,z)$ *are bounded in any compact domain of* $\mathbb{R}^{++}_2$. *We have an additional estimate for* $\rho_4(u-\hat{u})$ *similar to the above with* ρ_4 (A.2) *singular along* $x_1=0$.

Furthermore, we have the following estimate using the localized norm. There exist $D_1, D_2, ..D_n \subset \mathbb{R}_2^{++}$ *and* $D_S \in \mathbb{R}_2^{+}$ *depending on* x *in the* L^∞ *estimate and* x, z *in the* $C_{x_i}^{1/2}$ *estimate, such that*

$$\begin{aligned}|\rho_f(f-\hat{f})(x)| \leq & \sum_j C_{f,\infty,j}(x,\varphi,\psi_1,\gamma)||\omega\varphi||_{L^\infty(D_j)} \\ & + C_{f,\infty,S}(x,\varphi,\psi_1,\gamma)\max_{l=1,2}\left(\gamma_l[\omega\psi_1]_{C_{x_l}^{1/2}(D_S)}\right), \\ \frac{|\delta(\psi_f(f-\hat{f}),x,z)|}{|x-z|^{1/2}} \leq & \sum_j C_{f,i,j}(x,z,\varphi,\psi_1,\gamma)||\omega\varphi||_{L^\infty(D_j)} \\ & + C_{f,i,S}(x,z,\varphi,\psi_1,\gamma)\max_{l=1,2}\left(\gamma_l[\omega\psi_1]_{C_{x_j}^{1/2}(D_S)}\right)\end{aligned}$$

for $x_{3-i} = z_{3-i}, \varphi = \varphi_{elli}$ *and the same notation as above, where* $C_{f,\infty,S}, C_{f,i,S} = 0$ *for* $f = u, v$*. Similarly, we have an estimate for* $\rho_4(u-\hat{u})$ *using a localized norm with* $C_{f,\infty,S} = 0$ *similar to the above.*

Since the weights $\rho_{10} \sim |x|^{-3}, \psi_1 \sim |x|^{-2}, \psi_u$ are singular near $x = 0$, without subtracting the approximation $\hat{f}$ from f, $\rho_f f$ is not bounded near $x = 0$. We design the finite rank approximations $\hat{\mathbf{u}}, \widehat{\nabla \mathbf{u}}$ in section 4.3 in Part I [13].

Based on these finite rank approximations, we can decompose the perturbations.

LEMMA 2.4. *There exist* $m < 50$ *approximate solutions* $\hat{F}_i$ *to the linearized equations* $\partial_t W = \mathcal{L}W$ *of* (2.10) *around* $(\bar{\omega}, \bar{\theta}, \bar{c}_l, \bar{c}_\omega)$ *in Lemma* 2.2 *from given initial data* $\bar{F}_i(0)$ *with residual error* $\mathcal{R}$ *small in the energy norm. Further we can decompose the perturbation* $W = W_1 + \widehat{W}_2$ *with the following properties.* (a) $\hat{W}_2$ *is constructed based on* $\widehat{F}_i$*; see section* 4.2.4 *of Part* I [13]. (b) W_1 *satisfies the equations with the leading order linearized operator* $(\mathcal{L} - \mathcal{K})W_1$ *up to the small residual error* $\mathcal{R}$ *for some finite rank operator* $\mathcal{K}$*, and* W_1 *depends on* $\widehat{W}_2$ *weakly at the linear level via* $\mathcal{R}$*. The functionals* $a_i(W_1), a_{nl,i}(W)$ *in the construction of* $\widehat{W}_2$ *and* $\mathcal{K}$ *(see section* 4.2.4 *of Part* I [13]*) are related to the finite rank approximations in Lemma* 2.3*.*

Moreover, there exists an energy $E_4(t)$ *for* W_1, W *(see section* 5.6.3. *of Part* I [13]*) that controls the weighted* L^∞ *and* $C^{1/2}$ *seminorm of* W_1 *such that under the bootstrap assumption* $E_4(t) < E_{*0}$ *with* $E_{*0} > 0$*, we can establish nonlinear energy estimates for* $E_4(t)$ *using the estimates in Lemma* 2.3*.*

If the bounds in Lemma 2.3 are tight, and the residual error in the constructions of $(\bar{\omega}, \bar{\theta}), \widehat{F}_i$ are small enough, we can use Lemma 2.1 to obtain nonlinear stability.

LEMMA 2.5. *For* $E_* = 5 \cdot 10^{-6}$*, the coefficients in the nonlinear energy estimates of* $E_4(t)$ *satisfy the conditions* (2.17)*, and the statements in Theorem* 3 *hold true.*

The main purpose of Part II of our paper is the following. First, we construct the approximate $\hat{F}_i(t)$ in Lemma 2.4 numerically, and estimate its piecewise derivatives and the local residual error in section 3. Second, in section 4, we obtain sharp estimates of the constants in Lemma 2.3, which only depend on the weights. Third, we estimate piecewise bounds of the approximate steady state in Appendix C, the singular weights in Appendix A, and some explicit functions related to the approximate solutions in Appendix D. We remark that all of these estimates and constants depend on the given weights, some operators, and functions, e.g., the approximate steady state and the specific initial conditions. With these estimates and constants, we obtain the concrete values of the inequalities in (2.17) and Lemma 2.5, which are

given in Appendix D in Part I [13]. We further verify the inequalities for the stability conditions in Lemma 2.5.

Let us make a few comments on the above lemmas. First, our energy estimate is based on weighted L^∞ functional spaces, which is crucial for extracting the damping terms for the energy estimate. See section 2.7 of Part I [13] for the motivations. Given $\omega \in C^{1/2}$, we have $\mathbf{u} \in C^{3/2}, \nabla \mathbf{u} \in C^{1/2}$. To establish the nonlinear stability conditions (2.17) in Lemma 2.5, we need sharp constants in the estimates in Lemma 2.3. We use some techniques from optimal transport to obtain a sharp $C^{1/2}$ estimate of $\nabla \mathbf{u}$ in section 3 of Part I [13]. This corresponds to the limiting case in the $C^{1/2}_{x_i}$ estimate in Lemma 2.3 for a fixed x with $|x - z| \to 0$ and captures the most singular part in the estimates in Lemma 2.3. The constants in the sharp $C^{1/2}$ estimate established in Part I [13] are given by several integrals. In section 5 of the supplementary material (supplement.pdf [local/web 1.43MB]), we estimate these integrals.

Other parts of the estimates in Lemma 2.3 are more regular since we work with the regular part of the velocity integral with a desingularized kernel. Given $\omega \in C^{1/2}$, we can reduce the estimates of these more regular terms to estimate some explicit L^1 integrals. We can obtain sharp estimates of these more regular integrals using some numerical quadrature with computer assistance. See section 4.

By designing $\mathcal{K}$ to approximate the nonlocal terms, we can obtain much better linear stability estimates for $\mathcal{L} - \mathcal{K}$. After we have shown that the stability conditions (2.17) are satisfied, we have nonlinear stability estimate $E_4(t) < E_*$ for all $t > 0$ using Lemma 2.1, which implies the bounds in Theorem 3. The remaining steps of obtaining finite time blowup from smooth initial data and finite energy follow [15] and a rescaling argument. We remark that the variable $\widehat{W}_2$ in Lemma 2.4 (see the full definition in section 4.2.4 of Part I [13]) plays an auxiliary role, and we do not perform an energy estimate on $\widehat{W}_2$ directly.

Note that all the nonlocal terms in the linearized equations are not small. Without the sharp $C^{1/2}$ estimate, with the choice of energy E_4, the stability conditions in (2.17) and Lemma 2.5 fail in the weighted Hölder estimate. Without the finite rank approximations for the nonlocal terms in Lemmas 2.3, 2.4, the stability conditions for weighted L^∞ estimate also fail.

Rigorous numerics. We need to track two types of errors for rigorous numerics. The first type is the discretization error, e.g., the error terms in the trapezoidal rule and in the interpolating polynomials. The second type is the round-off error in the computation. We use numerical analysis to estimate *all* the discretization errors and use *only* the basic interval arithmetic [49, 53] (see, e.g., (A.4), (A.5)) in the INTLAB package [52] from MATLAB to track the round-off error.

In our nonlinear estimates, we use a singular weight φ like $|x|^{-3}$ near $x = 0$ to measure the residual error $\bar{\mathcal{F}}_i$. To obtain a small weighted residual error $|\varphi \bar{\mathcal{F}}_i|$ near $x = 0$, we choose the mesh y_i (C.2) representing the approximate profile to be exact floating point numbers to reduce the round-off error near $x = 0$.

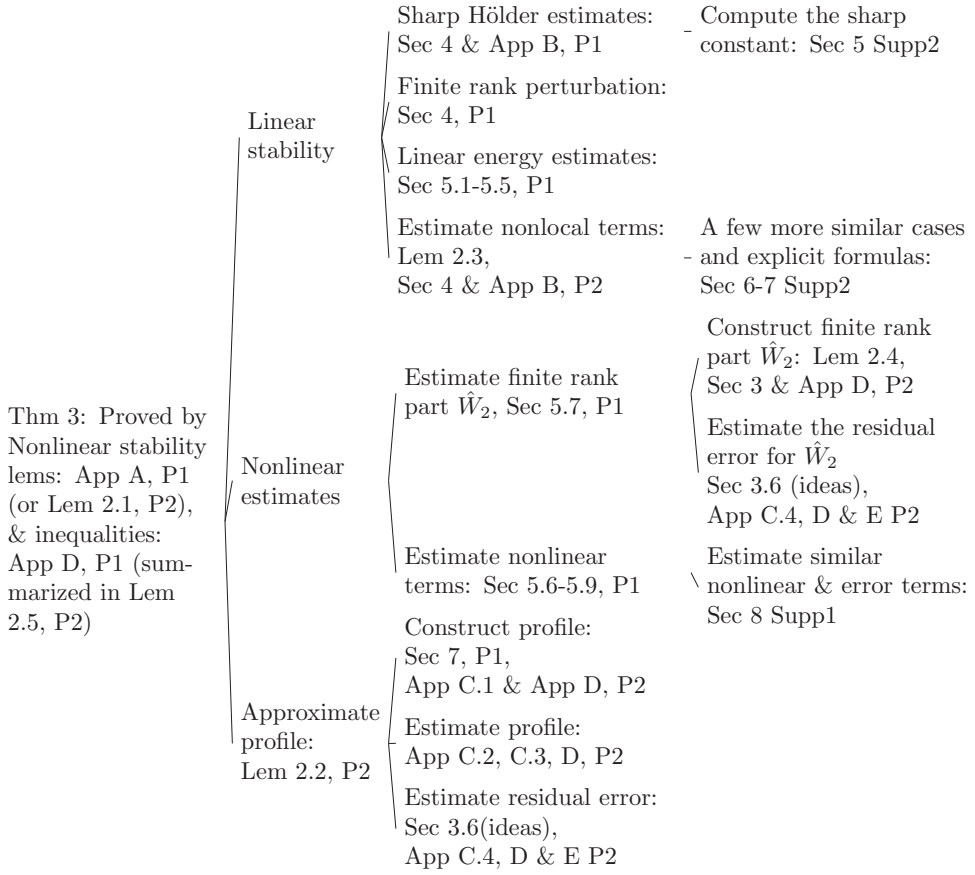
The codes for the computations are implemented in MATLAB and can be found in [10]. The estimates of the constants in Lemma 2.3, integrals in section 4, and constructions and estimates of the approximate space-time solutions in Lemma 2.4 and in section 3 are performed in parallel using the Caltech high Performance Computing.[1] Other computer-assisted estimates and the main part of the verifications are done on

[1] See more details for Caltech HPC Resources: https://www.hpc.caltech.edu/resources.

a Mac Pro (Rack,2019) with 2.5 GHz 28-core Intel Xeon W processor and 768 GB (6x128 GB) of DDR4 ECC memory.

2.3. Dependency tree. The following tree schematizes various intermediate steps and related sections that lead to the main stability result, Theorem 3, which implies the blowup result, Theorem 1, for the 2D Boussinesq equations. The blowup for the Euler equations in Theorem 2 is proved by a perturbative argument in section 6 in Part I [13].

Below, *Thm, Lem, App, Sec, P*1, *P*2, *Supp*1, *Supp*2 are short for theorem, lemma, appendix, section, paper I [13], paper II (the current paper), and the supplementary material for paper I [14] and this paper (supplement.pdf [local/web 1.43MB]), respectively. We present a few more detailed derivations in the supplementary materials [14], (supplement.pdf [local/web 1.43MB]), which expand and generalize discussions in the main papers and are less essential. Moreover, we present several explicit formulas we used in our computer-assisted estimates for the quantities derived in the main papers.



In section 8 in Supp2, we generalize the standard interpolation estimate in numerical analysis to derive higher order interpolation estimates, which are used to estimate the residual error effectively. See the discussion in section 3.6. In Appendix A, we derive piecewise bounds for various weights, which are used in the weighted estimates of the nonlocal terms, the residual error, and the linear, nonlinear estimates for stability.

In Appendix F, we collect the main notation used in this paper.

3. Construct and estimate the approximate solution to the linearized equations. As we described in section 2 of Part I [13] (see also the Introduction), we need to construct the approximate solutions to $e^{\mathcal{L}t}F_0$ for several initial data $\bar{F}_i, \bar{F}_{\chi,i}$. In this section, we discuss how to construct these space-time solutions numerically with some vanishing properties at the origin with rigorous error control.

The linearized equations associated with $\mathcal{L}$ read

(3.1)
$$\begin{aligned}
&\partial_t\omega = -(\bar{c}_l x + \bar{u})\cdot\nabla\omega + \eta + \bar{c}_\omega\omega - \mathbf{u}\cdot\nabla\bar{\omega} + c_\omega\bar{\omega} = \mathcal{L}_1(\omega,\eta,\xi),\\
&\partial_t\eta = -(\bar{c}_l x + \bar{u})\cdot\nabla\eta + (2\bar{c}_\omega - \bar{u}_x)\eta - \bar{v}_x\xi - \mathbf{u}_x\cdot\nabla\bar{\theta} - \mathbf{u}\cdot\nabla\bar{\theta}_x + 2c_\omega\bar{\theta}_x = \mathcal{L}_2(\omega,\eta,\xi),\\
&\partial_t\xi = -(\bar{c}_l x + \bar{u})\cdot\nabla\xi + (2\bar{c}_\omega + \bar{u}_x)\xi - \bar{u}_y\eta - \mathbf{u}_y\cdot\nabla\bar{\theta} - \mathbf{u}\cdot\nabla\bar{\theta}_y + 2c_\omega\bar{\theta}_y = \mathcal{L}_3(\omega,\eta,\xi)
\end{aligned}$$

with normalization condition

$$c_\omega = u_x(0), \quad c_l \equiv 0. \tag{3.2}$$

Although η,ξ represent θ_x,θ_y in the Boussinesq equations, we will consider initial data (ω_0,η_0,ξ_0) with $\partial_y\eta_0 \neq \partial_x\xi_0$. Thus, we do not have the relation $\partial_y\eta = \partial_x\xi$ and will treat η,ξ as two independent variables. The solutions $\omega(x),\eta(x)$ are odd in x_1, and $\xi(x)$ is even in x_1 with $\xi(0,y)=0$. We consider initial data $(\omega_0,\eta_0,\xi_0) = O(|x|^2)$ near $x=0$. Using a direct calculation, we can show that these vanishing conditions are preserved in time

$$\omega(t,x),\ \eta(t,x),\ \xi(t,x) = O(|x|^2). \tag{3.3}$$

We introduce the bilinear operator $B_{op,i}((\mathbf{u},M),G)$ for $(\mathbf{u},M), G = (G_1,G_2,G_3)$

(3.4)
$$\begin{aligned}
&\mathcal{B}_{op,1} = -\mathbf{u}\cdot\nabla G_1 + M_{11}(0)G_1, \quad \mathcal{B}_{op,2} = -\mathbf{u}\cdot\nabla G_2 + 2M_{11}(0)G_2 - M_{11}G_2 - M_{21}G_3,\\
&\mathcal{B}_{op,3} = -\mathbf{u}\cdot\nabla G_3 + 2M_{11}(0)G_3 - M_{12}G_2 - M_{22}G_3.
\end{aligned}$$

If $M = \nabla\mathbf{u}, M_{11} = u_x, M_{12} = u_y, M_{21} = v_x, M_{22} = v_y$, then we drop M to simplify the notation

(3.5)
$$\begin{aligned}
&\mathcal{B}_{op,1}(\mathbf{u},G) = -\mathbf{u}\cdot\nabla G_1 + u_x(0)G_1, \quad \mathcal{B}_{op,2} = -\mathbf{u}\cdot\nabla G_2 + 2u_x(0)G_2 - u_xG_2 - v_xG_3,\\
&\mathcal{B}_{op,3} = -\mathbf{u}\cdot\nabla G_3 + 2u_x(0)G_3 - u_yG_2 - v_yG_3.
\end{aligned}$$

The main result in this section is the following. Given n initial data $\bar{G}_i = (\bar{G}_{i,1},\bar{G}_{i,2},\bar{G}_{i,3})$ and n functions $c_i(t)(i=1,2,..,n)$ which are Lipschitz and bounded in t, we construct an approximate space-time solution $\hat{W}_i = (\hat{W}_{i,1},\hat{W}_{i,2},\hat{W}_{i,3}), \hat{G} = (\hat{G}_1,\hat{G}_2,\hat{G}_3)$ and the approximate stream functions $\hat{\phi}_i^N, \hat{\phi}^N$ and the error $\hat{\varepsilon}$ associated with $\hat{W}_{i,1}, \hat{G}_1$

$$\begin{aligned}
&\hat{G} = \sum_{i\le n}\int c_i(t-s)\hat{W}_i(s)ds, \quad \hat{\phi}^N = \sum_{i\le n}\int c_i(t-s)\hat{\phi}_i^N(s)ds,\\
&\hat{\varepsilon} = \sum_{i\le n}\int c_i(t-s)(\hat{W}_{i,1} + \Delta\hat{\phi}_i^N)(s)ds
\end{aligned} \tag{3.6}$$

with residual error

$$\mathcal{R} = \sum_{i\le n} c_i(t)(\hat{W}_i(0) - \bar{W}_i) + \int_0^t c(t-s)(\partial_t - \mathcal{L})\hat{W}_i(s)ds, \tag{3.7}$$

vanishing $O(|x|^3)$ near $x=0$. Moreover, we can decompose $\mathcal{R}=(\mathcal{R}_1,\mathcal{R}_2,\mathcal{R}_3)$ as follows:

$$\begin{aligned}(3.8)\quad &\mathcal{R}_j(t)=\mathcal{R}_{loc,0,j}(t)+\mathcal{R}_{nloc,j}(t),\\ &\mathcal{R}_{loc,0,j}=\sum_{i\le n}\int_0^t c_i(t-s)\mathcal{R}_{num,i,j}(s)ds, \mathcal{R}_{num,i,j}=O(|x|^3),\\ &\mathcal{R}_{nloc,j}=P_j-D_j^2P_j(0)\chi_{j,2},\quad P_j=-\mathcal{B}_{op,j}(\mathbf{u}(\bar{\varepsilon}),\hat{G})-\mathcal{B}_{op,j}(\mathbf{u}(\hat{\varepsilon}),(\bar{\omega},\bar{\theta}_x,\bar{\theta}_y),\end{aligned}$$

where χ_{j2} is given in (D.5), and $\bar{\varepsilon}=\bar{\omega}-(-\Delta)\bar{\phi}^N$ is the error of the approximate stream function for $(-\Delta)^{-1}\bar{\omega}$, and $\mathcal{R}_{num,j}(t,x)$ depends on $\hat{W}_i,\hat{\phi}_i$ in x locally. We have absorbed the initial error in $\mathcal{R}_{num}$. We derive the above decompositions and estimates of $\mathcal{R}_{loc,0,j},\hat{G},\hat{\phi}^N,\hat{\varepsilon}$, in sections 3.5–3.7. See (3.37), (3.35). The error in constructing the stream function $\hat{\phi}_i^N$ associated with $\hat{W}_{i,1}$ leads to a nonlocal error, e.g., $\mathbf{u}(\hat{\varepsilon})$, in constructing the velocity. We combine the estimate of the nonlocal error in P_j and perturbation in section 5.8 in Part I [13]. Furthermore, we track the piecewise bounds of the following quantities:

$$\begin{aligned}(3.9)\quad &\int_0^\infty |\partial_x^k\partial_y^l F(t)|dt,\ F=\hat{W}_{i,j},\quad F=\hat{\phi}_i^N,\ F=\hat{\phi}_i^N-\partial_{xy}\phi_i^N(0)xy,\ F=\hat{W}_{i,1}+\Delta\phi_i^N,\\ &F=c_j\hat{W}_{i,j}-x\partial_x\hat{W}_{i,j}+y\partial_y\hat{W}_{i,j}-D_j^2\hat{W}_{i,j}(0)f_{\chi,j},\ D^2=(\partial_{xy},\partial_{xy},\partial_{xx}),c=(1,1,3)\end{aligned}$$

for $j=1,2,3,i=1,2,..,n$, where $f_{\chi,j}$ is defined in (D.6). We track the C^2 bound of $\hat{W}_{i,j}$ and C^4 bounds for others following (3.37) and use these bounds to control $\widehat{W}_2$ in Lemma 2.4 and use them in the nonlinear energy estimates in section 5 in Part I [13].

In practice, we choose the initial data $\bar{F}_i$ given in Appendix C.2.1 in Part I [13] and $c_i(t)$ some functionals of the perturbation $W_1,\hat{W}_2$ related to the finite rank perturbation.

Numerical methods. We solve (3.1) using the numerical method outlined in section 7 of Part I [13] to obtain the solution (ω_k,η_k,ξ_k) at discrete time t_k. Since ξ is even with $\xi(0,y)=0$, we write $\xi=x\zeta$ for an odd function ζ. We use the adaptive mesh discussed in Appendix C.1 to discretize the spatial domain. Then we represent ω,η,ζ using the piecewise sixth order B-spline (C.6). See Appendix C.1. To solve the stream function $-\Delta\phi=\omega$ numerically, we use the B-spline based finite element method and obtain the numerical approximation ϕ^N for $(-\Delta)^{-1}\omega$. Then we can construct the velocity $\mathbf{u}^N=\nabla^\perp\phi^N$.

The gradients of several initial conditions $\bar{F}_i$ are relatively large and the linearized equations (3.1) involve $\nabla\hat{W}$. To obtain a better approximation of the solution, we represent ω,η,ζ using a mesh $Y\times Y$ with Y refining the mesh y (C.2) in Appendix C.1 by a factor of three:

$$Y_{3i+j}=y_i+(y_{i+1}-y_i)j/3,\quad 0\le j\le 3.$$

Since solving the Poisson equation is the main computational cost in each time step, we still represent ϕ^N using the coarse mesh $y\times y$ and solve it from a source term with the grid points value $\omega(y_i,y_j)$.

In the temporal variable, we use a third order Runge–Kutta method to update the PDE. To reduce the round-off error near $x=0$, where we require a very small

error in solving the linear PDE, we use a multilevel representation. We defer more details to section 7 in Part I [13]. To keep the residual error smooth near $x = 0$, we apply a weak numerical filter near $x = 0$ every three steps. We do not add the semianalytic part in constructing $(\omega_k, \eta_k, \xi_k)$ for efficiency consideration and that the far-field behavior of the solutions is changing over time.

After we obtain the numerical solution $(\omega_k, \eta_k, \xi_k, \phi^N_{k,1})$ at discrete time, we will perform two rank-one corrections and interpolate the solution in time using a cubic polynomial to obtain the approximate space-time solution $\hat{W}$ and estimate residual error in the energy space a posteriori.

3.1. A posteriori error estimates: Decomposition of errors. Since we cannot solve the Poisson equation exactly, we decompose the stream function $\bar{\phi}, \phi$ as follows:

$$\bar{\phi} = (-\Delta)^{-1}\bar{\omega} = \bar{\phi}^N + \bar{\phi}^e, \quad \phi = (-\Delta)^{-1}\omega = \phi^N + \phi^e, \tag{3.10}$$

where $\bar{\phi}^N, \phi^N$ constructed using the finite element method are the numeric approximation of the stream function, and the shorthand N, e denote *numeric, error,* respectively. We use similar notation below for other nonlocal terms since we cannot construct them exactly. We will construct $\bar{\phi}^N, \phi^N$ numerically and treat $\bar{\phi}^e, \phi^e$ as error. The reader should not confuse ϕ^N with the Nth power of ϕ. We will never use the power of ϕ throughout the paper. Similarly, we denote by $\mathbf{u}^N, \mathbf{u}^e$ the velocities corresponding to ϕ^N, ϕ^e. For example, we have

$$\mathbf{u}^N = \nabla^\perp \phi^N, \quad \mathbf{u}^e = \nabla^\perp \phi^e = \nabla^\perp (-\Delta)^{-1}(\omega - (-\Delta)\phi^N), \quad c^N_\omega = u^N_x(0), \quad c^e_\omega = u^e_x(0). \tag{3.11}$$

The above decomposition leads to the following decomposition of the operator $\mathcal{L}$:

$$\begin{aligned}
&\mathcal{L}_1 = \mathcal{L}^N_1 + \mathcal{L}^e_1 + \mathcal{L}^{\bar{e}}_1, \quad \mathcal{L}_2 = \mathcal{L}^N_2 + \mathcal{L}^e_2 + \mathcal{L}^{\bar{e}}_2, \quad \mathcal{L}_3 = \mathcal{L}^N_3 + \mathcal{L}^e_3 + \mathcal{L}^{\bar{e}}_3,\\
&\mathcal{L}^N_1 = \eta + \bar{c}^N_\omega \omega - (\bar{c}_l x + \bar{\mathbf{u}}^N)\cdot\nabla\omega + c^N_\omega \bar{\omega} - \mathbf{u}^N \cdot \nabla\bar{\omega},\\
&\mathcal{L}^e_1 = c^e_\omega \bar{\omega} - \mathbf{u}^e \cdot \nabla\bar{\omega}, \quad \mathcal{L}^{\bar{e}}_1 = \bar{c}^e_\omega \omega - \bar{\mathbf{u}}^e \cdot \nabla\omega,\\
&\mathcal{L}^N_2 = -(\bar{c}_l x + \bar{\mathbf{u}}^N)\cdot\nabla\eta + (2\bar{c}^N_\omega - \bar{u}^N_x)\eta - \bar{v}^N_x \xi - \mathbf{u}^N_x \cdot \nabla\bar{\theta} - \mathbf{u}^N \cdot \nabla\bar{\theta}_x + 2c^N_\omega \bar{\theta}_x,\\
&\mathcal{L}^e_2 = -\mathbf{u}^e_x \cdot \nabla\bar{\theta} - \mathbf{u}^e \cdot \nabla\bar{\theta}_x + 2c^e_\omega \bar{\theta}_x, \quad \mathcal{L}^{\bar{e}}_2 = -\bar{\mathbf{u}}^e \cdot \nabla\eta + (2\bar{c}^e_\omega - \bar{u}^e_x)\eta - \bar{v}^e_x \xi,\\
&\mathcal{L}^N_3 = -(\bar{c}_l x + \bar{\mathbf{u}}^N)\cdot\nabla\xi + (2\bar{c}^N_\omega - \bar{v}^N_y)\xi - \bar{u}^N_y \eta - \mathbf{u}^N_y \cdot \nabla\bar{\theta} - \mathbf{u}^N \cdot \nabla\bar{\theta}_y + 2c^N_\omega \bar{\theta}_y,\\
&\mathcal{L}^e_3 = -\mathbf{u}^e_y \cdot \nabla\bar{\theta} - \mathbf{u}^e \cdot \nabla\bar{\theta}_y + 2c^e_\omega \bar{\theta}_y, \quad \mathcal{L}^{\bar{e}}_3 = -\bar{\mathbf{u}}^e \cdot \nabla\xi + (2\bar{c}^e_\omega - \bar{v}^e_y)\xi - \bar{u}^e_y \eta,
\end{aligned} \tag{3.12}$$

where $\mathcal{L}^e_i, \mathcal{L}^{\bar{e}}_i$ denote the errors from $\psi^e, \bar{\psi}^e$, respectively. These operators depend on ω, η, ξ, and we drop the dependence in (3.12) to simplify the notation.

3.2. First correction and the construction of $\boldsymbol{\phi^N}$. According to the normalization condition and (3.3), the solution to (3.1) satisfies $\omega_x(0,t) = \eta_x(0,t) = 0$. To obtain an approximate solution with this condition, we make the first correction

$$\omega_k \to \omega_k - \omega_{k,x}(0,0)\chi_{11}, \quad \eta_k \to \eta_k - \eta_{k,x}(0,0)\chi_{21}, \tag{3.13}$$

where χ_{ij} are cutoff functions defined in (3.17) with $\chi_{ij} = x + O(|x|^4)$ near 0. We do not modify ξ_k since ξ_k already vanishes quadratically near $(0,0)$. We remark that the first correction does not change the second order derivatives of the solution near 0 and c_ω since

$$\partial_{xy}\chi_{11}(0) = \partial_{xy}\chi_{21}(0) = 0, \quad c_\omega(\chi_{11}) = -\partial_{xy}\phi_1(0) = 0,$$

where ϕ_1 is defined below:

$$\phi_1 = -\frac{xy^2}{2}\kappa_*(x)\kappa_*(y), \tag{3.14}$$

where $\kappa_*(x)$ is the cutoff function chosen in (D.5) in Appendix D.2 satisfying $\kappa_*(x) = 1 + O(|x|^4)$ near $x = 0$, and ϕ_1 satisfies $-\Delta\phi_1 = x + O(|x|^4)$. For the numeric stream function $\phi^N_{k,1}$ constructed at the beginning of section 3, we correct it as follows:

$$\phi^N_{k,1} \to \phi^N_{k,1} + \partial_x\Delta\phi^N_{k,1}(0)\phi_1 \triangleq \phi^N_k.$$

Since $\partial_x\Delta\phi_1(0) = -1$, this allows us to obtain

$$\begin{aligned} &\partial_x(-\Delta)\phi^N_k(0) = -\partial_x\Delta\phi^N_{k,1}(0) + \partial_x\Delta\phi^N_{k,1}(0) = 0, \\ &\Delta\phi^N_k = O(|x|^2), \quad \omega_k - (-\Delta)\phi^N_k = O(|x|^2). \end{aligned} \tag{3.15}$$

We further extend it to Lipschitz continuous solutions $\widehat{W}^{(1)} \triangleq (\hat{\omega}^{(1)}(t), \hat{\eta}^{(1)}(t), \hat{\xi}^{(1)}(t)), \hat{\phi}^{N,(1)}$ in time using a cubic polynomial interpolation in t. See section 3.4 for more details. Here, we use $\hat{f}^{(1)}$ to denote the solution with the first correction.

3.3. The second correction. The error

$$(\partial_t - \mathcal{L}_i)(\hat{\omega}^{(1)}(t), \hat{\eta}^{(1)}(t), \hat{\xi}^{(1)}(t))$$

may not vanish to the order $O(|x|^3)$, which is a property that we require in the energy estimate. Then we add the second correction

$$\begin{aligned} (\hat{\omega}^{(1)}(t), \hat{\eta}^{(1)}, \hat{\xi}^{(1)}(t), \hat{\phi}^{N,(1)}) \to &(\hat{\omega}^{(1)}(t) + a_1(t)\chi_{12}, \hat{\eta}^{(1)} + a_2(t)\chi_{22}, \hat{\xi}^{(1)}(t) \\ &+ a_3(t)\chi_{32}, \hat{\phi}^{N,(1)} + a_1(t)\phi_2), \end{aligned}$$

so that the error satisfies

$$\varepsilon^{(2)}_i \triangleq (\partial_t - \mathcal{L}_i)(\hat{\omega}^{(1)}(t) + a_1(t)\chi_{12}, \hat{\eta}^{(1)}(t) + a_2(t)\chi_{22}, \hat{\xi}^{(1)}(t) + a_3(t)\chi_{32}) = O(|x|^3) \tag{3.16}$$

near $x = 0$. We use the following functions for these two corrections:

$$\begin{aligned} &\chi_{11} = -\Delta\phi_1, \quad \phi_1 = -\frac{xy^2}{2}\kappa_*(x)\kappa_*(y), \quad \chi_{21} = x\kappa_*(x)\kappa_*(y), \\ &\chi_{12} = -\Delta\phi_2, \quad \phi_2 = -\frac{xy^3}{6}\kappa_*(x)\kappa_*(y), \quad \chi_{22} = xy\kappa_*(x)\kappa_*(y), \quad \chi_{32} = \frac{x^2}{2}\kappa_*(x)\kappa_*(y), \end{aligned} \tag{3.17}$$

where $\kappa_*(x)$ is chosen in (D.5), $\chi_{\cdot,1}$ is used for the first correction, and $\chi_{\cdot,2}$ for the second correction. We do not have χ_{31} since we do need the first correction for ξ (3.13). Since $\kappa_*(x)$ satisfies $\kappa_*(x) = 1 + O(|x|^4)$ near $x = 0$, the behaviors of the above functions near $x = 0$ are given by

$$\chi_{11} = y + l.o.t.,\ \chi_{21} = x + l.o.t.,\ \chi_{12} = xy + l.o.t.,\ \chi_{22} = xy + l.o.t.,\ \chi_{32} = x^2/2 + l.o.t.$$

We choose $\chi_{1j} = -\Delta\phi_j$ for the correction of ω so that its associated velocity $\nabla^\perp(-\Delta)^{-1}\chi_{1j}$ can be obtained explicitly. We do not need such a form for the correction of η, ξ since we do not compute the velocity of η, ξ.

For cutoff functions χ_1, χ_2, χ_3 with

$$c_\omega(\chi_1) = -\partial_{xy}(-\Delta)^{-1}\chi_1 = 0, \tag{3.18}$$

e.g., $\chi_i = \chi_{i2}$ chosen above, we have the following formulas of $\mathcal{L}_i(a_1(t)\chi_1, a_2(t)\chi_2, a_3(t)\chi_3)$ (3.1):

$$\begin{aligned}
\mathcal{L}_1(a_1\chi_1, a_2\chi_2, a_3\chi_3) &= a_1(t)\Big(-(\bar{c}_l x + \bar{\mathbf{u}})\cdot\nabla\chi_1 + \bar{c}_\omega\chi_1 - \mathbf{u}(\chi_1)\cdot\nabla\bar{\omega}\Big) + a_2(t)\chi_2,\\
\mathcal{L}_2(a_1\chi_1, a_2\chi_2, a_3\chi_3) &= a_2(t)\Big(-(\bar{c}_l x + \bar{\mathbf{u}})\cdot\nabla\chi_2 + (2\bar{c}_\omega - \bar{u}_x)\chi_2\Big)\\
&\quad - a_3(t)\bar{v}_x\chi_3 - a_1(t)\Big(\mathbf{u}(\chi_1)\cdot\nabla\bar{\theta}\Big)_x,\\
\mathcal{L}_3(a_1\chi_1, a_2\chi_2, a_3\chi_3) &= a_3(t)\Big(-(\bar{c}_l x + \bar{\mathbf{u}})\cdot\nabla\chi_3 + (2\bar{c}_\omega + \bar{u}_x)\chi_3\Big)\\
&\quad - a_2(t)\bar{u}_y\chi_2 - a_1(t)\Big(\mathbf{u}(\chi_1)\cdot\nabla\bar{\theta}\Big)_y,
\end{aligned}$$

where $\mathbf{u}(\chi_1)$ is the velocity associated with χ_1. We want to apply the above formulas to the second corrections $\chi_{i2}, i = 1,2,3$ in (3.17). We use the Hadamard product

$$(A\circ B)_i = A_i B_i \tag{3.19}$$

and (3.12) to simplify the notation as follows:

$$\begin{aligned}
&\mathcal{L}_i(a\circ\chi) = Cor_{ij}(x;\chi)a_j(t), \quad Cor_{ij}(x;\chi) = Cor^N_{ij}(x;\chi) + Cor^{\bar{e}}_{ij}(x;\chi),\\
&\mathcal{L}^N_i(a\circ\chi) \triangleq Cor^N_{ij}(x;\chi)a_j(t), \quad \mathcal{L}^{\bar{e}}_i(a\circ\chi) \triangleq Cor^{\bar{e}}_{ij}(x;\chi)a_j(t).
\end{aligned} \tag{3.20}$$

Note that $\mathcal{L}^e_i(a\circ\chi) = 0$ since we can obtain $\mathbf{u}(\chi_1)$ explicitly for $\chi_1 = \chi_{11}, \chi_{12}$ (3.17).

Next, we derive the equations for $a_i(t), i = 1,2,3$. Using (3.1) and the condition

$$\partial_{xy}\varepsilon^{(2)}_1(0) = \partial_{xy}\varepsilon^{(2)}_2(0) = \partial_{xx}\varepsilon^{(2)}_3(0) = 0,$$

from (3.16), we obtain the following ODEs for $a(t), b(t), c(t)$:

$$\begin{aligned}
\dot{a}_1(t) &= (-2\bar{c}_l + \bar{c}_\omega)a_1(t) + a_2(t) - F_1(t),\\
\dot{a}_2(t) &= (-2\bar{c}_l + 2\bar{c}_\omega - \bar{u}_x(0))a_2(t) - F_2(t),\\
\dot{a}_3(t) &= (-2\bar{c}_l + 2\bar{c}_\omega - \bar{u}_x(0))a_3(t) - F_3(t),
\end{aligned} \tag{3.21}$$

where $F(t) = (F_1(t), F_2(t), F_3(t))^T$ is the error associated to the second order derivatives of $(\partial_t - \mathcal{L})\hat{W}^{(1)}$ near 0. More precisely, we have

$$\begin{aligned}
F_1(t) &= \partial_{xy}(\partial_t - \mathcal{L}_1)\widehat{W}^{(1)}(0) = \frac{d}{dt}\hat{\omega}^{(1)}_{xy}(t,0)\\
&\quad - (-2\bar{c}_l + \bar{c}_\omega)\hat{\omega}^{(1)}_{xy}(t,0) - \hat{\eta}^{(1)}_{xy}(t,0) - c_\omega(t)\bar{\omega}_{xy}(0),\\
F_2(t) &= \partial_{xy}(\partial_t - \mathcal{L}_2)\widehat{W}^{(1)}(0) = \frac{d}{dt}\hat{\eta}^{(1)}_{xy}(t,0)\\
&\quad - (-2\bar{c}_l + 2\bar{c}_\omega - \bar{u}_x(0))\hat{\eta}^{(1)}_{xy}(t,0) - c_\omega(t)\bar{\theta}_{xxy}(0),\\
F_3(t) &= \partial^2_x(\partial_t - \mathcal{L}_3)\widehat{W}^{(1)}(0) = \frac{d}{dt}\hat{\xi}^{(1)}_{xx}(t,0)\\
&\quad - (-2\bar{c}_l + 2\bar{c}_\omega - \bar{u}_x(0))\hat{\xi}^{(1)}_{xx}(t,0) - c_\omega(t)\bar{\theta}_{xxy}(0).
\end{aligned} \tag{3.22}$$

Denote

$$D^2 = (D_1^2, D_2^2, D_3^2) = (\partial_{xy}, \partial_{xy}, \partial_x^2)^T. \tag{3.23}$$

Then we can simplify (3.22) as

$$F_i = D_i^2(\partial_t - \mathcal{L}_i)\hat{W}^{(1)}(0) = D_i^2(\partial_t - \mathcal{L}_i^N - \mathcal{L}_i^e - \mathcal{L}_i^{\bar{e}})\hat{W}^{(1)}(0). \tag{3.24}$$

Denote by M the coefficients in (3.21)

$$M = \begin{pmatrix} -2\bar{c}_l + \bar{c}_\omega & 1 & 0 \\ 0 & -2\bar{c}_l + 2\bar{c}_\omega - \bar{u}_x(0) & 0 \\ 0 & 0 & -2\bar{c}_l + 2\bar{c}_\omega - \bar{u}_x(0). \end{pmatrix} \triangleq M^N + M^{\bar{e}}, \tag{3.25}$$

where the last identity is based on the decomposition $\bar{c}_\omega = \bar{c}_\omega^N + \bar{c}_\omega^e, \bar{u}_x(0) = \bar{u}_x^N(0) + \bar{u}_x^e(0)$, and $M^{\bar{e}}$ only contains the contribution from $\bar{c}_\omega^e, \bar{u}_x^{\bar{e}}(0)$. According to the normalization condition (3.2), we have $\bar{u}_x(0)^e = \bar{c}_\omega^e$. It follows that

$$M^{\bar{e}} = \bar{c}_\omega^e I_3. \tag{3.26}$$

We simplify the ODE for $a = (a_1, a_2, a_3)^T$ as

$$\dot{a}_i(t) = M_{ij}a_j(t) - F_i(t), \quad \dot{a}(t) = Ma - F = Ma - e_i D_i^2(\partial_t - \mathcal{L}_i)\hat{W}^{(1)}(0). \tag{3.27}$$

Recall $\chi_{\cdot 2} = (\chi_{12}, \chi_{22}, \chi_{32})$ from (3.17). In the ith equation, the overall error for the approximate solution $\widehat{W}^{(1)} + a(t) \circ \chi_{\cdot 2}$ is

$$\begin{aligned}(\partial_t - \mathcal{L}_i)(\widehat{W}^{(1)} + a(t) \circ \chi_{\cdot 2}) &= (\partial_t - \mathcal{L}_i^N)(a(t) \circ \chi_{\cdot 2}) + \Big((\partial_t - \mathcal{L}_i^N)\widehat{W}^{(1)} \\ &\quad - \mathcal{L}_i^e(\widehat{W}^{(1)} + a(t) \circ \chi_{\cdot 2}) - \mathcal{L}_i^{\bar{e}}(\widehat{W}^{(1)} + a(t) \circ \chi_{\cdot 2})\Big) \triangleq J + I.\end{aligned} \tag{3.28}$$

Note that in the above notation, ∂_t acts on $a_i(t)\chi_{i,2}$. For J, using the ODE for $a(t)$, (3.27), (3.20), (3.24), and (3.25), we get

$$\begin{aligned}J &= (M_{ij}a_j - F_i)\chi_{i2} - Cor_{ij}^N(x; \chi_{\cdot 2})a_j \\ &= (M_{ij}^N \chi_{i2} - Cor_{ij}^N(x; \chi_{\cdot 2}))a_j + M_{ij}^{\bar{e}} a_j \chi_{i2} - D_i^2(\partial_t - \mathcal{L}_i^N - \mathcal{L}_i^e - \mathcal{L}_i^{\bar{e}})\widehat{W}^{(1)}(0)\chi_{i2} \\ &\triangleq J_1 + J_2 + J_3,\end{aligned}$$

where we have a summation over $j = 1, 2, 3$. Since $\mathcal{L}^e(a(t) \circ \chi_{\cdot 2}) = 0$, using the above decomposition and combining I, J_2, J_3 yields

$$\begin{aligned}I + J_2 + J_3 &= \Big((\partial_t - \mathcal{L}_i^N)\widehat{W}^{(1)} - D_i^2(\partial_t - \mathcal{L}_i^N)\widehat{W}^{(1)}(0)\chi_{i2}\Big) \\ &\quad - \Big(\mathcal{L}_i^e\widehat{W}^{(1)} - D_i^2\mathcal{L}_i^e\widehat{W}^{(1)}(0)\chi_{i2}\Big) \\ &\quad - \Big(\mathcal{L}_i^{\bar{e}}(\widehat{W}^{(1)} + a(t) \circ \chi_{\cdot 2}) - D_i^2\mathcal{L}_i^{\bar{e}}\widehat{W}^{(1)}(0)\chi_{i2} - M_{ij}^{\bar{e}}a_j\chi_{i2}\Big) \\ &\triangleq I_{i,N} + I_{i,e} + I_{i,\bar{e}}.\end{aligned} \tag{3.29}$$

Next, we check that $J_1, I_{i,N}, I_{i,e}, I_{i,\bar{e}}$ have a vanishing order $O(|x|^3)$. This is clear for $I_{i,N}, I_{i,e}$. Since we correct the second order derivatives and $\hat{\omega}^{(1)}, \hat{\eta}^{(1)}, \hat{\zeta}^{(1)}$ are odd with $\hat{\xi}^{(1)} = x\hat{\zeta}^{(1)}$, we get $\partial_x^i\partial_y^j I_{i,N}$, $\partial_x^i\partial_y^j I_{i,e} = 0, i + j \le 2$, at the origin. For J_1, we

note that it is a linear combination of a_j with given coefficients $M_{ij}^N - Cor_{ij}^N$. Its cubic vanishing order follows from the definition. For example, when $i=j=1$, we have

$$\begin{aligned} S &= a_1(t)\cdot(Cor_{11}^{\bar e}(x) - M_{11}^{\bar e}\chi_{12}) = a_1(t)\Big(-\bar{\mathbf{u}}^e\cdot\nabla\chi_{12} + \bar c_\omega^e\chi_{12} - \bar c_\omega^e\chi_{12}\Big)\\ &= a_1(t)\Big(-\bar{\mathbf{u}}^e\cdot\nabla\chi_{12}\Big). \end{aligned}$$

Since $\chi_{12} = xy + O(|x|^4)$ (3.17), $\bar u^e = \bar u_x^e(0)x + O(|x|^2), \bar v^e = -\bar u_x^e(0)y$ near 0, we have $S = O(|x|^3)$ near 0. The vanishing order of other terms in J_1 can be obtained similarly. Then for J_1, we estimate the weighted norm for $Cor_{ij}^{\bar e}(x) - M_{ij}^{\bar e}\chi_{i2}$ and then apply the triangle inequality to further bound J_1. Similarly, for a fixed i, we have the following vanishing order:

$$\begin{aligned} &\mathcal{L}_i^{\bar e}(a(t)\circ\chi_{\cdot 2}) - M_{ij}^{\bar e}a_j\chi_{i2} = Cor_{ij}^{\bar e}a_j(t) - M_{ij}^{\bar e}a_j\chi_{i2} = O(|x|^3),\\ &\quad D_i^2\mathcal{L}_i^{\bar e}(a(t)\circ\chi_{\cdot 2})(0) = M_{ij}^{\bar e}a_j. \end{aligned}$$

Thus, we can rewrite $I_{i,\bar e}$ as follows:

$$I_{i,\bar e} = -\Big(\mathcal{L}_i^{\bar e}(\widehat{W}^{(1)} + a(t)\circ\chi_{\cdot 2}) - D_i^2\mathcal{L}_i^{\bar e}(\widehat{W}^{(1)} + a(t)\circ\chi_{\cdot 2})(0)\chi_{i2}\Big), \tag{3.30}$$

which clearly has a cubic vanishing order. Note that $\widehat{W}^{(1)} + a(t)\circ\chi_{\cdot 2}$ is our final approximate solution for solving (3.1).

In summary, to estimate the error $(\partial_t - \mathcal{L})(\widehat{W}^{(1)} + a\circ\chi_{\cdot 2})$, we will estimate $J_1, I_{i,N}, I_{i,e}, I_{i,\bar e}$ separately. The term $I_{i,N}$ is the local error of solving (3.1) numerically, and $I_{i,e}, I_{i,\bar e}$ are due to the error of solving the Poisson equations for ω and $\widehat{\omega}^{(1)}$. Since we use a cubic polynomial interpolation to obtain the continuous function $\hat W^{(1)}(t)$, the errors $I_{i,N}, I_{i,e}$ are piecewise cubic polynomials in time, and we track the coefficients of these polynomials to verify that they are small. We discuss the estimate of nonlocal error in section 3.7.

3.4. Cubic interpolation in time. Given the numerical solution with the first correction $\widehat{W}_n^{(1)} = (\hat\omega_n^{(1)}, \hat\eta_n^{(1)}, \hat\xi_n^{(1)})$, we use a piecewise cubic interpolation to construct $\widehat{W}^{(1)}(t,x)$ over $(t,x)\in[0,T]\times\mathbb{R}_2^+$. We partition the whole time interval $[0,T]$ into small subintervals $[3mk, 3(m+1)k]$ with length $3k$. For $s\in[-3k/2, 3k/2]$ and $t_m = 3mk$, we construct

$$\begin{aligned} \hat W^{(1)}\left(s + t_m + \frac{3k}{2}\right) &= \frac{1}{16}\left(-W_0 + 9W_1 + 9W_2 - W_3\right)\\ &\quad + \frac{1}{24}(W_0 - 27W_1 + 27W_2 - W_3)\frac{s}{k}\\ &\quad + \frac{1}{4}(W_0 - W_1 - W_2 + W_3)\left(\frac{s}{k}\right)^2\\ &\quad + \frac{1}{6}(-W_0 + 3W_1 - 3W_2 + W_3)\left(\frac{s}{k}\right)^3\\ &\triangleq \sum_{i\le 3} C_i\cdot V\frac{1}{i!}\left(\frac{s}{k}\right)^i, \quad V = (W_0, W_1, W_2, W_3), \end{aligned}$$

where k is the time step, $W_i = \hat{W}^{(1)}_{3m+i}$ for $t_m = 3mk$, and $C_i \in \mathbb{R}^4$ is the coefficient determined by the interpolation formula. A direct calculation yields

$$\begin{aligned}\partial_t \widehat{W}^{(1)} - \mathcal{L}\widehat{W}^{(1)} &= \sum_{1\le i\le 3} \frac{C_i\cdot V}{k}\frac{1}{(i-1)!}\Big(\frac{s}{k}\Big)^{i-1} - \sum_{i\le 3}\mathcal{L}(C_i\cdot V)\frac{1}{i!}\Big(\frac{s}{k}\Big)^{i}\\ &= \sum_{i\le 2}\Big(\frac{C_{i+1}\cdot V}{k} - \mathcal{L}(C_i\cdot V)\Big)\frac{1}{i!}\Big(\frac{s}{k}\Big)^{i} - \mathcal{L}(C_4\cdot V)\frac{s^3}{6k^3}.\end{aligned}$$

To estimate $\partial_t\widehat{W}^{(1)} - \mathcal{L}\widehat{W}^{(1)}$, we will use the triangle inequality and estimate $\frac{C_{i+1}\cdot V}{k} - \mathcal{L}(C_i\cdot V), \mathcal{L}(C_4\cdot W)$ rigorously using the methods in sections 3.6, 3.7.

Applying the triangle inequality and integrating the error over $s\in[-\frac{3k}{2},\frac{3k}{2}]$ yields

(3.31)
$$\begin{aligned}&\int_{|s|\le 3k/2}|\partial_t\hat{W}^{(1)} - \mathcal{L}\hat{W}^{(1)}|ds \le \sum_{i\le 2}\Big|\frac{C_{i+1}\cdot V}{k} - \mathcal{L}(C_i\cdot V)\Big|\int_{|s|\le 3k/2}\frac{1}{i!}|\frac{s}{k}|^i\\ &+ |\mathcal{L}(C_4\cdot V)|\int_{|s|\le\frac{3k}{2}}\frac{1}{6}|\frac{s}{k}|^3 = k\left(\sum_{i\le 2}\Big|\frac{C_{i+1}\cdot V}{k} - \mathcal{L}(C_i\cdot V)\Big|C_I(i) + |\mathcal{L}(C_4\cdot V)|C_I(3)\right),\end{aligned}$$

where

$$C_I = \left[3, \frac{9}{4}, \frac{9}{8}, \frac{27}{64}\right].$$

3.4.1. Decomposing the time interval for parallel computing. To verify that the posteriori error is small, we need to estimate the error rigorously at each time step, which takes a significant amount of time. Consider a partition of the time interval $0 = T_0 < T_1 < \cdots < T_n = T$, where T is the final time of the computation. To reduce the computational time, we first solve the equations on $[0,T]$ without any rigorous verification and save the solution $(\omega_k,\eta_k,\xi_k,\phi^N_{k,1})$ at $t_k = T_i$. Since we do not need to perform verification at this step, the running time for each time step is short. Then we solve the equations on a smaller time interval $[T_i,T_{i+1}], i = 0,1,2\ldots,n-1$ using $W(T_i)$ as the initial data and then perform the verification in each time interval in parallel. At the end of each time interval $[T_i,T_{i+1}]$, we use the precomputed data $W(T_{i+1})$, which is the same as the initial data for next time interval $[T_{i+1},T_{i+2}]$ for verification. This guarantees that we use the same discrete solution $(\omega_k,\eta_k,\xi_k,\phi^N_{k,1})$ for verification in $[T_i,T_{i+1}]$ and $[T_{i+1},T_{i+2}]$.

3.5. Compactly supported in time. To construct an approximate solution, we do not need to solve the linearized equations (3.1) for all time. In fact, since the solution decays in certain norm as t increases, we stop the computation at time T if $\hat{W}^{(1)} - D^2\hat{W}^{(1)}\circ\chi$ is small in the energy norm. Then we extend $\hat{W}^{(1)}(t,\cdot)$ trivially for $t > T$

$$\widehat{W}^{(1)}(t,\cdot) = 0, \quad t > T.$$

As a result, the error satisfies

$$\mathcal{R}_i = (\partial_t - \mathcal{L}_i)\widehat{W}^{(1)} = (\partial_t - \mathcal{L}_i)\widehat{W}^{(1)}\mathbf{1}_{t\le T} - \delta_T(t)\widehat{W}^{(1)}_i(T).$$

Let $F = (F_1,F_2,F_3), F_i = D^2_i(\partial_t - \mathcal{L}_i)\widehat{W}^{(1)}\Big|_{x=0}$ for $t\le T$, where $D^2 = (D_{xy}, D_{xy}, D^2_x)$. Then similarly, we get

$$\begin{aligned}F_{ext} &\triangleq D^2(\partial_t - \mathcal{L})\widehat{W}^{(1)}|_{x=0}\cdot\mathbf{1}_{t\le T} - D^2\widehat{W}^{(1)}(T,0)\delta_T\\ &= F(t)\mathbf{1}_{t\le T} - F_{end}(T)\delta_T,\ F_{end}(T) \triangleq D^2\widehat{W}^{(1)}(T,0).\end{aligned}$$

We will test the above formulas with some Lipschitz function in time and the above formulas are well defined. Recall that the coefficients of the second correction a satisfy (3.27). Although $\hat{W}$ only has finite support in time, to achieve the vanishing order (3.16) for all time, we need to solve the ODE *e*xactly for all time. If we stop solving the ODE at time T, we cannot achieve (3.16) at time T. Moreover, we cannot solve the ODE using a numerical method, e.g., the Runge–Kutta method, since it leads to an error. Instead, we solve the ODE exactly by diagonalizing the system. We introduce the notation

$$\begin{aligned}&\lambda_1=-2\bar{c}_l+\bar{c}_\omega,\quad \lambda_2=\lambda_3=-2\bar{c}_l+2\bar{c}_\omega-\bar{u}_x(0),\quad \lambda_1-\lambda_2=-\bar{c}_l/2,\\ &\tilde{a}_1=a_1+\frac{a_2}{\lambda_1-\lambda_2},\quad \tilde{F}_1=F_1+\frac{F_2}{\lambda_1-\lambda_2},\quad \tilde{a}_i=a_i,\quad \tilde{F}_i=F_i,\ i=2,3,\end{aligned}\tag{3.32}$$

and similar notation for $\tilde{F}_{ext}$, where we have used (2.11) to get $\lambda_1-\lambda_2=-\bar{c}_l/2$. The coefficients satisfy $\lambda_1\approx -7, \lambda_2=\lambda_3\approx -5.5$. We diagonalize (3.21) as follows:

$$\frac{d}{dt}\tilde{a}_i=\lambda_i\tilde{a}_i-\tilde{F}_{ext,i}.$$

Using Duhamel's formula and the definition of $\tilde{F}_{ext,i}$ yields

$$\begin{aligned}\tilde{a}_j(t)&=e^{\lambda_j t}\tilde{a}_j(0)-\int_0^t e^{\lambda_j(t-s)}\tilde{F}_{ext,j}(s)ds\\ &=e^{\lambda_j t}\tilde{a}_j(0)-\int_0^{t\wedge T} e^{\lambda_j(t-s)}\tilde{F}_j(s)ds+\tilde{F}_{end}(T)e^{\lambda_j(t-T)}\mathbf{1}_{t\geq T}\triangleq S_1+S_2+S_3.\end{aligned}\tag{3.33}$$

For rank-one perturbation, the full solution $\hat{W}$ with two corrections in (3.6), (3.7) is given by

$$\widehat{W}=\widehat{W}^{(1)}+a\circ\chi_{\cdot 2},\quad \hat{\phi}^N=\hat{\phi}^{N,(1)}+a_1(t)\phi_2,\tag{3.34}$$

where $\chi_{\cdot 2},\phi_2$ are defined in (3.17). With the above extension and the decomposition of error (3.28)–(3.29), the residual error for rank-one perturbation (3.7) with $n=1$ is given by

$$\begin{aligned}\mathcal{R}&=c(t)(\widehat{W}_0^{(1)}+a_0\circ\chi_{\cdot 2}-\bar{W}_0)+\int_0^t c(t-s)(\partial_t-\mathcal{L})(\widehat{W}^{(1)}+a\circ\chi_{\cdot 2})ds\\ &=\mathcal{R}_{loc,0,\cdot}+\mathcal{R}_{nloc}\\ \mathcal{R}_{loc,0,\cdot}&=c(t)(\widehat{W}_0^{(1)}+a_0\circ\chi_{\cdot 2}-\bar{W}_0)-(\widehat{W}^{(1)}(T)-D^2\widehat{W}^{(1)}(T)\circ\chi_{\cdot 2})c(t-T)\mathbf{1}_{t\geq T}\\ &\quad+\int_0^{t\wedge T}c(t-s)\sum_{i\leq 3}e_i I_{i,N}(s))ds+\int_0^t c(t-s)\sum_{i\leq 3}e_i J_{1,i}(s))ds\\ &=\int_0^t c(t-s)\mathcal{R}_{num}(s)ds,\\ \mathcal{R}_{nloc}&=\int_0^{t\wedge T}c(t-s)\sum_{i\leq 3}e_i I_{i,e}(s)ds+\int_0^t c(t-s)\sum_{i\leq 3}e_i I_{i,\bar{e}}(s)ds,\end{aligned}\tag{3.35}$$

where $I_{i,N}, I_{i,e}, I_{i,\bar{e}}$ are given in (3.29), $J_{1,i}$ means J_1 (3.29) in the ith equation, and $\mathcal{R}_{num}$ is

$$\begin{aligned}\mathcal{R}_{num}(s)\triangleq{}&\delta_0\cdot(\widehat{W}_0^{(1)}+a_0\circ\chi_{\cdot 2}-\bar{W}_0)-\delta_T\cdot(\widehat{W}^{(1)}(T)-D^2\widehat{W}^{(1)}(T)\circ\chi_{\cdot 2})\\ &+e_j(\mathbf{1}_{t\leq T}I_{j,N}+J_{1,j}).\end{aligned}\tag{3.36}$$

We only integrate the integrals for $e_i I_{i,N}, e_i I_{i,e}$ up to $\min(t,T)$ since these two integrands (3.29) do not involve $a_i(t)$ and have compact support $[0,T]$ in time. We obtain the local part $\mathcal{R}_{loc,0,\cdot}$ in (3.8) for $n=1$. The first term is the initial interpolation error for $\bar{W}_0$, and we choose $a_0 \in \mathbb{R}^3$ to achieve vanishing order $\widehat{W}_0^{(1)} + a_0 \circ \chi - \bar{W}_0 = O(|x|^3)$. We use $\mathcal{R}_{loc,0,\cdot}, \mathcal{R}_{nloc}$ to denote the error that depends on the solution locally and nonlocally. We use the bootstrap assumption to obtain uniform control of $c(t)$ in t. See section 5.7 in Part I [13]. The error estimate of the local part $\mathcal{R}_{loc,0,\cdot}$ follows section 3.6. Moreover, we extract the essentially local part from $\mathcal{R}_{nloc}$ and can estimate it with $\mathcal{R}_{loc,0,j}$ together (3.38). We decompose the nonlocal part $\mathcal{R}_{nloc}$ in section 3.7. To control the terms involving a_i, e.g., $J_{1,i}$ above (3.29), we can estimate the weighted norm of the functions $Cor_{ij}^N(x) - M_{ij}^N \chi_{i2}$ and then only need to estimate the integral of $\tilde{a}_j$.

Denote $x \wedge y \triangleq \min(x,y)$. Since the factor $\lambda_j < 0$, using the formula of $\tilde{a}_j$ (3.33), we obtain

$$\int_0^\infty |S_1| dt = \frac{1}{|\lambda_j|} |\tilde{a}_j(0)|, \quad \int_0^\infty |S_3(t)| dt = \int_T^\infty |\tilde{F}_{end,j}(T)| e^{\lambda_j (t-T)} dt = \frac{1}{|\lambda_j|} |\tilde{F}_{end,j}(T)|,$$
$$\int_0^\infty |S_2(t)| dt \le \int_0^\infty \bigg(\int_0^{t \wedge T} e^{\lambda_j (t-s)} |\tilde{F}_j(s)| ds \bigg) dt = \int_0^T |\tilde{F}_j(s)| \bigg(\int_s^\infty e^{\lambda_j (t-s)} dt \bigg) ds$$
$$= \frac{1}{|\lambda_j|} \int_0^T |\tilde{F}_j(s)|.$$

It follows that

$$\int_0^\infty |\tilde{a}_j(t)| dt \le \frac{1}{|\lambda_j|} \left(|\tilde{a}_j(0)| + \int_0^T |\tilde{F}_j(s)| ds + |\tilde{F}_{end,j}(T)| \right).$$

In the estimate of the integral of $\tilde{F}_j$, (3.22), (3.32), we use $\bar{c}_\omega = \bar{c}_\omega^N + \bar{c}_\omega^e, c_\omega = c_\omega^N + c_\omega^e$ (3.11) and track the terms involving $\bar{c}_\omega^N, c_\omega^N$ in I_{e1} and error separately,

$$F_{j,e1} = \int_0^T |F_j^N(t)| dt, \ F_{j,e2} = \int_0^T |\bar{c}_\omega^e D_i^2 \widehat{W}^{(1)}(t,0)| dt, \ F_{j,e3} = \int_0^T |c_\omega^e(t,0) D_i^2 \bar{W}(0)| dt,$$
$$F_j^N = D_j^2 (\partial_t - \mathcal{L}_i^N) \widehat{W}^{(1)}(0), \quad D^2 = (\partial_{xy}, \partial_{xy}, \partial_{xx}).$$

From (3.11), we get $2\bar{c}_\omega^e - \bar{u}_x^e(0) = \bar{u}_x^e(0)$ and only 1 unit of error I_{e2} in $F_j(t), j = 2,3$. We track $\tilde{F}_j$ (3.32) similarly. Since $\widehat{W}^{(1)}, F, \tilde{F}, F^N, \tilde{F}^N$ (3.22) are cubic in time, we can estimate the above integrals following (3.31). Note that $|\tilde{F}_{end,j}(T)|$ does not involve the nonlocal error. Using the linear relation between $a_j, \tilde{a}_j$, we can estimate a_j.

Using the above estimates, we can represent the rank-one solution and estimate it as follows:

$$\hat{G}(t,x) = \int_0^t c(t-s) \hat{W}(s) ds, \quad \hat{W} = \hat{W}^{(1)} + a \circ \chi_{\cdot 2}$$
$$|\partial_x^i \partial_y^j G_l(t,x)| \le \sup_{t>0} |c(t)| \left(\int_0^T |\partial_x^i \partial_y^j \hat{W}_l^{(1)}(t)| dt + |\partial_x^i \partial_y^j \chi_{l2}| \int_0^\infty |a_l(t)| dt \right). \tag{3.37}$$

Similarly, we can bound other quantities for $\hat{G}$ and complete the estimates in (3.6).

We generalize the above formula and estimate directly to the finite rank perturbation operator using linearity. For different initial data $\bar{W}_0$ related to the finite

rank perturbation, we choose a different stopping time $T(\widehat{W}_0^{(1)})$ to save computation cost. In practice, we construct the numerical solution up to time $T(\widehat{W}_0^{(1)}) \le T = 12$. At that time, the solution $\hat{W}^{(1)}(T)$ is very small, which can be treated as a small perturbation. See the figures in section 4.3 in Part I [13].

Remark 3.1. Using linearity and the triangle inequality, we can assemble the estimates for $\mathcal{R}$ (3.7) from the estimates of each mode $\hat{W}_i$ in (3.6), (3.7). In practice, this means that we can implement the above estimate for each individual mode completely in parallel.

Finite support of the c_ω term in time. In section 5 of Part I [13], we need to use $c_\omega(f)$, where $c_\omega(f) = u_x(f)(0) = -\partial_{xy}(-\Delta)^{-1}f(0)$. Since we choose the cutoff function χ_{12} for the second correction of $\hat{\omega}$ with properties (3.17), (3.18), we get

$$c_\omega(\hat{W}_1^{(1)} + a_1(t)\chi_{12}) = c_\omega(\hat{W}_1^{(1)}),$$

and it is supported in $[0, T]$.

3.6. Ideas of estimating the norm of the error. In this section, we discuss how to estimate the error derived in the previous section, e.g., $I_{i,N}$ (3.29), a posteriori. The general idea is to first evaluate f on some grid points and estimate the higher order derivatives of f in a domain D. Then we can construct an approximation $\hat{f}$ of f by interpolating the values of f at different points. The approximation error $f - \hat{f}$ can be bounded by $C_k||f||_{C^k}h^k$, where h measures the size of the domain. If the mesh h is sufficiently small, the error term is small. See a simple second order error estimate in (C.12).

To develop an efficient method for rigorous estimates, we have the following considerations. First, we should evaluate as small a number of points as possible so that the method is efficient. Second, most functions f in the verification are complicated, e.g., $I_{i,N}$ (3.29), and it is difficult to obtain the sharp bound of the higher derivatives. Instead, we first estimate the piecewise derivatives of some simple functions, e.g., piecewise polynomials $(\hat{\omega}, \hat{\eta})$ or semianalytic solutions following Appendix C, D. Then we use the triangle inequality and the Leibniz rule to estimate the products of these simple functions, and their linear combinations. Yet, in general, this approach overestimates the derivatives significantly. To compensate for the overestimates, we use higher order interpolations and estimates with error bounds $Ch^k, k = 3, 4, 5$, which provide the small factor h^k. We develop three estimates based on different interpolations—the Newton interpolation, the Lagrangian interpolation, and the Hermite interpolation—in section 8 in the supplementary material (supplement.pdf [local/web 1.43MB]). The 1D interpolating polynomials are standard, and we generalize them to construct 2D interpolating polynomials.

We want to estimate the constant C in the error bound Ch^k as sharply as possible to reduce the computational cost and improve the efficiency. In fact, when $k = 4$, if we can obtain an interpolation method and reduce the constant C to $\frac{C}{16}$, to achieve the same level of error, we can increase h to $2h$. In this verification step, since the domain is 2D, it means that we can evaluate only $\frac{1}{4}$ of the grid point values of f, which can reduce the computational cost by 75%.

Using the above method, we can obtain a sharp estimate of the derivatives of f. Using the method in section 8 in the supplementary material (supplement.pdf [local/web 1.43MB]) and Taylor expansion, we can further estimate the weighted norm of f with a singular weight near 0. We discuss the estimate of the nonlocal error in section 3.7. Using these L^∞ estimates of f and its derivatives, we can further develop

a Hölder estimate for f. See section E.1. We remark that the numerical solutions are regular, e.g., the approximate steady state and the solutions to the linearized equations are $C^{4,1}$. We use these methods to estimate a piecewise $L^\infty(\varphi_{evo,i})$ norm of the local residual error $\mathcal{R}_{num,i}$ (3.36) and the $C^{1/2}_{x_i}$ partial Hölder seminorm of $\mathcal{R}_{num,i}\psi_i$, where $\varphi_{evo,i}, \psi_i$ are defined in (A.3).

We remark that the weights $\varphi_{evo,i}$ and $\varphi_i, i = 2, 3$ in the L^∞ energy estimate (see section 5 in [13]) for η, ξ are similar but with different coefficients $p_{5,\cdot}, p_{6,\cdot}$. Since φ_i and $\varphi_{evo,i}$ are equivalent, after we obtain the piecewise weighted $L^\infty(\varphi_{evo,i})$ estimate of the error, we can obtain a piecewise weighted $L^\infty(\varphi_i)$ estimate by estimating the ratio $\varphi_i/\varphi_{evo,i}$. Similarly, we can obtain a weighted $L^\infty(\varphi_{g,i})$ estimate of the error, where $\varphi_{g,i}$ is another weight in the energy estimate in section 5 in [13].

Estimate the local part of the residual error. Using the above methods, we can estimate the local part of the residual error $\bar{\mathcal{F}}_i$ for the approximate steady state and discuss the estimate in Appendix C.4. We further extract the local part of $\mathcal{R}_{nloc}$ (3.35), which has the form (3.8) obtained in section 3.7, and combine it with $\mathcal{R}_{loc,0,j}$ to get the essentially local residual error:

(3.38)
$$\mathcal{R}_{loc,i} = \mathcal{R}_{loc,0,i} + \mathcal{R}_{dif,i} + M, \quad \mathcal{R}_{dif,i} \triangleq D_i^2\mathcal{B}_{op,i}(\mathbf{u}(\bar{\varepsilon}), \hat{G})(0) \cdot (\chi_{i2} - f_{\chi,i}),$$
$$M \triangleq \mathcal{B}_{op,i}(\mathbf{u}(\hat{\varepsilon}), \bar{W}) - D_i^2\mathcal{B}_{op,i}(\mathbf{u}(\hat{\varepsilon}), \bar{W})(0)\chi_{i2} - \mathcal{B}_{op,i}(\mathbf{u}_A(\hat{\varepsilon}_1), (\nabla\mathbf{u})_A(\hat{\varepsilon}_1), \bar{W}),$$

where χ_{i2} is defined in (3.17). By definition (3.39) and following derivation of (3.24), we get

$$D_i^2\mathcal{B}_{op,i}(\mathbf{u}(\bar{\varepsilon}), \hat{G})(0) = u_x(\bar{\varepsilon})(0)V_i, \quad V = (\hat{G}_{1,xy}(0), \hat{G}_{2,xy}(0), \hat{G}_{3,xx}(0)).$$

To estimate each term, we follow section 3.6 and Appendix C.4. We perform the decomposition (C.18) $\mathbf{u}(\hat{\varepsilon}) = \mathbf{u}_A(\hat{\varepsilon}_1) + \hat{\mathbf{u}}(\hat{\varepsilon}_1) + \mathbf{u}(\hat{\varepsilon}_2)$ and similar decomposition for $\nabla\mathbf{u}(\hat{\varepsilon})$, with $(\bar{\varepsilon}, \chi_{\bar{\varepsilon}})$ in (C.18) replaced by $(\hat{\varepsilon}, \chi_{\hat{\varepsilon}})$, where $\chi_{\hat{\varepsilon}}$ is defined in (D.6). Using the linearity of $B_{op,i}$, we get

$$\mathcal{B}_{op,i}(\mathbf{u}(\hat{\varepsilon}), \bar{W}) - \mathcal{B}_{op,i}(\mathbf{u}_A(\hat{\varepsilon}_1), (\nabla\mathbf{u})_A(\hat{\varepsilon}_1), \bar{W}) = II_i(\hat{\varepsilon}_1) + II_i(\hat{\varepsilon}_2),$$
$$II_i(\hat{\varepsilon}_1) = \mathcal{B}_{op,i}(\hat{\mathbf{u}}(\hat{\varepsilon}_1), \widehat{\nabla\mathbf{u}}(\hat{\varepsilon}_1), \bar{W}), \; II_i(\hat{\varepsilon}_2) = \mathcal{B}_{op,i}(\mathbf{u}(\hat{\varepsilon}_2), \nabla\mathbf{u}(\hat{\varepsilon}_2), \bar{W}).$$

We have $\mathbf{u}_A = O(|x|^3), (\nabla\mathbf{u})_A = O(|x|^2)$ near 0, which implies $\mathcal{B}_{op,i}(\mathbf{u}_A(\hat{\varepsilon}_1), (\nabla\mathbf{u})_A(\hat{\varepsilon}_1), \bar{W}) = O(|x|^3)$ (3.4) and

$$\begin{aligned} M &= II_i(\hat{\varepsilon}_1) + II_i(\hat{\varepsilon}_2) - D_i^2\mathcal{B}_{op,i}(\mathbf{u}(\hat{\varepsilon}), \bar{W})(0)\chi_{i2} \\ &= II_i(\hat{\varepsilon}_1) + II_i(\hat{\varepsilon}_2) - D_i^2(II_i(\hat{\varepsilon}_1) + II_i(\hat{\varepsilon}_2))(0)\chi_{i2}. \end{aligned}$$

The term $I_{i,N}$ in $\mathcal{R}_{loc,0,i}$ (3.29), (3.35) is similar to II_i^N, and M has a similar form as $II_i(\bar{\varepsilon}_1) + II_i(\bar{\varepsilon}_2)$ in Appendix C.4. We have done the above decomposition for $\mathbf{u}(\bar{\varepsilon})$ in (C.18), (C.19) and refer therein for more details. Then the estimate of $\mathcal{R}_{loc,j}$ is similar to that in Appendix C.4. See section 5.8 in [13] for more discussion of the above forms.

Error for the initial data and at stopping time. The error $\widehat{W}^{(1)}(T) - D^2\widehat{W}^{(1)}(t) \circ \chi_{\cdot 2}$ at the stopping time has compact support and its estimate follows the methods in section 3.6. To bound the initial interpolation error $err_{in} \triangleq \widehat{W}_0^{(1)} + a_0 \circ \chi_{\cdot 2} - \bar{W}_0$ (3.35) in a large domain, we follow similar methods. The error involves $\bar{\omega}, \bar{\theta}$ which are supported globally. To bound err_{in} in the middle and far-field, since

$\hat{W}_0^{(1)} + a_0 \circ \chi_{i,2} = 0$, combining all the initial data from the finite rank perturbation (see Appendix C.2.1 of Part I [13]), we need to estimate

$$I_1 = c_\omega(\omega_1)\bar{\omega} - \hat{\mathbf{u}}(\omega_1)\cdot\nabla\bar{\omega}, \quad I_2 = 2c_\omega(\omega_1)\bar{\theta}_x - \hat{\mathbf{u}}\cdot\nabla\bar{\theta}_x - \hat{\mathbf{u}}_x\cdot\nabla\bar{\theta},$$
$$I_3 = 2c_\omega(\omega_1)\bar{\theta}_y - \hat{\mathbf{u}}\cdot\nabla\bar{\theta}_y - \hat{\mathbf{u}}_y\cdot\nabla\bar{\theta}$$

for large $|x|$. The approximation terms near 0 defined in section 4.2.1 of Part I [13] are supported near 0 and decay to zero as $|x|\to\infty$. In the far-field, $\hat{\mathbf{u}}(\omega_1)$ is only a rank-one term. We estimate the above terms using (C.21), (C.22) with $a = c_\omega(\omega_1)$ and the estimates in section C.4.

3.7. Posteriori error estimates of the velocity. In this section, we show that the nonlocal error in (3.35) has the desired forms in (3.8). Then we combine the estimate of such terms with the nonlinear energy estimate in section 5.8 in [13]. Using (3.5) and the definition of $\mathcal{L}^{\bar{e}}, \mathcal{L}^e$ (3.11), (3.12), we have

$$\mathcal{L}_j^{\bar{e}}(G) = \mathcal{B}_{op,j}(\mathbf{u}(\bar{\varepsilon}), G), \quad \mathcal{L}_j^{e}(G) = \mathcal{B}_{op,j}(\mathbf{u}(G + (-\Delta)\phi_G^N), \bar{W}), \tag{3.39}$$

where ϕ_G^N is the numerical stream function associated with G.

Given $c_i(t)$ Lipschitz in t and $\bar{W}_i(0), i = 1, 2.., n$, we construct $\hat{W}_i(t)$ following previous sections and $\hat{G}$ using (3.6). Using the derivations in (3.35), (3.29), (3.30) and the above relation, the contribution from the error type $I_{j,\bar{e}}$ term to the error (3.7) in the jth equation is the following:

$$\mathcal{R}_j^{\bar{e}} \triangleq \int c(t-s) I_{j,\bar{e}}(s) ds = -(\mathcal{R}_{j0}^{\bar{e}} - D_j^2 \mathcal{R}_{j0}^{\bar{e}}(0)\chi_{j2}),$$
$$\mathcal{R}_{j0}^{\bar{e}} \triangleq \sum_{i\le n} \int_0^t c_i(s) \mathcal{B}_{op,j}(\mathbf{u}(\bar{\varepsilon}), \hat{W}_i(t-s)) dt.$$

Since $\mathcal{B}_{op,j}$ is bilinear and $c_i(t)$ is spatial-independent and Lipschitz in t, we get

$$\mathcal{R}_{j0}^{\bar{e}} = \mathcal{B}_{op,j}(\mathbf{u}(\bar{\varepsilon}), \left(\sum_{i\le n} \int_0^t c_i(t-s)\hat{W}_i(s) ds\right) = \mathcal{B}_{op,j}(\mathbf{u}(\bar{\varepsilon}), \widehat{G}(t)). \tag{3.40}$$

Denote by $\widehat{G}^{(l)}, \hat{W}_j^{(l)}$ the approximate solution with extension in t in section 3.4, and the first correction $l = 1$ in section 3.2 or two corrections $l = 2$ in sections 3.2, 3.3. Let $\hat{\phi}_i^{(l)}$ be the stream function associated with $\hat{W}_i^{(1)}$ constructed numerically with the first correction for $l = 1$ and both corrections for $l = 2$. In particular, the full solution is given by $\hat{G} = \hat{G}^{(2)}, \hat{W}_i = \hat{W}_i^{(2)}, \hat{\phi}_i^N = \hat{\phi}_i^{N,(2)}$ (3.34). We construct the stream function $\hat{\phi}^{N,(l)}$ associated with $\hat{G}^{(l)}$ and error $\hat{\varepsilon}$ as follows:

$$\hat{\phi}^{N,(l)} \triangleq \sum_{i\le n} \int_0^t c_i(s)\hat{\phi}_i^{(l)}(t-s) ds, \quad \hat{\varepsilon} = \hat{W}^{(1)} + \Delta\hat{\phi}^{N,(1)}$$
$$= \sum_{i\le n} \int_0^t c_i(s)(\hat{W}_i^{(1)} + \Delta\hat{\phi}_i^{N,(1)})(t-s) ds.$$

Since we can obtain $\mathbf{u}(a(t)\chi_{12})$ exactly for the second correction (see section 3.3), we have

$$\hat{\varepsilon}(t) = \hat{W}^{(1)} - (-\Delta)\hat{\phi}^{N,(1)} = \hat{W}^{(2)} - (-\Delta)\hat{\phi}^{N,(2)} = \hat{W} - (-\Delta)\hat{\phi}^N. \tag{3.41}$$

In practice, we estimate $\hat{\varepsilon}$ using the first identity since it does not involve $a_i(t)$ and the integrand $\hat{W}_i^{(1)} + \Delta\hat{\phi}_i^{(1)}$ is piecewise cubic in time. We decompose $\hat{\varepsilon}$ as follows:

$$\hat{\varepsilon}_2 = \hat{\varepsilon}_{xy}(0)\Delta\Big(\frac{x^3 y}{2}\chi_{\hat{\varepsilon}}\Big), \quad \hat{\varepsilon} = (\hat{\varepsilon} - \hat{\varepsilon}_2) + \hat{\varepsilon}_2 \triangleq \hat{\varepsilon}_1 + \hat{\varepsilon}_2, \tag{3.42}$$

where $\chi_{\hat{\varepsilon}}$ is defined in (D.6). Since $\hat{\varepsilon}$ only vanishes $O(|x|^2)$ near 0, we perform the above decomposition so that $\hat{\varepsilon}_1 = O(|x|^3)$ near 0. See Appendix C.4 and section 5.8 in [13] for motivations of (3.42). We estimate $\hat{\varepsilon}_1, \hat{\varepsilon}_{xy}(0), \hat{\phi}^N$ following (3.37). We establish (3.6).

Similarly, using linearity, (3.39), and (3.41), we can rewrite the residual error in (3.35) from the $I_{j,e}$ term in (3.29) related to $\mathcal{L}_j^e(\cdot)$ as follows:

$$\begin{aligned}
\mathcal{R}_j^e &\triangleq \int c(t-s) I_{j,e}(s) ds = -(\mathcal{R}_{j0}^e - D_j^2 \mathcal{R}_{j0}^e(0)\chi_{j2}),\\
\mathcal{R}_{j0}^e &= \sum_{i\le n} \int_0^t c_i(t-s)\mathcal{L}_j^e(\hat{W}_i^{(1)})(s) ds\\
&= \sum_{i\le n} \int_0^t c_i(t-s)\mathcal{B}_{op,j}(\mathbf{u}(\hat{W}_i^{(1)} + \Delta\hat{\phi}_i^{N,(1)})(s), \bar{W}) ds\\
&= \mathcal{B}_{op,j}(\hat{\varepsilon}(t), \bar{W}),
\end{aligned}$$

which along with (3.39)–(3.40) for $\mathcal{R}_j^{\bar{e}}$ establishes the formula for $\mathcal{R}_{nloc}$ in (3.8).

4. Estimate the norm of the regular part of the velocity. In this section, we derive the constants in the upper bound in Lemma 2.3. We have constructed the finite rank approximation $\hat{f}$ for f in Lemma 2.3 in section 4.3 in Part I [13]. The estimate of the most singular part, e.g., $u_{x,a,b}(\omega\psi)$, in the $C^{1/2}$ estimate in Lemma 2.3 can be obtained using the sharp Hölder estimates in section 3 of Part I [13], where $u_{x,a,b}$ is defined via a localized kernel. In this section, we estimate other terms in Lemma 2.3, e.g., $I = \psi u_x(\omega) - u_{x,a,b}(\omega\psi) - \psi\hat{u}_x(\omega)$, involving the velocity with desingularized kernels, which are more regular.

In section 4.1, we outline the strategies in the estimate and decompose the integrals from the nonlocal terms into several parts based on their regularities. In section 4.2, we perform the L^∞ estimates in Lemma 2.3 and derive the constants. In section 4.3, we perform the Hölder estimate of different parts. In section 4.6, we combine the Hölder estimate of different parts, which provide the constants in Lemma 2.3. In particular, we reduce the L^∞ estimates and the $C^{1/2}$ estimates in Lemma 2.3 to bounding some explicit L^1 integrals depending on the weights, which can be estimated by a numerical quadrature with rigorous error control. We estimate these integrals with computer assistance. See the discussions in section 2.2.

We will apply the second estimates in Lemma 2.3 for the nonlocal error, e.g., $\mathbf{u}(\bar{\varepsilon})$ and $\bar{\varepsilon}$ is the error of solving the Poisson equations. Since we can estimate piecewise bounds of $\bar{\varepsilon}$ following section 3.6, instead of using the global norm, we improve the estimate using the localized norms, which are much smaller than the global norm. See section 4.7.

The kernels associated with $\mathbf{u}, \nabla\mathbf{u}$ are given by

$$\begin{aligned}
&K_1 \triangleq \frac{y_1 y_2}{|y|^4}, \quad K_2 \triangleq \frac{1}{2}\frac{y_1^2 - y_2^2}{|y|^4}, \quad K_u \triangleq \frac{y_2}{2|y|^2}, \quad K_v \triangleq -\frac{y_1}{2|y|^2},\\
&K_{u_x} = -K_1, \quad K_{u_y} = K_{v_x} = K_2.
\end{aligned} \tag{4.1}$$

Here, we have dropped the constant $\frac{1}{\pi}$, e.g., $u_x(\omega) = -\partial_{xy}(-\Delta)^{-1} = \frac{1}{\pi} K_{u_x} * \omega$. One needs to multiply $\frac{1}{\pi}$ back to obtain the final estimate.

Difficulties in the computations. In addition to the difficulties discussed in section 5.1 of Part I [13], e.g., singularities caused by the weights and kernels, the singular integral introduces several technical difficulties in our estimates. To address these difficulties, we need to consider different scenarios and decompose the domain of the integrals carefully in our computer-assisted estimates. Given $\omega\varphi \in L^\infty$, the velocity $\mathbf{u}$ and the commutator $\psi \cdot (\nabla \mathbf{u})(\omega) - (\nabla \mathbf{u})(\omega\psi)$ are only log-Lipschitz. The logarithm singularity introduces several difficulties. For example, if $\mathbf{u}$ is Lipschitz, a natural approach to estimate its Hölder norm in terms of $||\omega\varphi||_\infty$ is to estimate the piecewise bound of $\mathbf{u}$ and $\partial\mathbf{u}$, which are local in $\mathbf{u}$, and then use the method in section E.1. However, since $\mathbf{u}$ is only log-Lipschitz, we need to perform a decomposition of $\mathbf{u}$ into the regular part and the singular part carefully. For different parts, we will apply different estimates. See section 4.1.11 for ideas. For $\nabla\mathbf{u}$, the estimates are more involved since it is more singular.

4.1. Several strategies. We outline several strategies to estimate the nonlocal terms.

4.1.1. Integral with approximation. In our computation of $\mathbf{u}_A = \mathbf{u} - \hat{\mathbf{u}}, (\nabla\mathbf{u})_A = \nabla\mathbf{u} - \widehat{\nabla\mathbf{u}}$, where the approximation terms $\hat{\mathbf{u}}, \widehat{\nabla\mathbf{u}}$ are defined in section 4.3 of Part I [13], the rescaling argument still applies. Note that we do not have $\partial\mathbf{u}_A = (\partial\mathbf{u})_A$ since we design approximations for $\mathbf{u}, \nabla\mathbf{u}$ separately. We consider one approximation term $c(x)\int \mathbf{1}_{y\notin S} K(x_a, y)\omega(y)dy$ for $\int K(x,y)\omega(y)dy$ to illustrate the ideas, where S is the singular region associated with x_a. Suppose that K is $-d$-homogeneous. We want to estimate

$$I = \rho(x)\int_{\mathbb{R}^2} (K(x,y) - c(x)K(x_a,y)\mathbf{1}_{y\notin S})W(y)dy,$$

where W is the odd extension of ω from $\mathbb{R}^2_+$ to $\mathbb{R}^2$ (see (4.23)). Denote

$$f_\lambda(x) \triangleq f(\lambda x). \tag{4.2}$$

We choose $\lambda \asymp |x|$ and denote $x = \lambda\hat{x}, y = \lambda\hat{y}, x_a = \lambda\hat{x}_a$. Since $K(\lambda z) = \lambda^{-d}K(z)$, we have

$$\begin{aligned} I &= \rho_\lambda(\hat{x})\int_{\mathbb{R}^2} \Big(K(\lambda\hat{x},\lambda\hat{y}) - c(\lambda\hat{x})\mathbf{1}_{\lambda\hat{y}\notin S}K(\lambda\hat{x}_a,\lambda\hat{y})\Big)W(\lambda\hat{y})\lambda^2 d\hat{y} \\ &= \lambda^{2-d}\rho_\lambda(\hat{x})\int_{\mathbb{R}^2} \Big(K(\hat{x},\hat{y}) - c(\lambda\hat{x})\mathbf{1}_{\hat{y}\notin S/\lambda}K(\hat{x}_a,\hat{y})\Big)W_\lambda(\hat{y})d\hat{y}. \end{aligned} \tag{4.3}$$

The singular region becomes S/λ and close to $x_a/\lambda = \hat{x}_a$. For example, if $S = \{y : \max_i |y_i - x_{a,i}| \le a\}$, we have $S/\lambda = \{y : \max_i |y_i - \hat{x}_{a,i}| \le a/\lambda\}$. For the above integral, we will symmetrize the kernel and then estimate it using the norms $||W\varphi||_\infty$ and $[\omega\psi]_{C^{1/2}_{x_i}}, i = 1, 2$ (4.9).

The bulk and approximation. To take advantage of the scaling symmetry and overcome the singularity, in our computation for x away from the origin and not too large, we choose several dyadic rescaling parameters $\lambda = 2^i, i \in I$, e.g., $I = \{-4, -3, .., 10\}$. Then for any x with $\max(x_1, x_2) \in [2^i x_c, 2^{i+1}x_c]$, we can choose $\lambda = 2^i$ so that the rescaled $\hat{x} = \frac{x}{\lambda}$ satisfies

$$\hat{x} \in \begin{cases} [x_c, 2x_c]\times[0, 2x_c] \triangleq \Omega_1 & \text{if } x_2 \le x_1, \\ [0, x_c]\times[x_c, 2x_c] \triangleq \Omega_2 & \text{if } x_2 > x_1. \end{cases} \tag{4.4}$$

We also choose x_i ($(x_i, 0)$ is the singularity) and the size of the singular region t_i for the approximation term defined in section 4.3.2 of Part I [13] such that x_i/λ is on the grid point of the mesh and the boundary of the singular region $\{y : |x_i - y_1| \vee |y_2| \geq t_i/\lambda\}$, which aligns with one of the edges of a mesh cell. For example, this can be done by choosing the following y mesh in the near-field to discretize the y-integral, x_i, and t_i:

$$y_{1,i} = ih, \quad y_{2,i} = ih, \quad x_i = 2^{n_i}h, \quad t_i = 2^{m_i}h.$$

Then when we discretize the rescaled integral in y, e.g., (4.3), the singular region is the union of several mesh cells. For large y, it is away from the singularity $\hat{x}$. Then we can use an adaptive mesh in y_1, y_2 to discretize the integral.

We remark that in (4.3), if $x_a \neq 0$ and $x_a/\lambda \asymp x_a/|x|$ is too large or too small, since $c(x)$ is supported near x_a, $c(\lambda\hat{x}) = c(x)$ will be 0. This means that when we compute $\mathbf{u}_A(x), (\nabla\mathbf{u})_A$, if the coefficient of an approximation term with center x_i and parameter t_i is nonzero, e.g., $c(x) \neq 0$, then λ is comparable to x_i when we rescale the integral by λ. Thus $\hat{x}_i = x_i/\lambda$ is on the grid. We also choose t_i such that t_i/λ is a multiple of mesh size h for λ comparable to x_i.

Remark 4.1. Using the scaling symmetry and rescaling the integral by dyadic scales, we can compute the integral for $x \in [0, D]^2 \backslash [0, d]^2$ with roughly $O(\log(D/d))$ computational cost.

The near-field and the far-field. Recall the notation from section 4.3 in Part I [13]:

(4.5)
$$C_{u0} = x, \quad C_{v0} = -y, \quad C_{u_x 0} = 1, \quad C_{u_y 0} = C_{v_x 0} = 0,$$
$$C_{u_x} = -(x^2 - y^2), \quad C_{v_x} = 2xy, \quad C_{u_y} = 2xy, \quad C_u = -\Big(\frac{1}{3}x^3 - xy^2\Big), \; C_v = x^2 y - \frac{1}{3}y^3,$$
$$K_{ux0} = -\frac{4y_1y_2}{|y|^4}, \quad K_{00} = \frac{24y_1y_2(y_1^2 - y_2^2)}{|y|^8}, \quad \mathcal{K}_{00}(\omega) \triangleq \frac{1}{\pi}\int_{\mathbb{R}^2_{++}} K_{00}(y)\omega(y)dy.$$

If x is sufficiently small, i.e., $\max(x_1, x_2) < \min_{i\in I} 2^i x_c$, we choose $\lambda = \max(x_1, x_2)/x_c$ so that the rescaled $\hat{x} = \frac{x}{\lambda}$ is on the line $x_1 = x_c$ or $x_2 = x_c$. Assuming $\varphi(x) \geq |x|^{-\beta_1}|x|_1^{-\beta_2}$, $\rho \sim |x|^{-\alpha}$ near $x = 0$, and K is $-d$-homogeneous, then we get

(4.6)
$$\begin{aligned}|\rho(x)\int_{\mathbb{R}^{++}_2} K(x,y)\omega(y)dy| &\leq ||\omega_\lambda\varphi_\lambda||_{L^\infty}\rho_\lambda(\hat{x})\int_{\mathbb{R}^{++}_2}|K(\hat{x},\hat{y})|\varphi_\lambda(\hat{y})^{-1}\lambda^{2-d}d\hat{y}\\ &\leq ||\omega_\lambda\varphi_\lambda||_{L^\infty}\lambda^{\beta_1+\beta_2+2-d}\rho_\lambda(\hat{x})\int_{\mathbb{R}^{++}_2}|K(\hat{x},\hat{y})||\hat{y}|^{\beta_1}\hat{y}_1^{\beta_2}d\hat{y}.\end{aligned}$$

As $x \to 0$, $\lambda \to 0$. The factor $\lambda^{\beta_1+\beta_2+2-d}$ absorbs the large factor $\lambda^{-\alpha}$ in $\rho_\lambda(\hat{x})$. In our estimate of $\mathbf{u}_A, \nabla\mathbf{u}_A$, we have $\beta_1 + \beta_2 = 2.9$ for $\varphi_1, \varphi_{g,1}$, 2.5 for φ_{elli} (A.2), $(\alpha, d) = (2, 2)$ for $(\psi_1, \nabla\mathbf{u}_A)$ (A.1), and $(\alpha, d) = (3, 1)$ for $(\rho_{10}, \mathbf{u}_A)$ (A.2). We have $\beta_1 + \beta_2 + 2 - d - 2 > 0$.

In general, the above integral may not be integrable due to the growing weight $|y|^{\beta_1}y_1^{\beta_2}$. For $\mathbf{u}_A, \nabla\mathbf{u}_A$ with small x, it takes the form (see section 4.3 of Part I [13])

(4.7)
$$f(x) - C_{f0}(x)u_x(0) - C_f(x)\mathcal{K}_{00} = \int_{\mathbb{R}^2_{++}}(K_f^{sym} + C_{f0}(x)\frac{4}{\pi}\frac{y_1y_2}{|y|^4} - C_f(x)K_{00}(y))\omega(y)dy,$$

where C_{f0}, C_f, and K_{00} are defined in (4.5), and $f = u, v, u_x, v_x, u_y, v_y$. In particular, the associated kernel has a much faster decay rate $|y|^{-6}$, which will be shown in Appendix B.1.1. Thus, the integral is integrable.

Since $\lambda = \max(x_1, x_2)/x_c$ is very small, $\rho_\lambda(\hat{x})$ can be well approximated by the most singular power $c\lambda^{-\alpha}|x|^{-\alpha}$ for some $c > 0$, which can be estimated effectively after factorizing out $\lambda^{-\alpha}$.

Similarly, if x is sufficiently large, i.e., $\max(x_1, x_2) > \max_{i\in I} 2^{i+1}x_c$, we choose $\lambda = \frac{\max(x_1,x_2)}{x_c}$ so that the rescaled $\hat{x} = x/\lambda$ is on the line $x_1 = x_c$ or $x_2 = x_c$. Since λ is sufficiently large, we can estimate the weight $\rho_\lambda, \varphi_\lambda$ based on their asymptotic behavior.

Integral near 0. We have an approximation $I = -C_{f0}(x)K_{ux0}(y) - C_2(x)K_{00}(y)$ (4.5) for $K_{f0}^{sym}(x, y)$ with some smooth coefficients C_2 (C_2 may not be C_f). The term $C_{f0}(x)K_{ux0}(y)$ and K_f are both $-d$ homogeneous, $d = 1$ or 2. Since K_{ux0}, K_{00} are singular near 0, after we rescale the integral following (4.3), we decompose the symmetrized integral for y near 0 as follows:

(4.8)

$$
\begin{aligned}
II &= \int_{\mathbb{R}_2^{++}} \Big(K_f^{sym}(\hat{x}, \hat{y})\lambda^{2-d} - \lambda^{2-d}C_f(\hat{x})K_{ux0}(\hat{y}) - C_2(\lambda\hat{x})K_{00}(\hat{y})\lambda^{-2}\Big)\omega(\lambda\hat{y})d\hat{y} \\
&= \lambda^{2-d}\Big(\int_{\mathbb{R}_2^{++}} \Big(K_f^{sym}(\hat{x}, \hat{y}) - C_f(\hat{x})\mathbf{1}_{|\hat{y}|_\infty \geq k_{01}h}K_{ux0}(\hat{y}) \\
&\qquad - \lambda^{-4+d}C_2(\lambda\hat{x})\mathbf{1}_{|\hat{y}|_\infty \geq k_{02}h}K_{00}(\hat{y})\Big)\omega(\lambda\hat{y})d\hat{y} \\
&\quad - \int_{\mathbb{R}_2^{++}} \Big(C_f(\hat{x})\mathbf{1}_{|\hat{y}|_\infty \leq k_{01}h}K_{ux0}(\hat{y}) - \lambda^{-4+d}C_2(\lambda\hat{x})\mathbf{1}_{|\hat{y}|_\infty \leq k_{02}h}K_{00}(\hat{y})\Big)\omega(\lambda\hat{y})d\hat{y}\Big)
\end{aligned}
$$

for some small integers k_{0i} with $k_{0i}h < |\hat{x}|_\infty/2$, e.g., $k_{01} = 4, k_{02} = 20$, where $|a|_\infty = \max(a_1, a_2)$ and h is chosen in (4.14). We will estimate the first integral with regular integrand near $\hat{y} = 0$ using the method in section 4.1.3 and the last two integrals for $|\hat{y}|_\infty \leq k_{01}h, |\hat{y}|_\infty \leq k_{02}h$ analytically in section 4.4.1. We perform the above decomposition since $K_{00}(\hat{y}), K_{ux0}(\hat{y})$ are too singular to estimate them numerically.

We apply the above decompositions to the integrals in both L^∞ and $C^{1/2}$ estimates. We also apply the above decompositions to the approximation terms and estimate the integral of K_{ux0} separately near $y = 0$.

4.1.2. The scaling relations. We discuss several scaling relations, which will be useful in later computation. For a $-d$-homogeneous kernel K, i.e., $K(\lambda x) = \lambda^{-d}K(x)$, we have

$$
I(x) = \rho(x)\int K(x, y)\omega(y)dy = \rho_\lambda(\hat{x})\int K(\hat{x}, \hat{y})\omega_\lambda(\hat{y})\lambda^{2-d}d\hat{y} \triangleq \lambda^{2-d}I_\lambda(\hat{x}),
$$

where $x = \lambda\hat{x}, y = \lambda\hat{y}$. To compute the derivative of $I(x)$, using the chain rule, we have

$$
\partial_{x_i}I(x) = \lambda^{2-d}\frac{d\hat{x}_i}{dx_i}\partial_{\hat{x}_i}I_\lambda(\hat{x}) = \lambda^{1-d}\partial_{\hat{x}_i}I_\lambda(\hat{x}).
$$

For the L^∞ part, clearly, we get $|I(x)| = |I_\lambda(\hat{x})|$. To compute the Hölder norm, we use the following relation $|x - z| = \lambda|\hat{x} - \hat{z}|$ and

$$
\frac{|I(x) - I(z)|}{|x - z|^{1/2}} = \lambda^{-1/2}\frac{|I_\lambda(\hat{x}) - I_\lambda(\hat{z})|}{|\hat{x} - \hat{z}|^{1/2}}.
$$

In particular, for $i=1,2$, we have

(4.9)
$$||\omega_\lambda\varphi_\lambda||_\infty = ||\omega\varphi||_\infty, \quad [\omega_\lambda\psi_\lambda]_{C^{1/2}_{x_i}} = \lambda^{1/2}[\omega\psi]_{C^{1/2}_{x_i}}, \quad [f]_{C^{1/2}_{x_i}} := \sup_{y,z:y_i=z_i} \frac{|f(y)-f(z)|}{|y-z|^{1/2}}.$$

Using these scaling relations, we can perform the estimate in a rescaled domain with any $\lambda > 0$.

4.1.3. Mesh and the trapezoidal rule. After rescaling the integral with suitable scaling factor λ, we can restrict the rescaled singularity $\hat{x} \in [0,2x_c]^2 \backslash [0,x_c]^2$ (see (4.3), (4.4)).

If a domain Q is away from the singularity $\hat{x}$ of the kernel, applying (4.9), we get

$$\begin{aligned} \int_Q |K(\hat{x},y)||\omega_\lambda(y)|dy &\le ||\omega_\lambda\varphi_\lambda||_\infty \int_Q |K(\hat{x},y)|\varphi_\lambda^{-1}(y)dy \\ &= ||\omega\varphi||_\infty \int_Q |K(\hat{x},y)|\varphi_\lambda^{-1}(y)dy. \end{aligned} \tag{4.10}$$

Then, it suffices to estimate the integral of an explicit function $|K(\hat{x},y)|\varphi_\lambda^{-1}(y)$. If in addition, the region Q is small, e.g., Q is the grid $[y_i,y_{i+1}]\times[y_j,y_{j+1}]$ introduced below, we further apply

$$\int_Q |K(\hat{x},y)||\omega_\lambda(y)|dy \le ||\omega\varphi||_\infty ||\varphi_\lambda^{-1}||_{L^\infty(Q)} \int_Q |K(\hat{x},y)|dy.$$

Since the domain Q is small, the estimate is sharp. We use the following method to estimate $\int |K(\hat{x}_i,y)|dy$ for a suitable kernel K and $\hat{x}_i$ on the grid points.

We consider the estimate of the L^1 norm of some function f in $\mathbb{R}^{++}_2$, e.g., $f = K(\hat{x}_i,y)$ mentioned above. To discretize the integral, we design a uniform mesh in the domain $[0,b]^2$ covering Ω_1 and Ω_2 with mesh size h and adaptive mesh in the larger domain $[0,D]^2$

$$0 = y_0 < y_1 < \cdots < y_n = D, \quad y_i = ih,\ i \le b/h. \tag{4.11}$$

The finer mesh in the near-field $[0,b]^2$ allows us to estimate the integral with higher accuracy. We choose a sparser mesh in the far-field since y is away from the singularity $\hat{x}$ and the kernel decays in y. We partition the integral as follows:

$$\int_{\mathbb{R}^{++}_2} |f(y)|dy = \sum_{0\le i,j\le n-1} \int_{[y_i,y_{i+1}]\times[y_j,y_{j+1}]} |f(y)|dy + \int_{y\notin D} |f(y)|dy. \tag{4.12}$$

We focus on how to estimate the first part for nonsingular f. In section 4.4, we estimate the integral beyond $[0,D]^2$ using the decay of the integral. We will discuss how to estimate the integral near the singularity of the kernel in a later subsection.

Denote $Q = [a,b]\times[d,c], h_1 = b-a, h_2 = d-c$. We use the trapezoidal rule

$$\int_{[a,b]\times[c,d]} |f(y)|dy \le T(|f|,Q) + Err(f),$$

where

$$T(f,Q) \triangleq \frac{(b-a)(d-c)}{4}(f(a,c)+f(a,d)+f(b,c)+f(b,d)).$$

The error estimate of the above trapezoidal rule is not obvious due to the absolute sign. In fact, even if f is smooth, $|f|$ is only Lipschitz near the zeros of f. Since the set of zeros is hard to characterize and $|f|$ can have low regularity, we do not pursue a higher order quadrature rule. We have the following error estimate.

LEMMA 4.2 (trapezoidal rule for the L^1 integral). *For $f \in C^2(Q)$, we have*

$$\int_Q |f(y)|dy \le T(|f|, Q) + \frac{|Q|}{12}(h_1^2||f_{xx}||_{L^\infty(Q)} + h_2^2|f_{yy}||_{L^\infty(Q)}).$$

Remark 4.3. The above estimate shows that the trapezoidal rule remains second order accurate from the above. In particular, this error estimate is comparable to the case without taking the absolute value.

Proof. Define the linear interpolation of f in Q

$$L(f) = \sum_{i=1}^{4} \lambda_i(x) f_i, \quad E(f) = f - L(f),$$

where $\lambda_i(x)$ is linear and satisfies $\sum \lambda_i(x) = 1$ and $\lambda_i(x) \ge 0$ for $x \in Q$. Using the triangle inequality, we obtain

$$\int_Q |f|dy \le \int_Q |E(f)|dy + \int_Q \lambda_i(x)|f_i|dy = T(|f|, Q) + \int_Q |E(f)|dy.$$

We have the standard error bound for linear interpolation $E(f)$

$$|E(f)| \le \frac{||f_{xx}||_{L^\infty(Q)}}{2}|(x-a)(x-b)| + \frac{||f_{yy}||_{L^\infty(Q)}}{2}|(y-c)(y-d)|, \tag{4.13}$$

which can be obtained by first applying interpolation in x and then in y. It can also be established using the error estimate for the 2D Lagrangian interpolation with $k = 2$ in section 8 in the supplementary material (supplement.pdf [local/web 1.43MB]). Integrating the above estimate in x, y and using $\frac{1}{2}\int_0^1 t(1-t)dt = \frac{1}{12}$ concludes the proof. ∎

To estimate the integral $\int |K(x,y)|$ for all $\hat{x} \in \Omega_1, \Omega_2$ (4.4), we discretize $[0, 2a]^2$ using a uniform mesh with mesh size $h_x = h/2$. We use the above method to estimate $\int |K(\hat{x}_i, y)|dy$ for x_i on the grid points. After we estimate the derivatives of the kernel, we use the following lemma to estimate the integral for any x in a domain.

LEMMA 4.4. *Suppose that $K(x,y) \in C^2(P \times Q)$, $P = [a_1, b_1] \times [a_2, b_2], h_i = b_i - a_i, i = 1, 2$, and $Q = [a, b] \times [c, d]$. Let $L(K)(x,y) = \sum_{i,j=1,2} \lambda_{ij}(x) K((a_i, b_j), y)$ be the linear interpolation of $K(x,y)$ in x using $K((a_i, b_j), y), i, j = 1, 2$. Then for any $x \in P$, we have*

$$\int_Q |K(x,y)|dy \le \sum_{i,j=1,2} \lambda_{ij}(x) \int_Q |K((a_i, b_j), y)|dy + \Big(\frac{h_1^2}{8}||K_{xx}||_{L^\infty(P\times Q)} + \frac{h_2^2}{8}||K_{yy}||_{L^\infty(P\times Q)}\Big)|Q|.$$

The proof follows from (4.13), the triangle inequality, and $\frac{1}{2}|t(1-t)| \le \frac{1}{8}$ for $t \in [0, 1]$. We will apply the above lemma and sum Q over all the near-field domains $Q_{ij} = [y_i, y_{i+1}] \times [y_j, y_{j+1}]$ (4.11). Since $\sum_{ij} \lambda_{ij}(x) = 1$, we can simplify the first term as follows:

$$\sum_{i,j=1,2} \lambda_{ij}(x) \sum_{k,l \le n} \int_{Q_{kl}} |K((a_i, b_j), y)|dy \le \max_{1 \le i,j \le 2} \sum_{k,l \le n} \int_{Q_{kl}} |K((a_i, b_j), y)|dy.$$

Therefore, it suffices to estimate the integral for x on the grid points and the piecewise derivative bounds of the kernel.

We apply Lemmas 4.2, 4.4 to estimate the weighted integral related to the velocity. The integrands take the form (4.28), (4.29), (4.24). To estimate the error in the above integrals, we need to obtain a piecewise L^∞ estimate of the derivatives of the integrands in P, Q. We estimate the derivatives of the weights in Appendix A.1 and the kernel in Appendix B.

Parameters for the integrals. In our computation, we choose

$$h_x = 13 \cdot 2^{-12}, \quad h = 13 \cdot 2^{-11}, \quad x_c = 13 \cdot 2^{-5}, \tag{4.14}$$

which can be represented exactly in a binary system, to reduce the round-off error. The approximate values of the above parameters are $h_x \approx 0.0032, h \approx 0.0064, x_c \approx 0.4$. For $x \in [0, 2x_c]^2 \backslash [0, x_c]^2$ (4.4), we have

$$\max(x_1, x_2) \geq x_c = 64h = 128h_x. \tag{4.15}$$

In our decomposition of the integral, e.g., (4.24), (4.45), (4.49), we impose a constraint on the size of the singular region to satisfy $(k+1)h < x_c$ such that the region does not cover the origin.

4.1.4. Decomposition, commutators, and the Lipschitz norm. The most difficult part of the computation is to estimate the Hölder norm of $\nabla \mathbf{u}$, and we discuss several strategies. In this computation, we cannot first estimate the local Lipschitz norm of ∇u and then obtain the local Hölder norm due to the difficulties discussed at the beginning of section 4. We need to decompose the integral related to ∇u into several parts according to the distance between y and the singularity and use different estimates for different parts.

We focus on the integral related to u_x without subtracting any approximation term and assume that $x \in [0, 2x_c]^2 \backslash [0, x_c]^2$. The approximation term $\widehat{\nabla \mathbf{u}}$ is nonsingular and can be estimated using the method in section 4.1.3. Let h be the mesh size in the discretization of the integral in y. Suppose that

$$x \in \mathbb{R}_2^{++}, \quad x_2 \leq x_1, \quad x \in B_{i_1,j_1}(h_x) \subset B_{ij}(h), \quad j \leq i, \tag{4.16}$$

where $h_x = h/2$ and $B_{lm}(r)$ is defined as

$$B_{lm}(r) = [lr, (l+1)r] \times [mr, (m+1)r]. \tag{4.17}$$

Denote by $R(x, k)$ the rectangle covering x

$$R(x,k) \triangleq [(i-k)h, (i+1+k)h] \times [(j-k)h, (j+1+k)h] \tag{4.18}$$

for any $k > 0$. If $k \in Z^+$, the boundary of $R(x, k)$ is along with the mesh grid and is at least kh away from x. Denote by $R_s, R_{s,1}, R_{s,2}$ different symmetric rectangles with respect to x

$$\begin{aligned} R_s(x,k) &\triangleq [x_1 - kh, x_1 + kh] \times [x_2 - kh, x_2 + kh], \\ R_{s,1}(x,k) &\triangleq [x_1 - kh, x_1 + kh] \times [(j-k)h, (j+1+k)h], \\ R_{s,2}(x,k) &\triangleq [(i-k)h, (i+1+k)h] \times [x_2 - kh, x_2 + kh]. \end{aligned} \tag{4.19}$$

We have $R_s(x,k) \subset R_{s,i}(x,k) \subset R(x,k), i = 1, 2$. We introduce the upper and lower parts of $R(x, k)$:

$$R^+(x,k) \triangleq R(x,k) \cap \{y : y_2 \geq x_2\}, \quad R^-(x,k) \triangleq R(x,k) \cap \{y : y_2 \leq x_2\}. \tag{4.20}$$

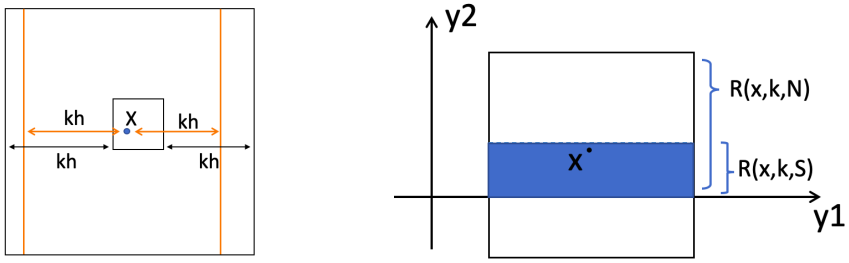


FIG. 1. *Left: The large box is $R(x,k)$ and the red box is $R_{s,1}(x,k)$. (Color images are available online.) The small box containing x has size $h \times h$. Right: The upper box is $R(x,k,N)$, and the shaded box is $R(x,k,S)$, the reflection of the region below the y-axis.*

We use similar notation for $R_s(x,k), R_{s,1}(x,k), R_{s,2}(x,k)$. We further introduce the intersection of the rectangle and four half planes with reflection:

$$\begin{aligned} &R(x,k,N) = R(x,k) \cap \{y : y_2 \geq 0\}, \quad R(x,k,S) = \mathcal{R}_2(R(x,k) \cap \{y : y_2 \leq 0\}), \\ &R(x,k,E) = R(x,k) \cap \{y : y_1 \geq 0\}, \quad R(x,k,W) = \mathcal{R}_1(R(x,k) \cap \{y : y_1 \leq 0\}), \end{aligned} \tag{4.21}$$

where N, E, S, W are short for *north, east, south, west,* respectively, and the reflection operators $\mathcal{R}_1, \mathcal{R}_2$ are given by

$$\mathcal{R}_1(y_1, y_2) = (-y_1, y_2), \quad \mathcal{R}_2(y_1, y_2) = (y_1, -y_2).$$

It is clear that $R(x,k,S) \subset \mathbb{R}^+_2, R(x,k,W) \subset \{y : y_1 \geq 0\}$. An illustration of these domains is given in Figure 1. If $x, y \in \mathbb{R}^{++}_2$, we have the equivalence

$$\begin{aligned} &(y_1, -y_2) \notin R(x,k) \iff (y_1, -y_2) \notin R(x,k) \cap \{y : y_2 \leq 0\} \iff y \notin R(x,k,S), \\ &(y_1, -y_2) \in R(x,k) \iff y \in R(x,k,S). \end{aligned} \tag{4.22}$$

The above notation will be very useful in our later decomposition of the symmetrized kernel.

Define the odd extension of ω in y from $\mathbb{R}^+_2$ to $\mathbb{R}_2$:

$$W(y) = \omega(y) \text{ for } y_2 \geq 0, \quad W(y) = -\omega(y_1, -y_2) \text{ for } y_2 < 0. \tag{4.23}$$

W is odd in both y_1 and y_2 variables. Since we fix x (4.16) below, for simplicity, we drop x in the R notation. For $k > k_2, k, k_2 \in Z^+$, we decompose the weighted $u_x(x)$ integral as follows:[2]

$$\begin{aligned} \psi(x) \int K_1(x-y) W(y) dy &= \psi(x) \int_{R(k)^c} K_1(x-y) W(y) dy \\ &+ \int_{R_{s,1}(k)} K_1(x-y) \psi(y) W(y) dy + \int_{R(k) \setminus R_{s,1}(k)} K_1(x-y) \psi(y) W(y) dy \\ &+ \int_{R(k) \setminus R(k_2)} K_1(x-y) (\psi(x) - \psi(y)) W(y) dy \\ &+ \int_{R(k_2)} K_1(x-y) (\psi(x) - \psi(y)) W(y) dy \\ &\triangleq I_1(x,k) + I_2(x,k) + I_3(x,k) + I_4(x,k,k_2) + I_5(x,k_2), \end{aligned} \tag{4.24}$$

[2]Since we have no flow boundary condition for the velocity and stream function $-\Delta\phi = \omega, \phi(0,y) = 0$, the Poisson integral formula for $\mathbf{u} = \nabla^\perp \phi$ is equivalent to $\nabla^\perp(-\Delta_{2D})^{-1} W$ for Δ_{2D} defined in the whole space case.

where

$$K_1(s) \equiv \frac{s_1 s_2}{|s|^4}.$$

We drop $-\frac{1}{\pi}$ in the integrand $-\frac{1}{\pi}K_1(s)$ for $u_x(x)$ (4.1) at this moment to simplify the notation. We will estimate different parts in section 4.3.

Principle and log-Lipschitz integrand. Our major motivation for the decomposition (4.24) and the integrand (4.28) with symmetrization is to obtain an integrand J which is at least locally log-Lipschitz satisfying $J(x,y) \lesssim |x-y|^{-1}$ for y close to x, and the more singular one J_S. We will estimate the integral of J using the trapezoidal rule in section 4.1.3 and J_S analytically.

4.1.5. Symmetrization. After we obtain the decomposition, we use the odd symmetry of W in y_1, y_2 to symmetrize the integral and reduce the integral over $\mathbb{R}_2$ to the first quadrant $\mathbb{R}_2^{++}$. This enables us to exploit the cancellation in the integral and obtain a sharper estimate. In our computation, we symmetrize the integrals in $I_1(x,k)$ and $I_4(x,k,k_2)$, which are more regular. For a given kernel $K(x,y)$, we denote by K^{sym} the symmetrization of K:

$$K^{sym}(x,y) \triangleq K(x,y) - K(x,-y_1,y_2) - K(x,y_1,-y_2) + K(x,-y). \tag{4.25}$$

We show how to symmetrize $I_1(x,k)$ as an example. Recall the notation in (4.21), (4.16). We assume $x_1 \geq x_2$. We choose $k < i$ so that $R(x,k) \subset \{y : y_1 > 0\}$ and $R(x,k,W) = \emptyset$. By definition (4.18), the domains $R(x,k), R(x,k,N), R^+(x,k)$, etc., are the same for all $x \in B_{i_1,j_1}(h_x)$. Yet, $R(x,k)$ may cross the boundary $y_2 = 0$, i.e., $R(x,k,S) \neq \emptyset$. See the right figure in Figure 1 for a possible configuration. Using the equivalence (4.22) and the property that W is odd in y_1 and y_2, for general $x \in \mathbb{R}_2^{++}$ (without $x_1 \geq x_2$), we can symmetrize $I_1(x,k)$ as follows:

$$\begin{aligned} I_1(x,k) = \psi(x) \int_{\mathbb{R}_2^{++}} \Big(& K_1(x-y)\mathbf{1}_{y\in R(k)^c} - K_1(x_1-y_1, x_2+y_2)\mathbf{1}_{y\notin R(k,S)} \\ & - K_1(x_1+y_1, x_2-y_2)\mathbf{1}_{y\notin R(k,W)} + K_1(x+y)\Big)\omega(y)dy. \end{aligned} \tag{4.26}$$

For $I_4(x)$ (4.24), we choose the weight $\psi(y)$ (A.1), (A.2) even in y_1, y_2. Then the symmetrization of I_4 is

$$\begin{aligned} I_4(x,k,k_2) = \int_{\mathbb{R}_2^{++}} & (K_1(x-y)\mathbf{1}_{y\in R(k)\setminus R(k_2)} - K_1(x_1-y_1, x_2+y_2)\mathbf{1}_{y\in R(k,S)\setminus R(k_2,S)} \\ & - K_1(x_1+y_1, x_2-y_2)\mathbf{1}_{y\in R(k,W)\setminus R(k_2,W)})(\psi(x)-\psi(y))W(y)dy. \end{aligned} \tag{4.27}$$

In (4.27), we do not have the term $K_1(x+y)$ since for $y \in \mathbb{R}_2^{++}$, $x+y \geq x_c > (k+1)h$ and $-y \notin R(k)$. See the discussion below (4.15). Thus after symmetrizing the kernel in I_4, we do not have such a term.

Though the symmetrized kernel is complicated, since these regions $R(l), R(l,\alpha)$ $\alpha = N, E, l = k, k_2$ (4.18), (4.21) can be decomposed into the union of the mesh girds $[y_i, y_{i+1}] \times [y_j, y_{j+1}]$, in each grid, the indicator functions are constants. See also Remark 4.6. In each grid $y \in [y_i, y_{i+1}] \times [y_j, y_{j+1}]$, we can write the integrand in $I_1 + I_4$ as

$$\begin{aligned} J &= K^{NC}(x,y)\cdot\psi(x) + K^C(x,y)\cdot(\psi(x)-\psi(y)), \\ \partial_{x_i} J &= (K^{NC}+K^C)\partial_{x_i}\psi(x) + \partial_{x_i}K^{NC}\cdot\psi(x) + \partial_{x_i}K^C(x,y)\cdot(\psi(x)-\psi(y)), \end{aligned} \tag{4.28}$$

where NC, C are short for *noncommutator, commutator,* respectively.

For y close to x, J is at least locally log-Lipschitz. See the principle before section 4.1.5 for motivation. For y away from x, e.g., $|y_1| \vee |y_2| \geq 4x_c$ in our computation, we have

$$J = K^{sym}(x,y)\psi(x). \tag{4.29}$$

In practice, we assemble the symmetrized integrand in $I_1 + I_4$ in $\mathbb{R}^{++}_2$ together. Using (4.28), we only need to assemble K^{NC}, K^C. We first initialize the integrand with $(K^{NC}, K^C) = (K^{sym}, 0)$. To assemble the integrand in the singular regions, we perform two replacements. In the first replacement, we pretend that $R(k_2) = \emptyset$ and replace the integrand in $R(k) \cap \mathbb{R}^{++}_2$. Based on $x \in B_{ij}(h)$ (4.16), we determine the regions $R(x,k), R(x,k,S)$ (4.18), (4.21). Since $x_1 \geq x_2$, we get $R(x,k,W) = \emptyset$. See Figure 1. We partition $R(k) \cap \mathbb{R}^{++}_2$ as follows:

$$R(k) \cap \mathbb{R}^{++}_2 = R(k,N) = (R(k,N) \backslash R(k,S)) \cup R(k,S) \triangleq D_1 \cup D_2. \tag{4.30}$$

According to (4.26), (4.27) ($R(k_2) = \emptyset$), for $i = 1, 2$, we first replace (K^{NC}, K^C) in D_i by

$$\begin{aligned} (K^{NC}, K^C) &= (K^{sym} - K^C_i, K^C_i),\ K^C_1 = K_1(x-y),\ K^C_2 \\ &= K_1(x-y) - K_1(x_1 - y_1, x_2 + y_2), \end{aligned} \tag{4.31}$$

respectively, where K^C is from the integrand in (4.27). We have i singular terms in D_i in (4.27).

In the second replacement, we replace the integrand in the smaller singular region $R(k_2) \cap \mathbb{R}^{++}_2 \subset R(k) \backslash \mathbb{R}^{++}_2$. Outside this region, we have obtained the symmetrized integrand using (4.31). Since we assume $x_1 \geq x_2$, we get $R(k,W) = \emptyset$ (see the discussion below (4.25)) and $\mathbf{1}_{y \notin R(k,W)} \equiv 1, \mathbf{1}_{y \in R(k,W)} = 0$. Similarly to $R(k) \cap \mathbb{R}^{++}_2$ (see Figure 1), we can decompose

$$R(k_2) \cap \mathbb{R}^{++}_2 = (R(k_2, N) \backslash R(k_2, S)) \cap R(k_2, S) \triangleq D_3 \cup D_4.$$

In $D_4 = R(k_2, S) \subset R(k_2), R(k, S)$, from (4.26), (4.27), we completely remove the $K_1(x-y), K_1(x_1 - y_1, x_2 + y_2)$ terms in the integrand and have

$$(K^{NC}, K^C) = (K_1(x+y) - K_1(x_1 + y_1, x_2 - y_2),\ 0).$$

In D_3, since $D_3 \subset R(k,N) = D_1 \cup D_2$ (4.30), there are two cases. In $D_3 \cap D_1, D_1 = R(k,N) \backslash R(k,S)$, we have three nonsingular terms from (4.26) and zero terms from (4.27) and get

$$(K^{NC}, K^C) = (K_1(x+y) - K_1(x_1 + y_1, x_2 - y_2) - K_1(x_1 - y_1, x_2 + y_2),\ 0).$$

In $D_3 \cap D_2, D_2 = R(k,S)$, we have two terms from (4.26) and one term from (4.27). We get

$$(K^{NC}, K^C) = (K_1(x+y) - K_1(x_1 + y_1, x_2 - y_2), -K_1(x_1 - y_1, x_2 + y_2)).$$

For $x_1 < x_2$, we assemble the integrand similarly. Using (4.28), we obtain the integrand $\partial_{x_i} J$ for the Hölder estimate.

***$C_y^{1/2}$* estimate of *u_y, v_x*.** In the $C_y^{1/2}$ estimate of u_y, v_x with kernel K_2 (4.1), we symmetrize the integrand $K(x-y)(\psi(x)-\psi(y)$; see (4.68) in section 4.3.9. In this case, the symmetrized integrand $W(y)T$ is similar to (4.26) with $\psi(x)$ replaced by $\psi(x)-\psi(y)$ and

$$T = (\psi(x)-\psi(y))\Big(K_2(x-y)\mathbf{1}_{y\in R(k)^c} - K_2(x_1-y_1,x_2+y_2)\mathbf{1}_{y\notin R(k,S)} \\ - K_2(x_1+y_1,x_2-y_2)\mathbf{1}_{y\notin R(k,W)} + K_1(x+y)\Big).$$

Due to the weight $(\psi(x)-\psi(y))$, we always have $K^{NC}=0$. We initialize the T using (4.28) with $K^C = K_2^{sym}$ (4.25). In the singular region $R(x,k)\cap\mathbb{R}_2^{++}$, we only need to perform one replacement. Similar to (4.31), we use (4.30) and replace the integrand as follows:

$$K^C = K_2^{sym} - K_2(x-y), y\in R(k,N)\backslash R(k,S),$$
$$K^C = K_2^{sym} - (K_2(x-y)-K_2(x_1-y_1,x_2+y_2)), y\in R(k,S).$$

We remove the most singular integrand in $R(k,N)\backslash R(k,S)$ and the most two singular integrands in $D_2=R(k,S)$ to make T locally log-Lipschitz. See the principle before section 4.1.5.

***L^∞* estimate.** For the L^∞ estimate, we do not multiply the integrand by the weight $\psi(x)$ or the commutator. We decompose the integral as (4.45) and symmetrize the nonsingular part in I_1 using (4.26) without the weight $\psi(x)$. Symmetrizing I_4 (4.45) is similar. We initialize the symmetrized integrand as K^{sym} (4.25) and then replace it in $R(k)\cap\mathbb{R}_2^{++}$. Without loss of generality, we assume $x_1\geq x_2$ and have the decomposition (4.30). Similar to (4.31), we replace the integrand as follows:

$$K^{sym} - K_1(x-y), y\in R(k,N)\backslash R(k,S),$$
$$K^{sym} - (K_1(x-y)-K_1(x_1-y_1,x_2+y_2)), y\in R(k,S).$$

That is, we remove one and two singular terms in $R(k,N)\backslash R(k,S), R(k,S)$, resepctively, making the integrand at least locally log-Lipschitz. See the principle before section 4.1.5.

4.1.6. Integral in domains depending on *x*. In the computation, we need to estimate several integrals in the domains $D(x)$ depending on x, e.g., I_3 in (4.24). Our fundamental idea is to cover $D(x)$ by some piecewise constant domains, which will be essentially treated as fixed domains. By refining the location of x, we can obtain tight covering.

We use the L^∞ estimate of I_3 to illustrate the ideas. A direct estimate yields

$$|I_3(x)| \leq ||W\varphi||_\infty \int_{R(k)\backslash R_{s,1}(k)} |K_1(x-y)|\psi(y)\varphi^{-1}(y)dy.$$

We cannot apply the method in section 4.1.3 to first estimate $I_3(x)$ for x on the grid points and then estimate $\partial^2 I_3(x)$ for the error since the kernel is singular and the error part associated with $\partial^2 I_3(x)$ is more singular (see Lemma 4.4).

Denote $f=\psi\varphi^{-1}$. We consider a change of variable $y=x+s$ to center our analysis around the singularity x. The domain for s is

$$\{y\in R(k)\backslash R_{s,1}(k)\} = \{s\in R(k)-x\}\cap\{|s_1|\geq kh\} \triangleq D(x,k). \tag{4.32}$$

It suffices to estimate

$$J=\int_{s\in D(x,k)}|K_1(-s)|f(x+s)dy,\quad f\geq 0, \tag{4.33}$$

for all $x\in B_{i_1,j_1}(h_x)$ (4.16). We want to further simplify the above domain so that it does not depend on x. Recall the location of x (4.16). To obtain a sharp estimate, we further partition the location of $x\in B_{i_1,j_1}(h_x)$ as follows:

$$A_a=[i_1h_x+ah_x/m,i_1h_x+(a+1)h_x/m],\ B_b\triangleq[j_1h_x+bh_x/m,j_1h_x+(b+1)h_x/m] \tag{4.34}$$

for some $m\in Z^+$ and $0\leq a,b\leq m-1$. Clearly, $A_a\times B_b$ is a partition of $B_{i_1j_1}(h_x)$. Recall (4.16) and (4.18). We have

$$R(x,k)=[(i-k)h,(i+1+k)h]\times[(j-k)h,(j+1+k)h].$$

Now, for $x\in A_a\times B_b$, since $|s_1|\geq kh$, we have

$$\begin{aligned}s_1=y_1-x_1&\in[(i-k)h-i_1h_x-(a+1)h_x/m,-kh]\cup[kh,(i+1+k)h-i_1h_x-ah_x/m]\\&\triangleq X_{l,a}\cup X_{r,a},\end{aligned} \tag{4.35}$$

where the subscripts l, r are short for left, right, respectively. Similarly, for s_2, we have

$$\begin{aligned}s_2=y_2-x_2&\in[(j-k)h-j_1h_x-(b+1)h_x/m,(j+k+1)h-j_1h_x-bh_x/m]\\&\triangleq[(j-k)h-j_1h_x-(b+1)h_x/m,-kh]\cup[-kh,kh]\\&\qquad\cup[kh,(j+1+k)h-j_1h_x-bh_x/m]\\&\triangleq Y_{d,b}\cup Y_{m,b}\cup Y_{u,b},\end{aligned} \tag{4.36}$$

where the subscripts d, m, u are short for down, middle, upper, respectively. Note that the intervals X,Y do not depend on x. We have

$$D(x,k)\subset(X_{l,a}\cup X_{r,a})\times(Y_{d,b}\cup Y_{m,b}\cup Y_{u,b}). \tag{4.37}$$

Now, we can decompose J (4.33) as follows:

$$J\leq\sum_{\alpha=l,r,\beta=d,m,u}J_{\alpha,\beta},\quad J_{\alpha,\beta}\triangleq\int_{X_{\alpha,a}\times Y_{\beta,b}}|K_1(-s)|f(s+x)dy,\ \alpha=l,r,\ \beta=d,m,u.$$

See the left figure in Figure 2 for different domains in the above decomposition. From the definitions of X,Y, the total width of the left and the right domains $X_{\alpha,a}\times(Y_{d,b}\cup Y_{m,b}\cup Y_{u,b}),\alpha=l,u$ is

$$|X_{l,a}|+|X_{r,a}|=h+h_x/m.$$

For a fixed x, from the definition (4.18), the width of $R(k)\backslash R_{s,1}(k)$ is h. We choose a large m and further partition the location of x so that we do not overestimate the region too much.

For a small domain $Q=[a,b]\times[c,d]$, we can estimate the integral as follows:

$$\int_Q|K_1(-s)|f(x+s)ds\leq\int_Q|K_1(-s)|ds||f||_{L^\infty(B_{i_1j_1}(h_x)+Q)}. \tag{4.38}$$

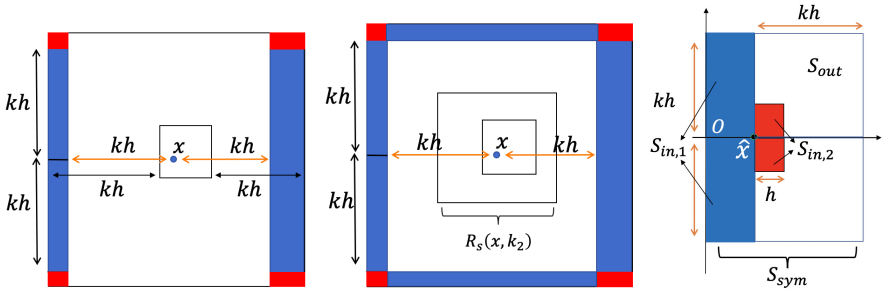


FIG. 2. *The largest box in the left and middle figure is $R(x,k)$. Left: The left and right blue regions are $X_{l,a}\times Y_{m,b}, X_{r,a}\times Y_{m,b}$. The four red regions correspond to $X_{\alpha,a}\times Y_{\beta,b}, \alpha=l,u,\beta=d,u$. Middle: Illustration of $R(x,k)\backslash R_s(x,k)$ and $R_s(x,k_2)$. $R(x,k)\backslash R_s(x,k)$ consists of the blue and the red regions. Right: different regions near the singularity for u/x_1. Blue, red, and white regions represent $S_{in,1}, S_{in,2}, S_{out}$, respectively.*

Since Q is given, $K_1(s)$ is explicit and has scaling symmetries, and we can estimate the integral of $|K_1(s)|$ easily. For example, if $Q=[ah,bh]^2$, we can use the scaling symmetries of $K_1(s)$ to obtain $\int_Q |K_1(-s)| = h^\beta \int_{[a,b]^2} |K_1(-s)|$ for some β. Moreover, for many kernels in our computations, e.g., $K(s)=\frac{s_1 s_2}{|s|^4}$, we have explicit formulas for the integral. See section 5.1 in the supplementary material (supplement.pdf [local/web 1.43MB]).

We apply the above method to estimate the integral in $X_{\alpha,a}\times Y_{\beta,b}, \alpha=l,r,\beta=d,u$ (red region in Figure 2; color images are available online). Since $Y_{m,b}=[-kh,kh]$, for the integral in $X_{\alpha,a}\times Y_{m,b}$ (blue region), we further decompose it,

$$J_{\alpha,m} = \sum_{-k\le t\le k-1} \int_{X_{\alpha,a}\times[th,(t+1)h]} |K_1(-s)| f(s+x)dy, \tag{4.39}$$

and then apply the above method to estimate it.

Next, we further simplify $||f||_{L^\infty(B_{i_1j_1}(h_x)+Q)}$ in the above estimate. From (4.16), we get

$$ih \le i_1 h_x < (i_1+1)h_x \le (i+1)h, \quad jh \le j_1 h_x < (j_1+1)h_x \le (j+1)h.$$

For $X_{l,a}$ (4.35) with $0\le a\le m-1$, we have the lower bound for the endpoint

$$(i-k)h - i_1 h_x - (a+1)h_x/m \ge (i-k)h - i_1 h_x - h_x \ge (i-k)h - ((i+1)h - h_x) - h_x$$
$$= -kh - h.$$

See the left figure in Figure 2. The width of the blue region is less than h. Similarly, we can cover the intervals of X,Y (4.35), (4.36) uniformly for $0\le a,b\le m-1$ and obtain

$$X_{l,a} \subset [(i-k)h - i_1 h_x - h_x, -kh] \subset [-(k+1)h, -kh],$$
$$X_{r,a} \subset [kh, (i+1+k)h - i_1 h_x] \subset [kh, (k+1)h],$$
$$Y_{d,b} \subset [-(k+1)h, -kh], \quad Y_{u,b} \subset [kh, (k+1)h].$$

Thus, we only need to estimate the L^∞ norm of f in

$$Q_{i_1j_1}(h_x) + [\alpha h, (\alpha+1)h]\times[\beta h, (\beta+1)h], \quad \alpha=-k-1,k, \quad \beta=-(k+1),-k,..,k.$$

These estimates are independent of the choice of m, a, b. Since the size of each domain is at most $2h \times 2h$, the above estimates based on (4.38) are sharp. We estimate the piecewise bound of the weights ψ, φ in Appendices A.1, A.2, A.3.

Using the above decomposition and estimates, we obtain the estimate of J (4.33) for $x \in A_a \times B_b$ (4.34). Similarly, we can estimate J for any $0 \le a, b \le m-1$. Taking the maximum of these m^2 estimates, we obtain the estimate of J and $I_3(x)$ for all $x \in B_{i_1 j_1}(h_x)$.

4.1.7. First generalization: Integral in a ring. We generalize the above ideas to estimate the integrals in domain $D = R(x,k)\backslash R(x,k_2) = R(k)\backslash R(k_2)$,

$$J = \int_{R(k)\backslash R(k_2)} |K(y-x)||f(y)|dy = \int_{s\in D(x,k)} |K(s)||f(x+s)dy|,$$
$$D(x,k) \triangleq R(k)\backslash R(k_2) - x$$

with $2 \le k_2 = k - \frac{i}{2} < k$ for some integer $i \ge 1$ and some kernel $K(z)$. Note that the inner region $R(k_2)$ is different from (4.32). See I_4 in (4.24) for an example of this integral region. Suppose $x \in B_{ij}(h)$ (4.16). We partition the location of x similarly to (4.34) and introduce p_l, q_l:

$$\begin{aligned}
&A_a = [ih + ah/m, ih + (a+1)h/m],\\
&B_b = [jh + bh/m, jh + (b+1)h/m],\ 0 \le a, b \le m-1,\\
&p_1 = -k_2 - a/m,\ p_2 = k_2 + (m-a-1)/m,\\
&p_3 = -k_2 - b/m,\ p_4 = k_2 + (m-b-1)/m,\\
&q_1 = -k - (a+1)/m,\ q_2 = k + (m-a)/m,\\
&q_3 = -k - (b+1)/m,\ q_4 = k + (m-b)/m.
\end{aligned} \tag{4.40}$$

For a fixed $x \in A_a \times B_b$, by comparing the boundaries of the following four rectangles, we get

$$D_{in} \triangleq [p_1 h, p_2 h] \times [p_3 h, p_4 h] \subset R(k_2) - x \subset R(k) - x \subset [q_1 h, q_2 h] \times [q_3 h, q_4 h] \triangleq D_{out}.$$

To obtain the above inclusions, for example, for $s = y - x, y \in R(k_2)$, we use

$$\min_{y\in R(k_2)} y_1 - x_1 = ih - k_2 h - x_1 \le ih - k_2 h - (ih + ah/m) = -k_2 h - ah/m = p_1 h,$$

uniformly for $x \in A_a \times B_b \subset B_{ij}(h)$. For $R(k) - x \subset D_{out}$, we have $q_1 h \le \min_{y\in R(k)} y_1 - x_1$. Other bounds for the inclusions are obtained similarly. This yields $D(x,k) \subset D_{ring}$, where

$$D_{ring} \triangleq D_{out}\backslash D_{in} \tag{4.41}$$

is fixed for $x \in A_a \times B_b$.

It suffices to estimate the integral J in D_{ring}. We partition $s \in D_{ring}$ using mesh

$$Z_1 = \{-k \le i \le k, i \in \mathbb{Z}\} \cup \{p_1, p_2, q_1, q_2\}, \quad Z_2 = \{-k \le i \le k, i \in \mathbb{Z}\} \cup \{p_3, p_4, q_3, q_4\}, \tag{4.42}$$

and then order them in an increasing order $z_{l,1} < z_{l,2} < \cdots < z_{l,2k+5} \in Z_l, l = 1, 2$. Note that we do not multiply $z_{l,c}$ by h here. We estimate the integral J_{cd} in each grid $Q_{c,d} = [z_{1,c}h, z_{1,c+1}h] \times [z_{2,d}h, z_{2,d+1}h]$ following (4.38) and using the norm $||f||_{L^\infty(x+Q_{c,d})}$. We turn off the integral in region $Q_{c,d}$ if $Q_{c,d} \subset D_{in}$ since it is not in D_{ring} (4.41).

Uniform covering. For fixed c,d, we want to cover $x+Q_{c,d}$ uniformly for $x \in A_a \times B_b$ and all $0 \le a,b \le m-1$ (the subpartition of x) to bound $||f||_{L^\infty(x+Q_{c,d})}$. Since we add four extra points p_l, q_l in Z_1 and Z_2 and order them in an increasing order, the region $Q_{c,d}$ can change for fixed c,d but with different a,b. We show that the $2k+4$ intervals $[z_{1,c}, z_{1,c+1}], 1 \le c \le 2k+4$ can be covered by $[\alpha_l, \beta_l]$ uniformly for a,b

$$\begin{aligned} &[\alpha_l,\beta_l], \quad \alpha_l \in Z_1^l, \beta_l \in Z_1^u,\ Z_1^l \triangleq \{-(k+1) \le i \le k, i \in \mathbb{Z}\} \cup \{-s_0-2, s_0\}, \\ &Z_1^u \triangleq \{-k \le i \le k+1, i \in \mathbb{Z}\} \cup \{-s_0, s_0+2\}, \quad s_0 = \lfloor k_2 \rfloor, \end{aligned} \tag{4.43}$$

with α_l, β_l increasing. From (4.40) and the definition of s_0, we get

$$p_1 \in [-s_0-2, -s_0],\ p_2 \in [s_0, s_0+2],\ q_1 \in [-k-1, -k],\ q_2 \in [k, k+1]. \tag{4.44}$$

The uniform covering is based on the following observations. Suppose that $u_i \le v_i, i = 1,2,\ldots,n$ (u_i, v_i may not be increasing). Let us denote by $\{U_i\}$ the reordering of $\{u_i\}$ in an increasing order and denote by $\{V_i\}$ the reordering of $\{v_i\}$ in an increasing order. Then we have $U_i \le V_i$. In fact, for any $k \le n$, from $u_i \le v_i$, V_k is larger than u_j with at least k different indexes j. Since U_k is the k-smallnest value in $\{u_i\}_i$, we get $V_k \ge U_k$.

From (4.42), (4.44), since $q_2 = \max_c z_{1,c}, q_1 = \min_c z_{1,c}$, we get

$$\begin{aligned} &\{z_{1,c}, c \le 2k+4\} = \{-k \le i \le k, i \in \mathbb{Z}\} \cup \{p_1, p_2, q_1\}, \\ &-k-1 \le q_1, -s_0-2 \le p_1, s_0 \le p_2, \\ &\{z_{1,c+1}, c \le 2k+4\} = \{-k \le i \le k, i \in \mathbb{Z}\} \cup \{p_1, p_2, q_2\}, \\ &p_1 \le -s_0,\ p_2 \le s_0+2, q_2 \le k+1. \end{aligned}$$

We can bound each component in Z_1^l (4.43) by a component in the above list. Using the above observations, after reordering two sequences in an increasing order, which gives $\{\alpha_c\}, \{z_{1,c}\}_{c \le 2k+4}$, we get $\alpha_c \le z_{1,c}, c \le 2k+4$ (4.43). Similarly, we obtain $z_{1,c+1} \le \beta_c$, yielding $[z_{1,c}, z_{1,c+1}] \in [\alpha_c, \beta_c], c \le 2k+4$.

Similarly, we obtain $[z_{2,d}, z_{2,d+1}] \subset [\alpha_d, \beta_d]$. Thus, we get $[z_{1,c}, z_{1,c+1}] \times [z_{2,d}, z_{2,d+1}] \in [\alpha_c, \alpha_{c+1}] \times [\beta_d, \beta_{d+1}]$ uniformly for the subpartition of $x \in A_a \times B_b$ with $0 \le a,b \le m-1$, and can cover $x+Q_{cd}$ by $B_{i_1 j_1}(h_x) + [\alpha_c h, \alpha_{c+1} h] \times [\beta_d h, \beta_{d+1} h]$ (4.16) for $x \in B_{i_1 j_1}(h_x) \subset B_{ij}(h)$.

4.1.8. Second generalization: The boundary terms. We generalize the method to estimate some boundary terms. We estimate the x_1-derivative of $I_3(x)$ (4.24) to illustrate the ideas. In $\partial_1 I_3$, we have an extra boundary term I_{32}

$$\begin{aligned} \partial_1 I_3(x) = &\int_{R(k) \setminus R_{s,1}(k)} \partial_{x_1} K_1(x-y)(W\psi)(y) dy \\ &- \int_{(j-k)h}^{(j+1+k)h} K_1(x-y)(W\psi)(y)\Big|_{y_1 = x_1 - kh}^{x_1+kh} dy_2 \triangleq I_{31} + I_{32}, \end{aligned}$$

where we have used the domain for $R(x,k)$ (4.18).

For I_{31}, we apply the method in section 4.1.6 to estimate it. Denote $\Gamma_k \triangleq [j-k)h, (j+1+k)h]$. Using a change of variable $y = x+s$, we can rewrite I_{32} as follows:

$$\begin{aligned} I_{32} = -\int_{s_2 \in \Gamma_k - x_2} &(K_1(-kh, -s_2)(W\psi)(x_1+kh, x_2+s_2) \\ &- K_1(kh, -s_2)(W\psi)(x_1-kh, x_2+s_2)) ds_2. \end{aligned}$$

We partition the location of x and assume $x \in A_a \times B_b \subset B_{i_1,j_1}(h_x)$ (4.34). From (4.36), we have

$$s_2 \in \Gamma_k - x_2 \subset Y_{d,b} \cup Y_{m,b} \cup Y_{u,b}.$$

Using the above decomposition and $|W\psi(x)| \le ||W\varphi||_\infty f(x)$, $f = \psi\varphi^{-1}$, we obtain

$$|I_{32}| \le ||W\varphi||_\infty \sum_{\alpha=\pm,\beta=d,m,u} M_{\alpha,\beta},$$
$$M_{\alpha,\beta} \triangleq \int_{Y_{\beta,b}} |K_1(-\alpha kh, -s_2)| \cdot |f(x_1 + \alpha kh, x_2 + s_2)| ds_2$$

for $\alpha = \pm, \beta = u, m, d$. For $\beta = u, d$, the domain $Y_{\beta,b}$ is small $|Y_{\beta,b}| \le h$. We apply the method in (4.38) to estimate $M_{\alpha,\beta}$. The only difference is that we need consider a 1D integral here

$$\int_Q |K_1(-\alpha kh, -s_2)| ds_2$$

for some interval Q, rather than a 2D integral in (4.38). For $M_{\alpha,m}$, we decompose the domain $Y_{m,b}$ into small intervals with length h similar to (4.39) and then apply the method in (4.38).

We combine these estimates to bound I_{32} for $x \in A_a \times B_b$. Then, we maximize the estimates over $0 \le a, b \le m-1$ to bound I_{32} for $x \in B_{i_1,j_1}(h_x)$.

4.1.9. Third generalization. In some of the computations, we need to estimate

$$J = \int_{R(k)\backslash R_s(k_2)} |K(x-y)| f(y) dy$$

for some $k_2 < k$ with $2k_2, k \in Z^+$, where $R_s(k)$ is defined in (4.19). Similarly, we use

$$R_s(k_2) \subset R_s(k) \subset R(k), \quad R(k)\backslash R_s(k_2) = R(k)\backslash R_s(k) \cup R_s(k)\backslash R_s(k_2),$$

and a change of variable $y = x + s$ to obtain

$$J = \left(\int_{s \in R(k)-x, |s_1| \vee |s_2| \ge kh} + \int_{k_2 h \le |s_1| \vee |s_2| \le kh} \right) K(-s) f(x+s) ds \triangleq J_1 + J_2.$$

Compared to $R(k)\backslash R_{s,1}(k)$, the domain $R(k)\backslash R_s(k)$ contains two more parts,

$$X_{m,a} \triangleq [-kh, kh], \quad X_{m,a} \times Y_{u,b}, \quad X_{m,a} \times Y_{d,b},$$

i.e., the upper and lower blue regions in the right figure in Figure 2. The integral in these regions is estimated to be similar to that in $X_{\alpha,a} \times Y_{m,b}$ (4.37), and the estimate of J_1 is similar to J in (4.33).

For J_2, the domain is simpler. Since $2k_2 \in Z^+$, we partition the domain into $h_x \times h_x$ grids,

$$J_2 = \sum_{(c,d) \in S_k \backslash S_{k_2}} \int_{[ch_x,(c+1)h_x] \times [dh_x,(d+1)h_x]} |K(-s)| f(s+x) ds,$$
$$S_l \triangleq \{-k \le c < k, -k \le d < k\}.$$

For each integral, we estimate it using the method in (4.38). The remaining steps are the same as those of J in (4.33) studied previously.

Remark 4.5. In the estimates in sections 4.1.6–4.1.9, we use the important property that the weights are locally smooth to move them outside the integral. Moreover, we use the fact that the singular region depends on x monotonously to cover it effectively. Since the integral $\int_Q |K_1(s)|dy$ for different Q, a, b in the above estimates does not depend on x, we first compute these integrals once and store them and then use them in later estimate of different x.

4.1.10. Taylor expansion near the singularity. We need to estimate the integral

$$J(x) \triangleq \int_D \partial_{x_i}\Big(K(x-y)(\psi(x)-\psi(y))W(y)\Big)dy$$

for $k_2 < k$ in some region D close to the singularity x. For example, $D = R(x,k_2)\backslash R(x,k_3)$, $R(x,k_3)\backslash R_{s1}(x,k_3)$ in $\partial_{x_i} I_{5,0}, \partial_{x_i} I_{5,1}$ (4.51). To obtain a sharp estimate, we perform Taylor expansion on $\psi(x)$. We focus on ∂_{x_1}. Denote $z = x-y, x_m = \frac{x+y}{2}$. A direct computation yields

$$I = \partial_{x_1}(K(x-y)\psi(x)-\psi(y)) = (\partial_1 K)(x-y)(\psi(x)-\psi(y)) + K(x-y)\partial_1\psi(x).$$

Using Taylor expansion of ψ at x_m and following (B.26), we get

$$\begin{aligned}
&\psi(x)-\psi(y) = (x-y)\cdot\nabla\psi(x_m)+\varepsilon_1, \quad \psi_x(x) = \psi_x(x_m)+\varepsilon_2,\\
&|\varepsilon_1| \le \sum_{i+j=2} c_{ij}||\partial_x^i\partial_y^j\psi||_{L^\infty(Q(y))}|z_1|^i|z_2|^j,\\
&|\varepsilon_2| \le \frac{1}{2}(||\partial_{xx}\psi||_{L^\infty(Q(y))}|z_1| + ||\partial_{xx}\psi||_{L^\infty(Q(y))}|z_2|),
\end{aligned}$$

where $c_{20} = \frac{1}{4}, c_{11} = \frac{1}{2}, c_{02} = \frac{1}{4}$, and we have written $z_i = x_i - y_i$ and $Q(y)$ is one of the four quadrants $D \cap \{y : sgn(y_i - x_i) = \pm 1\}$ covering both x, y. Combining the term with the same derivative of ψ, we need to estimate the following integrals:

$$\begin{aligned}
&\left|\int_D \psi_x(x_m)(\partial_1 K(z)z_1 + K(z))W(y)dy\right|, \quad \left|\int_D \psi_y(x_m)\partial_1 K(z)z_2 W(y)dy\right|,\\
&\int_D |\partial_x^i\partial_y^j\psi|_{L^\infty(Q(y))}|\partial_1 K(z)z_1^i z_2^j W(y)|dy, i+j=2,\\
&\int_D |\partial_x^{i+1}\partial_y^j\psi|_{L^\infty(Q(y))}|K(z)z_1^i z_2^j W(y)|dy, i+j=1.
\end{aligned}$$

We partition the region of $z = x-y \in x-D$, e.g., $D = R(k_2)\backslash R(k_3)$ (4.51), into small mesh and estimate the piecewise bounds of weights and each integral following sections 4.1.6–4.1.9.

We estimate the integral of $|\partial_1^i\partial_2^j K(z)z_1^k z_2^l|$ in section 5.1 in the supplementary material (supplement.pdf [local/web 1.43MB]).

4.1.11. Hölder estimate of log-Lipschitz function. In some computation, we need to perform a $C^{1/2}$ estimate of some log-Lipschitz function. We consider an example to illustrate the ideas:

$$\begin{aligned}
F(x) = &\int_{\max_i |x_i-y_i|\le b} K(x,y)f(y)dy, \quad |K(x,y)| \le C_1|x-y|^{-1},\\
&|\partial K(x,y)| \le C_2|x-y|^{-2},
\end{aligned}$$

for some constant C_1, C_2. Given $f \in L^\infty$, F is log-Lipschitz. To estimate $[f]_{C_x^{1/2}}$, we cannot first estimate the piecewise values of f and $\partial_x f$ and then combine them to obtain the $C_x^{1/2}$ estimate. Instead, given x, z, for a to be determined, we decompose F into the smooth part and the singular part

$$F_1(x) \triangleq \int_{a \le \max_i |x_i - y_i| \le b} K(x,y) f(y) dy, \quad F_2(x) \triangleq \int_{\max_i |x_i - y_i| \le a} K(x,y) f(y) dy.$$

Using the assumptions of the kernel, we have

$$|\partial_{x_1} F_1(x)| \le C_3 \log \frac{b}{a} ||f||_\infty, \quad |F_2(x)| \le C_4 |a| \cdot ||f||_\infty,$$

where the constants C_3, C_4 depend on b, C_1, C_2. Applying the above estimates, we obtain

$$\begin{aligned} \frac{|F(x) - F(z)|}{|x_1 - z_1|^{1/2}} &\le \frac{|F_1(x) - F_1(z)| + |F_2(x) - F_2(z)|}{|x_1 - z_1|^{1/2}} \\ &\le \Big(C_3 \log \frac{b}{a} \cdot |x_1 - z_1|^{1/2} + 2C_4 |a| |x_1 - z_1|^{-1/2} \Big) ||f||_\infty. \end{aligned}$$

We optimize the estimates by choosing $a = C_5 |x_1 - z_1|$ for some constant C_5 depending on C_3, C_4. Then we establish the estimate. The above simple estimates show that the choice of a depends on $|x - z|$. Thus, in our later Hölder estimates, we perform decomposition guided by the above estimates and optimize the choice of size of the singular region $[-a, a]^2$. On the other hand, since for different $|x - z|$ we need to choose different a, it increases the technicality of the computer-assisted estimates.

4.2. L^∞ estimate. Let $\widehat{u_x}$ be the approximation term of u_x (see section 4.3 of Part I [13]). We focus on the estimate of the piecewise L^∞ norm of $u_{x,A} = u_x - \widehat{u_x}$, which is a representative case. For simplicity, we assume the rescaling factor $\lambda = 1$. We assume that x satisfies (4.16) without loss of generality. We want to estimate $u_{x,A}$ for *all* $x \in B_{i_1 j_1}(h_x)$.

We can write $u_{x,A} = u_x - \hat{u}_x$ as follows:

$$u_{x,A} = \int (K(x-y) - \hat{K}(x,y)) W(y) dy, \quad K_A \triangleq K(x-y) - \hat{K}(x,y),$$

where $\hat{K}(x,y)$ is the kernel for the approximation term and W is the odd extension of ω (see (4.23)). From sections 4.3.2 and 4.3.3 of Part I [13], we remove the singular part in $\hat{K}$, and then $\hat{K}$ is nonsingular. Given x with (4.16), similar to (4.24), for $k \ge k_2$, we perform the following decomposition:

(4.45)

$$\begin{aligned} u_{x,A} &= \left(\int_{R(k)^c} + \int_{R(k) \setminus R_s(k_2)} + \int_{R_s(k_2)} \right) K(x-y) W(y) dy - \int \hat{K}(x,y) W(y) dy \\ &\triangleq I_1 + I_2 + I_3 + I_4, \end{aligned}$$

where $R_s(k)$ is the symmetric singular region (4.19). See section 4.2.3 for the choice of k.

Since $I_1 + I_4$ is nonsingular, we use the ideas in section 4.1.5 to symmetrize the kernels in $I_1 + I_4$. Then we use the method in section 4.1.3 to estimate it.

Remark 4.6. In our computation, the domain $[0, D]^2 \cap R(k)^c$ can be decomposed into the union of small grids $[y_i, y_{i+1}] \times [y_j, y_{j+1}]$ (4.11) since the boundary of $R(x,k)$

aligns with the mesh (4.18). In particular, in each grid, the indicator function is constant, and the integrand is smooth in y.

Next we consider I_2. The domain of the integral is close to the singularity. If we use the method in section 4.1.3 to estimate it, the error will be quite large since $\partial^2 K(x-y)$ is very singular. We want to estimate I_2 using $||W\varphi||_\infty$ and the singular part I_3 using $[W\psi_1]_{C^{1/2}}$. Since $K(z)$ is singular of order -2, we expect an estimate

$$|I_2| + |I_3| \lesssim \log \frac{k}{k_2} \varphi^{-1}(x) ||W\varphi||_{L^\infty[R(k)]} + \psi^{-1}(x) k_2^{1/2} [W\psi]_{C_x^{1/2}}.$$

Note that the weights φ, ψ have a different order of singularity for small x and a different rate of decay. Moreover, we need to control the right hand side using the energy, which assigns different weights to two norms (seminorms). Thus, to obtain a sharp estimate, we need to optimize the choice of k_2.

First, we consider $k_2 = 2, 2+\frac{1}{2}, .., k$; we use the method in section 4.1.9 to estimate I_2. We also consider very small $k_2 < 2$. In this case, we further decompose I_2 as follows:

$$I_2 = \left(\int_{R(k)\setminus R_s(2)} + \int_{R_s(2)\setminus R_s(k_2)} \right) K(x-y) W(y) dy \triangleq I_{21} + I_{22}.$$

For I_{21}, we apply the method in section 4.1.9. For I_{22}, we use a change of variables $y = x + sh$,

$$|I_{22}| = \left| \int_{k_2 \le |s_1| \vee |s_2| \le 2} K(-sh) W(x+sh) h^2 ds \right|.$$

Since the region is very small, $x + sh \in B_{i_1 j_1}(h_x) + [-2h, 2h]$, and $K_1(hs) = h^{-2} K_1(s)$, and we get

$$|I_{22}| \le ||W\varphi||_\infty ||\varphi^{-1}||_{L^\infty(B_{i_1 j_1}(h_x) + [-2h,2h])} \int_{k_2 \le |s_1| \vee |s_2| \le 2} |K(s)| ds.$$

The integral can be computed explicitly and has the order $\log \frac{2}{k_2}$.

It remains to estimate the most singular part I_3 for different k_2. Using a change of variables $y = x + sh$, the scaling symmetries, and the above derivations, we get

$$I_3 = \int_{[-k_2, k_2]^2} K(-s) W(x+sh) ds.$$

To use the Hölder norm of $W\psi$, we decompose it as follows:

(4.46)

$$\begin{aligned} I_3 &= \int_{[-k_2,k_2]^2} K(-s)(W\psi)(x+sh) \left(\frac{1}{\psi(x+sh)} - \frac{1}{\psi(x)} \right) + K(-s) \frac{(W\psi)(x+sh)}{\psi(x)} ds \\ &\triangleq I_{31} + I_{32}. \end{aligned}$$

For I_{32}, using the Hölder seminorm, the odd symmetry of $K(s) = c\frac{s_1 s_2}{|s|^4}$ in s_1, and $|(W\psi)(x+sh) - (W\psi)(x-sh)| \le \sqrt{2s_1 h}$, we get

$$\begin{aligned} |I_{32}| &\le \frac{h^{1/2}}{\psi(x)} [W\psi]_{C_x^{1/2}} \int_{[0,k_2]\times[-k_2,k_2]} |K(s)| \sqrt{2s_1} ds \\ &= \frac{2k_2^{1/2} h^{1/2}}{\psi(x)} [W\psi]_{C_x^{1/2}} \int_{[0,1]^2} |K(s)| \sqrt{2s_1} ds, \end{aligned}$$

where we used the scaling symmetry of K and a change of variables $s \to k_2 s$ in the last equality.

4.2.1. The commutator. For I_{31}, we apply the simple Taylor expansion to $f = \psi^{-1}$,

(4.47)
$$|f(x+sh) - f(x)| \le |f_x(x)hs_1 + f_y(x)hs_2| + h^2\left(\frac{m_{20}s_1^2}{2} + m_{11}s_1s_2 + \frac{m_{02}s_2^2}{2}\right),$$

where m_{ij} is the bound for the second derivatives of ψ^{-1},

$$m_{ij}(s) = \max_{B_{i_1j_1}(h)+I(\mathrm{sgn}(s_1))\times I(\mathrm{sgn}(s_2))} ||\partial_x^i\partial_y^j(\psi^{-1})||_{L^\infty}, \quad I_+ = [0, k_2h], \quad I_- = [-k_2h, 0].$$

Note that m_{ij} is constant in each quadrant of $[-k_2, k_2]$. We plug in the expansion (4.47) to estimate I_{31}. We only discuss a typical term $m_{20}s_1^2h^2$,

$$I_{31,20} \triangleq h^2 \int_{[-k_2,k_2]^2} |K(-s)(W\psi)(x+sh)|m_{20}(s)\frac{s_1^2}{2}ds.$$

If $k_2 \ge 2$, we can further partition $[-k_2, k_2]^2$ into $B_{2p,2q}(1/2) = [p, p+1/2] \times [q, q+1/2], -k_2 \le p, q \le k_2 - 1/2$, where we use the notation (4.17). For each grid $B_{2p,2q}(1/2)$, the sign of s and $m_{20}(s)$ are fixed, and we have

$$\begin{aligned}&\int_{B_{2p,2q}(\frac{1}{2})} |K(-s)|(W\psi)(x+sh)m_{20}(s)\frac{s_1^2}{2}ds\\ &\quad\le m_{20}||W\varphi||_\infty \int_{B_{2p,2q}\left(\frac{1}{2}\right)} \frac{|K(s)|s_1^2}{2}\left(\frac{\psi}{\varphi}\right)(x+sh)ds.\end{aligned}$$

The last integral can be estimated using the method in (4.38). Combining the estimate of integral in different regions $B_{2p,2q}(1/2)$, we obtain the estimate of $I_{31,20}$. Similarly, we can estimate the contributions of other terms in (4.47) to I_{31}.

For small $k_2 \le 2$, we do not partition the domain. We denote $D(k_2) = B_{i_1,j_1}(h_x) + [-k_2h, k_2h]^2$. For $s \in [-k_2, k_2]$, we use $x + sh \subset D(k_2) \subset D(2)$ to get

$$\begin{aligned}|f(x+sh) - f(x)| &\le ||f_x||_{L^\infty(D(k_2))}s_1h + ||f_y||_{L^\infty(D(k_2))}s_2h,\\ |W\psi(x+sh)| &\le ||W\varphi||_\infty \left\|\frac{\psi}{\varphi}\right\|_{L^\infty(D(2))}.\end{aligned} \tag{4.48}$$

Plugging the above estimate into I_{31}, we get

$$\begin{aligned}I_{31} \le &\sum_{(i,j)=(1,0),(0,1)} h||\partial_x^i\partial_y^j(\psi^{-1})||_{L^\infty(D(k_2))}||W\varphi||_\infty \left\|\frac{\psi}{\varphi}\right\|_{L^\infty(D(2))}\\ &\times \int_{[-k_2,k_2]^2} |K(s)s_1^is_2^j|ds.\end{aligned}$$

Using the scaling symmetry, we can reduce the last integral to $k_2^{i+j}\int_{[-1,1]^2}|K(s)s_1^is_2^j|ds$.

We apply the above estimates to a list of k_2 and bound different norms using $\max(||\omega\varphi||_\infty, \max_i \gamma_i[\omega\psi_1]_{C^{1/2}_{x_i}(\mathbb{R}_2^+)})$. Then by optimizing the k_2, we obtain the sharp estimate of $u_{x,A}$.

In (4.47), we do not bound $f(x+sh) - f(x)$ directly using the estimate (4.48) since s is large. Instead, we perform a higher order expansion.

Estimate of u_y, v_x. The estimates of u_y, v_x follow similar strategies and estimates. The only difference is the estimate of the most singular term similar to I_{32} (4.46) for u_y, v_x due to a different symmetry property of the kernel. We estimate it using a combination of norms $||\omega\varphi||_\infty$ and seminorms $[\omega\psi]_{C^{1/2}_{x_i}}$, and defer it to section 6.1 in the supplementary material (supplement.pdf [local/web 1.43MB]).

4.2.2. Estimate of $\mathbf{u}_A$. The estimate of $\mathbf{u}_A$ is much simpler since it is more regular. Let K and $\hat{K}$ be the kernel of u, v and its approximation term, respectively. For $f = u$ or v, we perform a decomposition similar to (4.45)

$$f_A = \left(\int_{R(k)^c} + \int_{R(k)\setminus R_s(k)} + \int_{R_s(k)} \right) K(x-y)W(y)dy - \int \hat{K}(x,y)W(y)dy \tag{4.49}$$
$$\triangleq I_1 + I_2 + I_3 + I_4.$$

The estimates of $I_1 + I_4$ follow the method for $u_{x,A}$. For I_2, we use the method in section 4.1.6. For I_3, since K has a singularity of order $|x|^{-1}$, which is locally integrable, we use a change of variable $y = x + sh$ to obtain

$$I_3 = h \int_{[-k,k]^2} K(-s)W(x+sh)ds.$$

Then we partition $[-k,k]^2$ into small grids and use the method in (4.38) to estimate the integral in each grid. Here, we get a factor h in the change of variables since $K(\lambda s) = \lambda^{-1}K(s)$.

4.2.3. Choice of parameters. Recall the choice of several parameters a, h, h_x from (4.14). We choose $3 \le k \le 10$. We choose k for the size of the singular region kh (4.45), (4.49) not so small such that the error $h^2\partial^2 K$ in Lemma 4.2, which has the order $h^2|x-y|^{-\alpha-2}$ near the singularity, is smaller than the main term K, which has the order $|x-y|^{-\alpha}, \alpha = 1, 2$. Since we will estimate $I_1 + I_4$, I_2, I_3 in the decomposition separately using the triangle inequality, we do not choose k to be too large so that we can exploit the cancellation in $I_1 + I_4$.

4.3. Hölder estimates. We want to estimate $\frac{|f(x)-f(z)|}{|x-z|^{1/2}}$ for any $x, z \in \mathbb{R}^{++}_2$ with $x_1 = z_1$ or $x_2 = z_2$ and some function f, e.g., $f = u_{x,A}$. Without loss of generality, we assume $|z| > |x|$. Then in the $C^{1/2}_x$ estimate, we have $x_1 < z_1, x_2 = z_2$; in the $C^{1/2}_y$ estimate, we have $x_1 = z_1, x_2 < z_2$. Applying the rescaling argument in section 4.1, we can restrict $\hat{x} = \frac{x}{\lambda}$ to $\hat{x} \in [0, 2x_c]^2 \setminus [0, x_c]^2$. For this reason, we assume $\lambda = 1$ for simplicity. We will only estimate the Hölder difference for comparable x, z: $|x| \asymp |z|$. If $|z| \gg |x|$, we simply apply the L^∞ estimate to $f(x), f(z)$ and use the triangle inequality.

We focus on the Hölder estimate of $u_{x,A}$, which is a representative and the most important nonlocal term to estimate in our energy estimate.

4.3.1. $C^{1/2}_x$ estimate. Recall I_i from the decomposition (4.24) and $K_1(s) = \frac{s_1 s_2}{|s|^4}$. We apply the same decomposition to $u_{x,A}(z)$. We assume that the approximation term $\hat{u}_x$ (see section 4.3 of Part I [13]) takes the following form:

$$\hat{u}_x(x) = \int \hat{K}_1(x,y)W(y)dy, \quad I_6(x) \triangleq \psi(x)\hat{u}_x(x) = \psi(x)\int \hat{K}_1(x,y)W(y)dy, \tag{4.50}$$

with a nonsingular kernel $\hat{K}_1$. We first discuss how to estimate the regular part I_1, I_3, I_4 in (4.24) and I_6, which are Lipschitz. We will apply the sharp Hölder estimate

in Lemmas 3.1–3.5 in section 3 of Part I [13] to estimate the most singular part I_2. The most technical part is to estimate I_5, which is log-Lipschitz since the kernel $K_1(x-y)(\psi(x)-\psi(y))$ has a singularity of order -1. We assemble the estimates of different parts to estimate $[u_{x,A}\psi]_{C_x^{1/2}}$ in section 4.6.

4.3.2. Estimates of the regular terms I_1, I_3, I_4, I_6. Recall I_1, I_3, I_4 from (4.24) and I_6 from (4.50). Since the integrands in I_1, I_3, I_4 are supported at least $k_2 h$ away from the singularity x, if W is in some suitable weighted L^∞ space, I_1, I_3, I_4 are piecewise smooth and their derivatives can be bounded by $||W\varphi||_{\infty(\mathbb{R}_2^{++})} = ||\omega\varphi||_\infty$. Their derivatives jump when $R(x,k), R(x,k_2)$ change, or equivalently, x moves from one grid to another. For $x \in B_{i_1,j_1}(h_x)$ (4.16), these rectangle domains are the same, and these functions are smooth. The approximation term I_6 (4.50) is locally smooth in x. To exploit the cancellation, we combine the estimates of I_1, I_4, I_6 together. We symmetrize the kernel in $I_1(x) + I_4(x) - I_6(x)$ following section 4.1.5 and use the method in section 4.1.3 to estimate the derivatives of $I_1(x) + I_4(x) - I_6(x)$. See also (4.28), (4.29) for the form of the symmetrized integrands in these integrals.

We estimate the piecewise Lipschitz norm of I_3 using the method in sections 4.1.6, 4.1.8. We choose integer k, k_2 in the decomposition (4.24). Then in each grid $[y_i, y_{i+1}] \times [y_j, y_{j+1}]$, the indicator functions in $I_1 + I_4 - I_6$, e.g., $\mathbf{1}_{R(k)^c}, \mathbf{1}_{R(k)\setminus R(k_2)}$, are constant. See Remark 4.6. We will combine the estimates of different terms in section 4.6, e.g., $I_1 + I_4 - I_6, I_3$ and part of I_5. defined later in (4.51), and obtain some Hölder continuous functions when x moves from one grid to another. We assemble the Hölder estimates in section 4.6.

4.3.3. $C_x^{1/2}$ estimate of I_2. We first estimate the second term I_2 in (4.24). Recall $R(x,k), R_{s1}(x,k), R_s(x,k)$ from (4.18), (4.19) and the location of x (4.16). We have

$$x_2 - (j-k)h \le (j+1)h - (j-k)h = (k+1)h,$$
$$(j+1+kh) - x_2 \le (j+1+kh) - jh = (k+1)h.$$

Since $x_2 = z_2$, using Lemma 3.1 from section 3 of Part I [13] with $(a, b_1, b_2) = (kh, x_2 - (j-k)h, (j+1+k)h - x_2)$ and $|b_1|, |b_2| \le (k+1)h$, we obtain

$$\frac{1}{|x-z|^{1/2}}|I_2(x,k) - I_2(z,k)| \le C_1\left(\frac{(k+1)h}{|x-z|}\right)[W\psi]_{C_x^{1/2}}$$
$$= C_1\left(\frac{(k+1)h}{|x-z|}\right)[\omega\psi]_{C_x^{1/2}}.$$

We only apply the Hölder estimate to $|x-z| \le \frac{kh}{2}$ (rescaled x, z) and the assumption $a \ge \frac{1}{2}|x_1 - z_1|$ in Lemma 3.1 in Part I [13] is satisfied. For $I_2(x,k)$ associated with other terms u, v, u_y, v_x, we can estimate it using similar ideas and Lemmas 3.1–3.5 in Part I [13]. The $C_y^{1/2}$ estimate of $I_2(x,k)$ is completely similar. See section 4.3.8 for more details.

4.3.4. $C_x^{1/2}$ estimate of I_5. For I_5 (4.24), $K_1(x-y)(\psi(x)-\psi(y))$ is singular of order -1 near $y = x$. Given $W \in L^\infty(\varphi)$, I_5 is log-Lipschitz. There are several approaches to estimate its Hölder norm (see, e.g., section 4.1.11). We use part of the $C_x^{1/2}$ seminorm of ω to get a better estimate. We choose $k_3 = k_2 - \frac{i}{2} \ge 2, i = 0, 1, 2, .., 2k_2 - 4$, and further decompose I_5 as follows:

(4.51)
$$I_5(x,k_2)=\left(\int_{R(k_2)\backslash R(k_3)}+\int_{R(k_3)\backslash R_{s,1}(k_3)}+\int_{R_{s,1}(k_3)}\right)K_1(x-y)(\psi(x)-\psi(y))W(y)dy$$
$$\triangleq I_{5,0}(x,k_2,k_3)+I_{5,1}(x,k_3)+I_{5,2}(x,k_3).$$

The domain in $I_{5,0}$ depends on x. For x in a grid cell, it does not change with x. We estimate $\partial_{x_1}I_{5,0}$ using Taylor expansion in section 4.1.10 and following the method in section 4.1.7. We estimate the x-derivative of $I_{5,1}$ using the method in sections 4.1.6, 4.1.8. We have

(4.52)
$$\partial_{x_1}I_{5,1}=\int_{R(k_3)\backslash R_{s,1}(k_3)}\partial_{x_1}(K_1(x-y)(\psi(x)-\psi(y)))W(y)dy$$
$$-\int_{(j-k_3)h}^{(j+1+k_3)h}K_1(x-y)(\psi(x)-\psi(y))W(y)\Big|_{y_1=x_1-k_3h}^{x_1+k_3h}dy_2.$$

We estimate the first part following section 4.1.10 and the second part following section 4.1.8.

For $I_{5,2}$, we will estimate it using a method similar to that of I_2. See the left figure in Figure 3 for the domains of the integrals in $I_{5,2}(x), I_{5,2}(z)$. The integrand satisfies

$$K_1(x-y)(\psi(x)-\psi(y))W(y)=\psi(x)K_1(x-y)(\psi^{-1}(y)-\psi^{-1}(x)(W\psi)(y)$$
$$\approx\psi(x)\partial_i(\psi^{-1}(x))\cdot K_1(x-y)(y_i-x_i)(W\psi)(y).$$

Thus, $I_{5,2}(x)$ can be seen as a weighted version of I_2 (4.24) with a weight $\psi(x)\partial_i(\psi^{-1}(x))$, a more regular kernel $K_1(x-y)(y_i-x_i)$, and a smaller domain $R_{s,1}(k_3)$. Since the kernel is more regular and the domain is smaller, our estimate for $I_{5,2}$ is much smaller than that of I_2.

Now, we justify this approach. Using a change of variables $y=x+s, s\in R_{s,1}(k_3)-x$ and the above identity yields

$$I_{5,2}(x,k_3)=\psi(x)\int_{R_{s,1}(k_3)-x}K_1(-s)(\psi^{-1}(x+s)-\psi^{-1}(x))(W\psi)(x+s)ds.$$

Using Newton's formula $f(1)=f(0)+f'(0)+\int_0^1(1-t)f''(t)dt$ for $f(t)=\psi^{-1}(x+ts)$, we get

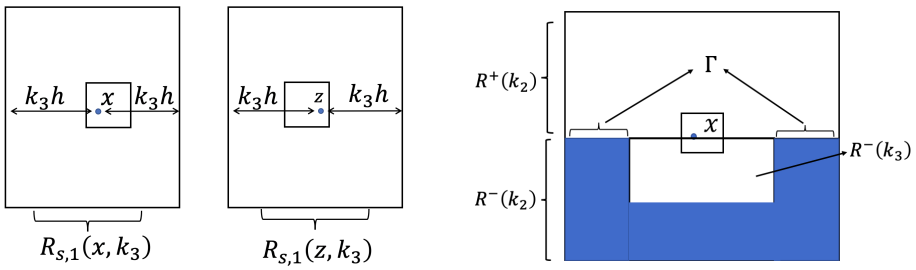


FIG. 3. *Left:* $R_{s,1}(x,k_3)$ *and* $R_{s,1}(z,k_3)$ *with* $x_2=z_2$. *The small square is a mesh grid containing* x *or* z. x,z *can have different locations relative to the grids. Right: The large rectangle is* $R(k_2)$, *the upper part is* $R^+(k_2)$, *and the lower part is* $R^-(k_2)$. *The blue region is* $R^-(k_2)\backslash R^-(k_3)$. Γ *is part of its boundary.*

$$\begin{aligned}
&\psi^{-1}(x+s)-\psi^{-1}(x)\\
&\quad= s\cdot\nabla\psi^{-1}(x)+\int_0^1(1-t)\Big(s\cdot(\nabla^2\psi^{-1})(x+ts)\cdot s\Big)dt\\
&\quad= \sum_{i=1,2} s_i\partial_i(\psi^{-1})(x)+\sum_{0\le i\le 2}\binom{2}{i}s_1^i s_2^{2-i}\int_0^1(1-t)\partial_1^i\partial_2^{2-i}(\psi^{-1})(x+ts)dt.
\end{aligned}$$

Denote

$$\begin{aligned}
&Q_{ij}(x)=\psi(x)\int_0^1(1-t)\partial_1^i\partial_2^j(\psi^{-1})(x+ts)dt,\ i+j=2,\quad D(x)=R_{s,1}(x,k_3)-x,\\
&Q_{ij}(x)=\psi(x)\cdot\partial_1^i\partial_2^j(\psi^{-1})(x)=-\frac{\partial_1^i\partial_2^j\psi(x)}{\psi(x)},\ i+j=1,\\
&P_{ij}(x)=\int_{D(x)}K_1(-s)s_1^i s_2^j(W\psi)(x+s)ds.
\end{aligned}$$

Using the above expansion and notation, we get

$$I_{5,2}(x,k_3)=\sum_{i+j=1}P_{ij}Q_{ij}+\sum_{i+j=2}\binom{2}{i}P_{ij}Q_{ij}.$$

Next, we use the above decomposition to estimate $I_{5,2}(x,k_3)-I_{5,2}(z,k_3)$. The leading order terms are $P_{ij}Q_{ij}$ with $i+j=1$. By the definition of $R_{s,1}$ (4.19), we observe that if $x_2=z_2$, we have

$$D(x)=R_{s,1}(x,k_3)-x=R_{s,1}(z,k_3)-z=D(z).$$

Suppose that $x_1<z_1$. We perform a decomposition

$$\begin{aligned}
&|P_{ij}(x)Q_{ij}(x)-P_{ij}(z)Q_{ij}(z)|\le J_1+J_2,\\
&J_1\triangleq|Q_{ij}(z)(P_{ij}(x)-P_{ij}(z))|,\quad J_2\triangleq|P_{ij}(x)(Q_{ij}(x)-Q_{ij}(z))|.
\end{aligned}\tag{4.53}$$

Using $D(x)=D(z)$, we bound J_1 as follows:

$$\begin{aligned}
|J_1|&\le|Q_{ij}(z)|\Big|\int_{D(x)}K_1(-s)s_1^i s_2^j((W\psi)(x+s)-(W\psi)(z+s))ds\Big|\\
&\le|Q_{ij}(z)|\cdot|x-z|^{1/2}||\omega\psi||_{C_x^{1/2}}\int_{s\in D(x)}|K_1(s)s_1^i s_2^j|ds.
\end{aligned}$$

The term Q_{ij} only depends on the weight and is smoother than P_{ij}. We can estimate $Q_{ij}(x)-Q_{ij}(z)$ by bounding $\partial_1 Q_{ij}$ since Q_{ij} is locally smooth. For P_{ij} in J_2, we use the method in (4.38) to bound it by $C||\omega\varphi||_\infty$ with some constant C. Then we obtain the estimate

$$|J_2|\le C_2|x-z|\cdot||\omega\varphi||_{L^\infty}$$

for some constant C_2. Note that the second order term $P_{ij}Q_{ij}, i+j=2$ is much smaller than the leading order terms. For $|x-z|$ not too small, we can estimate its contribution trivially

$$\frac{1}{|x-z|^{1/2}}|P_{ij}(x)Q_{ij}(x)-P_{ij}(z)Q_{ij}(z)|\le\frac{1}{|x-z|^{1/2}}(|P_{ij}(x)Q_{ij}(x)|+|P_{ij}(z)Q_{ij}(z)|).\tag{4.54}$$

We optimize the above two estimates.

In summary, to obtain the above estimates, we estimate piecewise bounds for $|Q_{ij}(x)|$, $P_{ij}(x), |\partial_k Q_{ij}(x)|$ and the integrals $\int_{D(x)} |K_1(s)s_1^i s_2^j|ds,\ i+j=1,2$.

The above estimate of $I_5(x,k_2)$ can be generalized to the $C_x^{1/2}$ estimate of u, v, v_x, u_y. Yet, it does not apply to the $C_y^{1/2}$ estimate of $\mathbf{u}, \nabla\mathbf{u}$ since it requires the estimate of $(W\psi)(x+s)-(W\psi)(z+s)$ for s in some rectangle $R=D(x)=D(z)$. However, since W is discontinuous across the boundary $y=0$, $W\psi \notin C_y^{1/2}(R)$ if $x+s, z+s$ are not in the same half plane. If $x_1<x_2$, then the rectangles $R(x,k_2), R(z,k_2)$ will not intersect the boundary and the previous estimate holds true. If $x_1>x_2$, we consider two modifications for different kernels in the following subsections.

4.3.5. Ideas of the $C_y^{1/2}$ estimates of I_5. The main idea in the following $C_y^{1/2}$ estimates is to use a combination of the estimates for the log-Lipschitz function in section 4.1.11 and the estimate in section 4.3.4. The latter provides better estimates, and we try to use this method *as much as possible.* Following the ideas in section 4.1.11, we decompose $I_5(x)$ into the singular part and nonsingular part with different size k_3 of the singular region

$$I_5(x) = I_{5,S}(x,k_3) + I_{5,NS}(x,k_3).$$

Although we cannot apply the second method to the whole $I_5(x)$, we can apply it to the integrals in the upper part of the regions, e.g., $R^+(k_2), R^+(k_3)$ (4.20), since these integrals only involve $W\psi$ in $\mathbb{R}^+_2$ and we have $W\psi \in C^{1/2}$. Thus, we will further decompose some of the regions into the upper part and the lower part and then apply the first method to the lower part, and the second method to the upper part.

4.3.6. $C_y^{1/2}$ estimate of the velocity with a kernel of the first type. The kernels

$$K = \frac{y_1 y_2}{|y|^4}, \quad \frac{y_2}{|y|^2} \tag{4.55}$$

associated with $u_x = -\partial_{xy}(-\Delta)^{-1}\omega, u = -\partial_y(-\Delta)^{-1}\omega$ vanish when $y_2=0$. We call them the first type kernel. Let K be a kernel of the first type. We use the following decomposition:

$$\begin{aligned} I_5(x,k_2) &= \left(\int_{R^+(k_2)} + \int_{R^-(k_2)}\right) K(x-y)(\psi(x)-\psi(y))W(y)dy \\ &\triangleq I_5^+(x,k_2) + I_5^-(x,k_2). \end{aligned} \tag{4.56}$$

See the right figure in Figure 3 for $R^{\pm}(k_2)$. Since $R^+(x,k_2), R^+(z,k_2) \subset \mathbb{R}^+_2$, we can decompose

$$I_5^+ = I_{5,1}^+ + I_{5,2}^+$$

into the integral in the regions $I_{5,1}^+ : R^+(k_2)\backslash R^+_{s,2}(k_2)$ and $I_{5,2}^+ : R^+_{s,2}(k_2)$ and apply the same argument as that for $I_{5,1}(x,k_3), I_{5,2}(x,k_3)$ in section 4.3.4 to obtain the desired estimates by restricting all the derivations in $R^+(x,k_2), R^+(z,k_2)$. Note that here, we do not further choose smaller window $R^+(x,k_3)$ to decompose $I_5^+(x,k_2)$, i.e., $k_3=k_2$ and $I_{5,0}=0$ in (4.51). For $I_{5,1}^+$, similar to (4.52), we get a boundary term from $\partial_2(R^+(k_2)\backslash R^+_{s2}(k_2)) = [(i-k_2)h, (i+1+k_2h)] \times \{x_2+k_2h\}$. See (4.19), (4.18) for $R^+(k), R^+_{s2}(k)$.

For the lower part $I_5^-(x,k_2)$, it is log-Lipschitz if $W\in L^\infty(\varphi)$. We cannot bound its derivative using $||W\varphi||_\infty$. We face the difficulty discussed at the beginning of section 4.

Alternatively, we follow the ideas in section 4.1.11. We decompose it into the smooth part and rough part. We introduce $0<k_3<k_2$ and consider the following decomposition:

(4.57)
$$\begin{aligned} &I_5^-(x,k_2)\\ &=\left(\int_{R^-(k_2)\backslash R^-(k_3)}+\int_{R^-(k_3)\backslash R^-_{s,2}(k_3)}+\int_{R^-_{s,2}(k_3)}\right)K(x-y)(\psi(x)-\psi(y))W(y)dy\\ &\triangleq I^-_{5,0}(x,k_2)+I^-_{5,1}(x,k_3)+I^-_{5,2}(x,k_3). \end{aligned}$$

See the right figure in Figure 3 for an illustration of different domains. Recall that $k_2\in Z_+$. We choose $k_3=k_2-\frac{i}{2}\geq 2, i=0,1,2\ldots,2k_2-4$. Since the integrand in $I^-_{5,0}$ supports at least k_3h away from the singularity, $I^-_{5,0}(x,k_2)$ is piecewisely smooth. We can estimate $\partial_{x_2}I^-_{5,0}(x,k)$ following sections 4.1.7, 4.1.10. The domain $R^-(k_2)\backslash R^-(k_3)$ is not piecewise constant since the upper part of its boundary, i.e.,

$$\Gamma=\{(y_1,x_2): y_1\in[(i-k_2)h,(i+1+k_2)h]\backslash[(i-k_3)h,(i+1+k_3)h]\},$$

depends on x_2. See Figure 3 for an illustration of Γ. Taking x_2-derivative on $I^-_{5,1}$, we get

(4.58)
$$\begin{aligned} |\partial_{x_2}I^-_{5,0}(x,k_2)|\leq&\left|\int_{R^-(k_2)\backslash R^-(k_3)}\partial_{x_2}J(x,y)W(y)dy\right|+\left|\int_{y\in\Gamma}J(x,y)W(y)dy_1\right|,\\ &J(x,y)=K(x-y)(\psi(x)-\psi(y)). \end{aligned}$$

Since $y\in\Gamma\subset\{y: y_2=x_2\}$ and $K(y_1,0)\equiv 0$, the second term vanishes. The first term can be estimated using a change of variables $y=x+s$ and the method in sections 4.1.10, 4.1.7, since its support is at least k_3h away from the singularity.

For $I^-_{5,1}$, it is also piecewise Lipschitz, and we estimate the x_2-derivative similarly to $I_{5,1}$ in (4.52)

(4.59)
$$\begin{aligned} |\partial_{x_2}I^-_{5,1}|\leq&\left|\int_{R^-(k_3)\backslash R^-_{s,2}(k_3)}\partial_{x_2}J(x,y)W(y)dy\right|\\ &+\left|\int_{(i-k_3)h}^{(i+1+k_3)h}J(x,y)W(y)\Big|_{y_2=x_2-k_3h}dy_1\right|. \end{aligned}$$

Different from $I_{5,1}$ in (4.52), the boundary term in the above estimate only involves the lower part $y_2=x_2-k_3h$ since the domain in $I^-_{5,1}$ is $R^-(k_3)\backslash R^-_{s,2}(k_3)$.

For $I^-_{5,2}$, the kernel satisfies $K(x-y)(\psi(x)-\psi(y))\sim|x-y|^{-1}$ for small $|x-y|$ and is locally integrable. We estimate its piecewise L^∞ bound using the method in section 4.2.1 for the commutator.

The above decomposition can be applied to estimate

$$\begin{aligned} \frac{|I_5^-(x,k_2)-I_5^-(z,k_2)|}{|x-z|^{1/2}}\leq&\min_{k_3=k_2-\frac{i}{2}}\frac{|(I^-_{5,0}+I^-_{5,1})(x,k_3)-(I^-_{5,0}+I^-_{5,1})(z,k_3)|}{|x-z|^{1/2}}\\ &+\frac{|I^-_{5,2}(x,k_3)|+|I^-_{5,2}(z,k_3)|}{|x-z|^{1/2}} \end{aligned}$$

for $|x-z|$ not too small, e.g., $|x-z| \geq d_s = \frac{h}{10}$. When $|x-z|$ is sufficiently small, the second term in the above estimate can be very large.

According to the analysis in section 4.1.11, for $|x-z|$ very small, we need to choose $k_3 h \sim |x-z|$ to get the sharp estimate. Thus, we consider one more decomposition for $a \leq 1$,

$$\begin{aligned} I_5^-(x,k_2) = &\int_{R^-(k_2)\setminus R_s^-(a)} K(x-y)(\psi(x)-\psi(y))W(y)dy \\ &+ \int_{R_s^-(a)} K(x-y)(\psi(x)-\psi(y))W(y)dy \triangleq I_{5,3}^-(x,a) + I_{5,4}^-(x,a). \end{aligned} \tag{4.60}$$

The above decomposition is slightly different from (4.57). We choose $R_s^-(a)$ rather than $R^-(a)$, since we need to choose the singular region with size going to 0 as $|x-z| \to 0$. Yet, $R^-(a)$ (4.18) does not satisfy this requirement for $a \to 0$. We can estimate the derivative of $I_{5,3}^-(x,a)$ following sections 4.1.6–4.1.8 and the L^∞ norm of $I_{5,4}^-(x,a)$ following section 4.2.1. Again, in the computation of $\partial_{x_2} I_{5,3}^-(x,a)$, the boundary term vanishes due to $K(y_1,0) \equiv 0$. In summary, we can obtain the estimate

$$|\partial_{x_2} I_{5,3}^-(x,a)| \leq A(x) + B(x)\log(1/a), \quad |I_{5,4}^-(x,a)| \leq C(x)ah, \tag{4.61}$$

for any $a \leq 1$, where $A(x), B(x)$ can be estimated following the method in Appendix B.5.1, and the estimate of $C(x)$ follows the method in section 4.2.1. Using the above estimates and the ideas in section 4.1.11, we can estimate $d_y(I_5^-(\cdot,k_2),x,z)$ for small $|x-z|$ by optimizing a, where d_y is defined below:

$$d_y(f,x,z) = |f(x)-f(z)||x-z|^{-1/2}. \tag{4.62}$$

We will assemble these estimates in section 4.6.

4.3.7. $C_y^{1/2}$ estimate of the velocity with a kernel of the second type. For the kernels $K_2 = \frac{y_1^2-y_2^2}{|y|^4}$ and $\frac{y_1}{|y|^2}$, they do not vanish on $y_2 = 0$ in general. We call them the second type kernel.

If we use the strategies in the previous subsection, the boundary term in the computation of $\partial_{x_2} I_{5,0}^-(x,k_3), \partial_{x_2} I_{5,1}^-(x,k_3)$ or $\partial_{x_2} I_{5,3}^-(x,k_3)$ does not vanish on Γ and can be large. To avoid picking up a boundary term on Γ and to apply the ideas in section 4.3.5, we consider another estimate on $I_5(x,k_2)$. For $k_3 = k_2 - \frac{i}{2}, i = 0,1,\ldots,2k_2-4$, we perform the following decomposition:

$$\begin{aligned} I_5(x,k_2) = &\left(\int_{R(k_2)\setminus R(k_3)} + \int_{R^-(k_3)\setminus R_{s,2}^-(k_3)} + \int_{R^+(k_3)} + \int_{R_{s,2}^-(k_3)} \right) \\ &\times K(x-y)(\psi(x)-\psi(y)W(y)dy \\ \triangleq\ & I_{5,0} + I_{5,1} + I_{5,2} + I_{5,3}. \end{aligned} \tag{4.63}$$

Following the ideas in section 4.1.11, we estimate the derivative of the regular part and then the L^∞ norm of the singular part. Indeed, we can estimate the y-derivative of $I_{5,0}$ following sections 4.1.10, 4.1.7, $I_{5,1}$ following the estimates of $I_{5,1}, I_{5,1}^-$ in (4.52), (4.59), and the L^∞ norm of $I_{5,3}$ following section 4.2.1. The estimate of $I_{5,1}$ is similar to that of I_4 in section 4.3.2. For $I_{5,2}$, since $R^+(k_3)$ is in $\mathbb{R}_2^+$, we decompose

$$I_{5,2} = I_{5,2,1} + I_{5,2,2}$$

into the integral in the regions $I_{5,2,1} : R^+(k_3)\backslash R^+_{s,2}(k_3)$ and $I_{5,2,2} : R^+_{s,2}(k_3)$, and then estimate them following the method in the estimate of $I_{5,1}, I_{5,2}$ in section 4.3.4.

After we estimate these quantities, we can estimate $d_y(I_5, x, z)$ (4.62) for $|x-z|$ not too small by optimizing k_3. To estimate $d_y(I_5, x, z)$ (4.62) for sufficiently small $|x-z|$, following (4.60), we use the decomposition

$$\begin{aligned} I_5(x,k_2) = &\int_{R(k_2)\backslash R_s(a)} K(x-y)(\psi(x)-\psi(y)W(y)dy \\ &+ \int_{R^+_s(a)} K(x-y)(\psi(x)-\psi(y))W(y)dy \\ &+ \int_{R^-_s(a)} K(x-y)(\psi(x)-\psi(y))W(y)dy \triangleq I_{5,4}+I_{5,5}+I_{5,6}. \end{aligned} \tag{4.64}$$

Then we estimate the derivative of $I_{5,4}$ and the L^∞ norm of $I_{5,6}$ as follows:

$$|\partial_{x_2} I_{5,4}| \le A(x) + B(x)\log(1/a), \quad |I_{5,6}| \le C(x)ah, \tag{4.65}$$

where the estimates of A, B are given in Appendix B.5.1, and the estimate of C follows the method in section 4.2.1. The Hölder estimate of $I_{5,5}$ follows the method in the estimate of $I_{5,2}$ in section 4.3.4. With these estimates, we can further bound $d_y(I_5, x, z)$,

$$d_x(f,x,z) \triangleq \frac{|f(x)-f(z)|}{|x_1-z_1|^{1/2}}, \quad d_y(f,x,z) \triangleq \frac{|f(x)-f(z)|}{|x_2-z_2|^{1/2}}$$

for sufficiently small $|x-z|$ by optimizing a. See section 4.6.

Remark 4.7. We do not apply the above computation with smaller window $[-ah, ah]^2$ in the $C^{1/2}_x$ estimate, since it leads to a worse estimate. See also the discussions in section 4.3.5.

4.3.8. Hölder estimate of u, v, u_y, v_x. The ideas of the Hölder estimate for other terms are similar. For a kernel K associated with $\mathbf{u}, \nabla\mathbf{u}$, we perform another decomposition similar to (4.24):

$$\begin{aligned} \psi(x)\int K(x-y)W(y)dy = &\int \Big(\psi(x)\mathbf{1}_{R(k)^c} + \mathbf{1}_{R_s(k)}\psi(y) + \mathbf{1}_{R(k)\backslash R_s(k)}\psi(y) \\ &+ \mathbf{1}_{R(k)\backslash R(k_2)}(\psi(x)-\psi(y)) + \mathbf{1}_{R(k_2)}(\psi(x)-\psi(y))\Big)K(x-y)W(y)dy \\ \triangleq\ & I_1(x,k) + I_2(x,k) + I_3(x,k) + I_4(x,k,k_2) + I_5(x,k_2). \end{aligned} \tag{4.66}$$

Here, we use $R_s(x,k)$ (4.19), which is symmetric with respect to both x_1 and x_2, rather than $R_{s,1}(x,k)$, since the singular region in the sharp Hölder estimate of $[u_y]^{1/2}_{C^{1/2}_{x_i}}, [v_x]_{C^{1/2}_{x_i}}, [u_x]_{C^{1/2}_y}$ in Lemmas 3.3–3.5 in Part I [13] needs to be symmetric in both x_1, x_2. Denote by $I_{f6}(x,k_2)$ the approximation term for $f = u_x, u_y, v_x, u, v$. It takes a form similar to (4.50).

We consider two cases of $\hat{x} \in [0,2x_c]^2\backslash[0,x_c]^2$ (4.4). In the first case, we consider $\hat{x} \in [x_c, 2x_c]\times[0,2x_c] \triangleq D_{X1}$, where we have $\hat{x}_1 \ge c\hat{x}_2$ for some constant $c > 0$. In the second case, we consider $\hat{x} \in [0,x_c]\times[x_c,2x_c] \triangleq D_{X2}$, where we have $\hat{x}_1 \le c\hat{x}_2$. We distinguish these two cases since in the second case, the singular region does not touch the boundary, and we can apply the method in section 4.3.4.

$C_x^{1/2}$ estimate of u_y, v_x. In the $C_x^{1/2}$ estimate of u_y, v_x, we follow section 4.3.2 to estimate the regular part $I_1 + I_4 - I_6$ and I_3. We follow section 4.3.3 and use Lemma 3.4 in section 3 of Part I [13] to estimate I_2. For I_5, we follow section 4.3.4.

$C_y^{1/2}$ estimate of u_x. We perform the decomposition (4.66) rather than (4.24). The estimates of $I_1 + I_4 - I_6$, I_3 follow section 4.3.2. For I_2, we use Lemma 3.3 in section 3 of Part I [13]. We follow section 4.3.6 to estimate I_5 if $\hat{x} \in D_{X1}$, and section 4.3.4 if $\hat{x} \in D_{X2}$.

We remark that we use the decomposition (4.66) rather than (4.24) since in Lemma 3.3 in section 3 of Part I [13], we need to assume that the singular region around x is symmetric in both x_1 and x_2. The same reasoning applies to the $C_{x_i}^{1/2}$ estimate of u_y, v_x.

$C_x^{1/2}$ and $C_y^{1/2}$ estimate of u, v. The Hölder estimates of u, v are substantially easier since u, v are more regular. We perform $C_x^{1/2}, C_y^{1/2}$ of $\rho \mathbf{u}_A$ for another weight $\rho = \psi_u$ (A.1). Below, we only use the weighted L^∞ norm $||\omega\varphi||_\infty$. We decompose the integral as follows:

$$\begin{aligned} \rho(x) \int K(x-y) W(y) dy &= \int (\mathbf{1}_{R(k)^c} \rho(x) + \mathbf{1}_{R(k)} \rho(x)) K(x-y) W(y) dy \\ &\triangleq I_1(x,k) + I_2(x,k). \end{aligned} \tag{4.67}$$

We choose k smaller than that in (4.24) for $\nabla \mathbf{u}$ since the kernel for $\mathbf{u}$ is more regular. We follow section 4.3.2 to estimate $I_1 - I_6$. For I_2, we follow the ideas in sections 4.1.11, 4.3.6, 4.3.7 to estimate the log-Lipschitz function. We choose a list of k_2 and associated region $S(k_2)$ and decompose I_2 as follows:

$$\begin{aligned} I_2(x,k) &\triangleq \left(\int_{R(k) \setminus R(k_2)} + \int_{R(k_2) \setminus S(k_2)} + \int_{S(k_2)} \right) \rho(x) K(x-y) W(y) dy \\ &\triangleq I_{20}(x,k_2) + I_{21}(x,k_2) + I_{22}(x,k_2). \end{aligned}$$

For large $k_2 = k, k - 1/2, \ldots, 2$, we choose $S(k_2) = R_{s,i}(k_2)$ in the $C_{x_i}^{1/2}$ estimate, $i = 1, 2$. For $k_2 < 2$, we choose $S(k_2) = R_s(k_2)$. For $I_{20}(x,k_2), I_{21}(x,k_2)$, we estimate its derivatives following the estimate of $I_{50}, I_{5,1}$ (4.51), (4.52), respectively, or section 4.1.7 when $k_2 \geq 2$, and the estimate of I_{54} when $k_2 < 2$ in section 4.3.7. For $I_{22}(x,k_2)$, we estimate its L^∞ norm following the estimate of I_{53} when $k_2 \geq 2$, and the estimate of I_{56} when $k_2 < 2$ in section 4.3.7. The estimate is simpler since the above kernel is much simpler than $K(x-y)(\psi(x) - \psi(y))$ in section 4.3.7.

4.3.9. Special case: $C_y^{1/2}$ estimate of u_y, v_x. In this case, we apply Lemma 3.5 from section 3 of Part I [13] to estimate the most singular part. Since in Lemma 3.5 from section 3 of Part I, we do not localize the integral, we perform the following decomposition:

$$\begin{aligned} \psi(x) \int K(x-y) W(y) dy &= \int \Big(\psi(y) + \mathbf{1}_{R(k_2)^c} (\psi(x) - \psi(y)) + \mathbf{1}_{R(k_2)} (\psi(x) - \psi(y)) \Big) \\ &\quad \times K(x-y) W(y) dy \\ &\triangleq I_1(x,k) + I_2(x,k) + I_3(x,k). \end{aligned} \tag{4.68}$$

For I_1, we apply Lemma 3.5 from Part I [13]. We follow section 4.3.7 to estimate I_3 if $\hat{x} \in D_{X1}$, and section 4.3.4 if $\hat{x} \in D_{X2}$. We follow section 4.3.2 to estimate $I_2 - I_6$,

where I_6 is the approximation term for u_y, v_x similar to (4.50). The symmetrized integrand is discussed in the paragraph "$C^{1/2}$ estimate of u_y, v_x" in section 4.1.5. There are additional difficulties since the weight $\psi(y)$ and the symmetrized integrand $I(x,y) = K^C(x,y)(\psi(x) - \psi(y))$ for some kernel K^C (see similar derivations in (4.28), (4.29)) are singular near 0. Note that we do not have the K^{NC} term. See the paragraph $C^{1/2}$ *estimate of* u_y, v_x before section 4.1.6.

The integral of $I(x,y)$ near 0 or in the far-field require some additional estimates, which we discuss below. Since y is away from the singularity x in these cases, the symmetrized integral is given by $I = K^{sym}(x,y)(\psi(x) - \psi(y))$. See (4.29) and section 4.1.5 for related discussions.

Estimate the integral near 0. To estimate the $D_1 = \partial_{x_2}$ derivative, we use

$$\begin{aligned} |D_1 I| &= |D_1 K^{sym}(\psi(x) - \psi(y)) + K^{sym} \cdot D_1 \psi(x)| \\ &\le |D_1 K^{sym} \cdot \psi(x) + K^{sym} \cdot D_1 \psi(x)| + |D_1 K^{sym} \cdot \psi(y)|. \end{aligned}$$

For y close to 0, since ψ is singular, $\psi(y)$ is much larger than $\psi(x)$, and $K^{sym}(x,y)$ is not singular. The main term in $D_1 I$ is given by $D_1 K^{sym} \psi(y)$. It follows that

$$\begin{aligned} \int_Q |D_1 I \cdot W(y)| dy \le ||W\varphi||_\infty \Big(||\varphi^{-1}||_{L^\infty(Q)} \int_Q |D_1 K^{sym} \psi(x) + K^{sym} \cdot D_1 \psi(x)| dy \\ + \left| \frac{\psi}{\varphi} \right|_{L^\infty(Q)} \int_Q |D_1 K^{sym}| dy \Big), \end{aligned}$$

where Q is some grid near the origin. The integrands in both integrals do not involve the singular weight, and we can estimate them for each grid point x using the previous methods.

To estimate the X-discretization error, we need to estimate the integral of $\partial_{xi}^2 \partial_{x_2} J$. Since $\psi(y)$ is independent of x, we get

$$\begin{aligned} I = K^{sym}(x,y) \left(\frac{\psi(x)}{\psi(y)} - 1 \right) \psi(y), \int_Q |\partial_{x_i}^2 \partial_{x_2} I \cdot W(y)| dy \\ \le ||W\varphi||_\infty \left| \frac{\psi}{\varphi} \right|_{L^\infty(Q)} \int_Q \left| \partial_{x_i}^2 \partial_{x_2} K^{sym}(x,y) \left(\frac{\psi(x)}{\psi(y)} - 1 \right) \right| dy. \end{aligned}$$

The last integrand is not singular in y near $y = 0$, and we estimate it using the previous method, e.g., section 4.1.3.

For u_y, v_x, we have a rank-one approximation $K_{app}(x,y)$ from $C_{u_y} \chi_0 K_{00}$ (4.5) (see section 4.3.2 from Part I [13]). The full integrand with approximation term and weight is given by

$$\begin{aligned} I_{app} &= K^{sym}(x,y)(\psi(x) - \psi(y)) - K_{app}(x,y)\psi(x) \\ &= (K^{sym}(x,y) - K_{app}(x,y))\psi(x) - K^{sym}(x,y)\psi(y) = I_{app,1} + I_{app,2}. \end{aligned}$$

For y away from the singularity x and 0, $I_{app,1}$ has the same form as the previous case, e.g., the $C_x^{1/2}$ estimate. We improve the error estimate $\partial_i^2 \partial_{x_2} I_{app}$ using the cancellation between the full symmetrized kernel $K(x,y)$ and K_{app} from Lemma B.2 and the estimate in (B.15) in Appendix B.1.1 and the property that $\psi(y)$ is much smaller than $\psi(x)$ for $|y|$ much larger than $|x|$.

Estimate in the far-field. For the tail part in this case, we have an improvement for small $|x|$ where $\chi_0(x) = 1$ due to the approximation term near 0

$$\hat{f} = C_{f0}(x,y)u_x(0) + C_f(x,y)\mathcal{K}_{00} = C_f(x,y)\mathcal{K}_{00},$$

where $f = u_y, v_x$ and $\mathcal{K}_{00}$ is defined in (4.5), and we have used $C_{f0}(x,y) = 0$. Its associated integrand is given by

$$K_{app} \triangleq \pi^{-1} C_f(x,y) K_{00}(y),$$

where K_{00} is defined in (4.5). To estimate it, we use the following decomposition:

$$D_1(I - \psi(x)K_{app}) = D_1((K^{sym} - K_{app}) \cdot \psi(x)) - D_1 K^{sym} \cdot \psi(y) \triangleq P_1 + P_2.$$

We estimate P_1 using the method in section 4.4. Due to the approximation, $K^{sym} - K_{app}$ has a much faster decay for large y beyond $[0,D]^2$. See (B.15) and Appendix B.1.1. For P_2, we have

$$\int_{\Omega^c} |P_2||W(y)|dy \le ||W\varphi||_\infty \int_{\Omega^c} |D_1 K^{sym}| \frac{\psi}{\varphi}(y)dy,$$

where $\Omega = [0,D]^2$ with large D. The last integral is computed using the method in section 4.4.

4.4. Estimate the integrals near 0 and in the far-field. We use a combination of uniform mesh and adaptive mesh to compute the integral in a finite domain $[0,D]^2$, e.g., $D = 1000$. See section 4.1.3. Since the kernel decays and the singularity is in the near-field, the integral beyond this domain is small, and we estimate it directly. In addition, for y near 0, we estimate the integrals (the last two integrals in (4.8)) from the approximations $u_x(0), K_{00}$ (4.7), which is singular of order $|y|^{-2}$ or $|y|^{-4}$. For simplicity, we consider $\lambda = 1$. The estimates can be generalized to another scaling parameter λ. To estimate $\int_D k(y)\omega(y)dy$ for D near 0 or D in the far-field, following (4.10), we only need to estimate $\int_D |k(y)|\varphi^{-1}(y)dy$. Since $|y|$ is either very small or very large, we can use the asymptotics of φ in these estimates.

4.4.1. Near-field estimate. First, we estimate $\int_{[0,R_1]^2} |k(y)|\varphi^{-1}(y)dy$ for $k(y) = \frac{y_1 y_2}{|y|^4}, \frac{y_1 y_2 (y_1^2 - y_2^2)}{|y|^8}$ related to $u_x(0), K_{00}$ (4.7). We partition $[0,R_1]$ into

$$0 = z_0 < z_1 < \cdots < z_n = R_1$$

with z_1 much smaller than R_1. Denote $Q_{ij} = [z_{i-1}, z_i] \times [z_{j-1}, z_j]$. Clearly, we have

$$\int_{[0,R_1]^2} |k(y)|\varphi^{-1}(y)dy \le \sum_{1 \le i,j \le n} I_{ij}, \quad I_{ij} \triangleq \int_{Q_{ij}} |k(y)|\varphi^{-1}(y)dy.$$

For $I_{ij}, (i,j) \ne (1,1)$, we apply a trivial bound

$$I_{ij} \le ||\varphi^{-1}||_{L^\infty(Q_{ij})} \int_{Q_{ij}} |k(y)|dy \le |Q_{ij}| \cdot ||k||_{L^\infty(Q_{ij})} ||\varphi^{-1}||_{L^\infty(Q_{ij})}. \tag{4.69}$$

For $k(y) = \frac{y_1 y_2}{|y|^4}, \frac{y_1 y_2 (y_1^2 - y_2^2)}{|y|^8}$, the estimate of $||k||_{L^\infty(Q_{ij})}$ is established in Appendix B. It remains to estimate the first term I_{11}. Denote $r = y_1$. Suppose that

$$\varphi(x) \ge q|x|^a (\cos\beta)^b, \quad b \le 0.$$

See (A.2). If $k(y)=\frac{y_1y_2}{|y|^4}$ and $a<0$, we get

$$\begin{aligned}I_{11}&\le q^{-1}\int_0^{\sqrt{2}r}\int_0^{\pi/2}\frac{\sin\beta\cos\beta}{r^2}r^{-a}(\cos\beta)^{-b}rdrd\beta\\&=q^{-1}\int_0^{\sqrt{2}r}r^{-a-1}dr\int_0^{\pi/2}\sin\beta(\cos\beta)^{-b+1}d\beta\\&=q^{-1}\frac{(\sqrt{2}r)^{-a}}{-a}\int_0^1 t^{-b+1}dt=q^{-1}\frac{(\sqrt{2}r)^{-a}}{-a}\frac{1}{2-b}.\end{aligned}$$

If $k(y)=\frac{y_1y_2(y_1^2-y_2^2)}{|y|^8}$, we get $|k(y)|\le\frac{1}{4}\frac{\sin 4\beta}{r^4}$. Since $b\le 0$, if $a<-2$, we get $\varphi\ge qr^a$ and

$$\begin{aligned}I_{11}&\le q^{-1}\int_0^{\sqrt{2}r}\int_0^{\pi/2}\frac{1}{4}\frac{|\sin 4\beta|}{s^4}s^{-a}sdsd\beta=\frac{1}{4q}\int_0^{\sqrt{2}r}s^{-a-3}ds\frac{1}{4}\int_0^{2\pi}|\sin\beta|d\beta\\&=\frac{1}{4q}\frac{(\sqrt{2}r)^{-a-2}}{-2-a}\int_0^{\pi/2}\sin\beta d\beta=\frac{1}{4q}\frac{(\sqrt{2}r)^{-a-2}}{-2-a}.\end{aligned}$$

4.4.2. Far-field estimate. Denote $a\vee b=\max(a,b)$. To estimate the far-field integral $I\triangleq\int_{y_1\vee y_2\ge R_0}|k(y)|\varphi^{-1}(y)dy$, we first pick a sufficiently large R and then partition the domain

$$0=z_0<z_1<\cdots<z_m=R_0<z_{m+1}<\cdots<z_n=R_1<+\infty.$$

Denote $Q_{ij}=[z_{i-1},z_i]\times[z_{j-1},z_j]$. Clearly, we have

$$\begin{aligned}&I=\sum_{m+1\le\max(i,j)\le n}I_{ij}+J,\quad I_{ij}\triangleq\int_{Q_{ij}}|k(y)|\varphi^{-1}(y)dy,\\&J=\int_{y_1\vee y_2\ge R_1}|k(y)|\varphi^{-1}(y)dy.\end{aligned}$$

For I_{ij}, we apply the trivial estimate (4.69). Suppose that

$$\varphi\ge qr^a(\cos\beta)^b,\quad |k(y)|\le|y|^{-p},\quad b\in[-1,0],\quad p+a>2.$$

We get

$$J\le\frac{1}{q}\int_{R_1}^{\infty}\int_0^{\pi/2}r^{-p-a}(\cos\beta)^{-b}rdrd\beta=\frac{1}{q}\frac{R_1^{-p-a+2}}{|p+a-2|}\int_0^{\pi/2}(\cos\beta)^{-b}d\beta.$$

Using Hölder's inequality and $b\in[-1,0]$, we get

$$\int_0^{\pi/2}(\cos\beta)^{-b}d\beta\le\left(\int_0^{\pi/2}\cos\beta d\beta\right)^{-b}\left(\int_0^{\pi/2}1\right)^{1+b}=(\pi/2)^{1+b}.$$

It follows that

$$J\le\frac{1}{q}\frac{R_1^{-p-a+2}}{|p+a-2|}(\pi/2)^{1+b}.$$

Application. We apply the above calculations to estimate the integral and its derivatives beyond the mesh $[0,D]^2$ (4.12). Since the domain is far away from the singularity, the integrand is the symmetrized kernel, e.g., (4.29). From Appendix B.1.1 and Lemma B.2 in Appendix B, for $\mathbf{u}_A, \nabla\mathbf{u}_A, \partial_i(\rho\mathbf{u}_A), \partial_i(\psi\nabla\mathbf{u}_A)$, the integrand in the far-field (y is large) satisfies

$$|K(x,y)| \le C(x)\mathrm{Den}^{-k}$$

with some $k \ge 2$ and coefficients $C(x)$, where Den is defined in (B.20).

In our computation, we rescale x to $\hat{x}$ and restrict it to the near-field $[0,b]^2$ with $b<2$. Note that $y \notin [0,D]^2$ and $|y| \ge D \gg b$. From (B.20), we get

$$\mathrm{Den} \ge \min_{|z_1|\le x_1, |z_2|\le x_2} |y-z|^2 \ge \min_{|z_1|\le x_1, |z_2|\le x_2} (|y|-|z|)^2 \ge (|y|-|x|)^2 = |y|^2\left(1-\frac{|x|}{|y|}\right)^2.$$

Since $\frac{|x|}{|y|} \le \sqrt{2}b/D$, we get

$$\mathrm{Den} \ge (1-C_s)^2|y|^2, \quad C_s = \sqrt{2}b/D.$$

It follows that

$$\int_{y\notin[0,D]^2} |K(x,y)|\varphi^{-1}(y)dy \le (1-C_s)^{-2k}C(x)\int_{y\notin[0,D]^2} |y|^{-2k}\varphi^{-1}(y)dy.$$

Using the method in section 4.4.2, we can estimate the above integral.

4.5. Estimate for very small or large x. The rescaling argument and the methods in the previous subsections apply to the estimate of $\mathbf{u}_A(x), \nabla\mathbf{u}_A(x)$ for $x \in [0,x_M]^2\backslash[0,x_m]^2, 0<x_m<x_M$. For very small or large x, we cannot use a finite number of dyadic scales $\lambda = 2^i$ to rescale x such that $x/\lambda \in [0,2x_c]^2\backslash[0,x_x]^2$. Instead, we choose $\lambda = \frac{\max(x_1,x_2)}{x_c}$. We want to estimate the rescaled integral with a $-d$-homogeneous kernel K

$$p(x)\int K(x-y)W(y)dy = p_\lambda(x)\int K(\hat{x}-\hat{y})\lambda^{2-d}W_\lambda(\hat{y})dy,$$

uniformly for all small $\lambda \ll 1$ or large $\lambda \gg 1$, where p is some weight and p_λ is defined in (4.2). The rescaled singularity $\hat{x} = x/\lambda$ satisfies $\max_i \hat{x}_i = x_c$. We simplify $\hat{x}, \hat{y}$ as x, y.

We can use the asymptotic of the weights to estimate the integral (see, e.g., (4.6).) The new difficulty is that the estimate involves the rescaled weight $p_\lambda(y)$. Since λ is not fixed and depends on x that tends to 0 or ∞, we cannot evaluate $p_\lambda(y)$ and the integrand directly. In the following derivation, λ is comparable to $|x|$, which is either very small or very large.

For y away from the singular region, the integrand of the regular part is given by $J = K(x,y)\cdot p_\lambda(x)$ (4.29). We choose a radial weight p defined in Appendix A.1 $p(x) = \sum_{1\le i\le n} q_i|x|^{a_i}$. See $\psi_1, \psi_u, \psi_{du}$ (A.1). We introduce the asymptotics of these weights

$$R_{\lim} \triangleq \lim_{x\to A}\frac{D_1 p_\lambda(x)}{p_\lambda(x)}, \quad p_{lim} = q_i|x|^{a_i},$$

with $(A,i)=(0,1)$ or $(A,i)=(\infty,n)$, where (q_n,a_n) denotes the last power in the weight. We use the following decomposition to compute D_1J with $D_1=\partial_{x_i}$:

$$\begin{aligned}|D_1J| &= |D_1(K(x,y)\cdot p_\lambda(x))| = |D_1K(x,y)\cdot p_\lambda(x)+K(x,y)\cdot D_1p_\lambda(x)|\\ &= \Big|p_\lambda(x)\Big\{D_1K(x,y)+R_{lim}K(x,y)+\left(\frac{D_1p_\lambda(x)}{p_\lambda(x)}-R_{lim}\right)K(x,y)\Big\}\Big|.\end{aligned}$$

Since we consider a very small λ or very large λ, the error term $\frac{D_1p_\lambda(x)}{p_\lambda(x)}-R_{lim}$ is small. Hence, we use a triangle inequality to bound D_1J:

$$|D_1J|\le p_\lambda(x)\Big|D_1K(x,y)+R_{lim}K(x,y)\Big|+p_\lambda(x)\Big|\left(\frac{D_1p_\lambda(x)}{p_\lambda(x)}-R_{lim}\right)K(x,y)\Big|.$$

The advantage of the above decomposition is that the main term $D_1K(x,y)+R_{lim}$ $K(x,y)$ does not depend on λ so that we can estimate it using previous methods.

Since the estimate of a derivative of u,v does not involve the commutator (see, e.g., (4.67)), we can apply the above method to compute the integral of D_1u for small x or large x.

For y near the singular region, from (4.28), the symmetrized integrand is given by

$$J=K^C(p_\lambda(x)-p_\lambda(y))+K^{NC}p_\lambda(x),$$

where we use p for the weight. First, we have

$$|D_1J|=|D_1K^C(p_\lambda(x)-p_\lambda(y))+D_1K^{NC}p_\lambda(x)+(K^C+K^{NC})D_1p_\lambda(x)|.$$

Denote $K=K^C+K^{NC}$. We use the following method to bound D_1J:

$$\begin{aligned}|D_1J| &\le p_\lambda(x)\left|D_1K^C\cdot\left(1-\frac{p_\lambda(y)}{p_\lambda(x)}\right)+D_1K^{NC}+K\cdot\frac{D_1p_\lambda}{p_\lambda}\right|\\ &\le p_\lambda(x)\Bigg\{\left|D_1K^C\cdot\left(1-\frac{p_{lim}(y)}{p_{lim}(x)}\right)+D_1K^{NC}+K\cdot\frac{D_1p_{lim}}{p_{lim}}\right|\\ &\quad+\left|D_1K^C\left(\frac{p_\lambda(y)}{p_\lambda(x)}-\frac{p_{lim}(y)}{p_{lim}(x)}\right)\right|+K\left|\frac{D_1p_{lim}}{p_{lim}}-\frac{D_1p_\lambda}{p_\lambda}\right|\Bigg\}.\end{aligned}$$

The second and the third terms on the right hand side can be seen as error terms. The main term $|D_1K^C\cdot(1-\frac{p_{lim}(y)}{p_{lim}(x)})+D_1K^{NC}+K\cdot\frac{D_1p_{lim}}{p_{lim}}|$ does not depend on λ, and the singularity x is in the near-field and away from 0. We can apply all the delicate decompositions developed in previous sections to estimate D_1J.

In the Hölder estimates, we need various bounds for the weights p_λ. Using the asymptotics of $p(x)$, we can estimate the derivatives of p_λ for very small λ or very large λ uniformly. See Appendices A.1, A.2. Once we obtain the estimates of ψ_λ and the weight φ_λ in the L^∞ norm $||\omega_\lambda\varphi_\lambda||_\infty$, we can use the methods in the previous subsections and the scaling relations in section 4.1.2 to perform the Hölder estimates.

The L^∞ estimate follows similar ideas and is much easier. We defer more details to section 7 in the supplementary material (supplement.pdf [local/web 1.43MB]).

We remark that since we have much larger damping coefficients in the energy estimates (see section 5 in Part I [13]) near $x=0$ and in the far-field, the estimates of the nonlocal terms in these regions, though technical, only have minor effects on the nonlinear stability estimates.

4.6. Assemble the Hölder estimates. In section 4.3, we decompose the velocity in several parts and estimate them separately using the norms $||\omega\varphi||_\infty, [\omega\psi]_{C^{1/2}_{x_i}}$. In this section, we assemble these estimates and estimate

$$\delta(f,x,z) \triangleq \frac{|f(x)-f(z)|}{|x-z|^{1/2}}$$

for $f = \psi_u \mathbf{u}_A, \psi\nabla\mathbf{u}_A$ with weights in (A.1). To obtain better estimates, we combine some of the estimates.

In the proof of the first inequality in Lemma 2.3, we combine and bound different norms using $\max(||\omega\varphi||_\infty, \max_{j=1,2}\gamma_j[\omega\psi_1]_{C^{1/2}_{x_j}(\mathbb{R}^+_2)})$. We apply the second inequality to the error $\varepsilon = \omega - (-\Delta)\phi^N$ (3.10) and can evaluate the localized norm using piecewise bounds of the error. See section 4.7.

To illustrate the ideas, we focus on the $C^{1/2}_x$ estimate, $x \in [x_c, 2x_c]\times[0,2x_c]$, i.e., x_1 is large relative to x_2, $z_1 \geq x_1$, and $x_2 = z_2$. For general pairs (x,z), we can rescale (x,z) to $(\lambda x, \lambda z)$ such that $\lambda x \in [0,2x_c]^2\backslash[0,x_c]^2$. Using the scaling relations in (4.1.2), we can estimate the rescaled version of $\delta(f,x,z)$. See also the discussion at the beginning of section 4.3.

We assume that $z_1 \in [x_c, 2(1+\nu)x_c]$ with $\nu < 1$. For $z_1 \geq 2(1+\nu)x_c$, we have $z_1 > (1+\nu)x_1$. Since z_1, x_1 are large relative to z_2, x_2, respectively, we have

$$|x-z| = |z_1 - x_1| \asymp |z_1| \gtrsim |x|, |z|.$$

Then, we can use the L^∞ estimate and triangle inequality to estimate $\delta(f,x,z)$. Note that we can estimate the piecewise L^∞ norm of $|x|^{-1/2}\rho(x)\mathbf{u}_A(x)$ and $|x|^{-1/2}\psi\nabla\mathbf{u}_A$ following section 4.2, where ρ, ψ are the weights in the Hölder estimate of $\rho\mathbf{u}_A, \psi\nabla\mathbf{u}_A$. See section 7.4 in the supplementary material (supplement.pdf [local/web 1.43MB]) for more details.

We focus on $f = \psi u_{x,A}$. We partition the domain $D_\nu = [x_c, 2(1+\nu)x_c]\times[0,2x_c]$ into $h_x \times h_x$ grids $D_{ij}, 1 \leq i \leq 2(1+\nu)x_c/h_x, 1 \leq j \leq 2x_c/h_x$. We apply the decomposition (4.67) with the same parameters k, k_2 to x in different grids D_{ij}. For $x \in D_{ij}$, using the method in section 4.3, we obtain the estimate

(4.70)
$$f(x) = I_1(x) + I_2(x) + I_3(x) + I_4(x) + I_5(x) - I_6(x), \quad I_5 = I_{5,0} + I_{5,1} + I_{5,2},$$
$$|\partial_x(I_1 + I_4 + I_{5,0} - I_6)| \leq a_{ij,1}||\omega\varphi||_\infty,\ |\partial_x I_3| \leq a_{ij,2}||\omega\varphi||_\infty,\ |\partial_x I_{5,1}| \leq a_{ij,3}||\omega\varphi||_\infty$$

for some constants $a_{ij,l}, b_{ij} \geq 0$, where $I_{5,1}, I_{5,2}$ are defined and estimated in section 4.3.4.

For $x, z \in D_\nu$ with $x_2 = z_2, z_1 \leq z_1$, we have $x \in D_{i_1,j}, z \in D_{i_2,j}$ for some $i_1 \leq i_2$. We apply the method in section 4.3.3 to estimate $\delta(I_2,x,z)$ and the method in section 4.3.4 to estimate J_1 related to $\delta(I_{52},x,z)$ (4.53). These estimates contribute to the bound $C_{hol}[\omega\psi]_{C^{1/2}_x}$ for some $C_{hol} > 0$, which can be computed.

Regularity of the combination. While $I_1 + I_4 + I_{5,0} - I_6, I_3, I_{5,1}$ are only piecewise smooth and can be discontinuous when x crosses the grids D_{ij}, the sum $I_{lip} = I_1 + I_4 + I_{5,0} - I_6 + I_3 + I_{5,1}$ is continuous and Lipschitz in x_1 for fixed x_2. In fact, by definition (4.24), (4.51), we get

$$
\begin{aligned}
I_{lip} = &\int_{R(k)\setminus R_{s,1}(k_3)} K_1(x-y)(\psi(x)-\psi(y))W(y)dy \\
&+ \psi(x)\int_{R(k)^c} K_1(x-y)W(y)dy + \int_{R(k)\setminus R_{s,1}(k)} K_1(x-y)\psi(y)W(y)dy - I_6 \\
= &\psi(x)\int_{R_{s,1}^c(k_3)} K_1(x-y)W(y)dy - \int_{R_{s,1}(k)\setminus R_{s,1}(k_3)} K_1(x-y)W(y)\psi(y)dy - I_6.
\end{aligned}
$$

For fixed x_2, since I_6 (4.50) for the approximation term is smooth in x and the domain $R_{s,1}(l)$ (4.19) depends on x_1 continuously, we obtain that I_{lip} is continuous in x_1 when x crosses the grids D_{ij}. Since $I_{lip}(x)$ is smooth for $x \in D_{ij}$, we get that I_{lip} is continuous and Lipschitz in x_1 with piecewise Lipschitz norm bounded by $a_{ij,1} + a_{ij,2} + a_{ij,3}$.

Similarly, for the case in section 4.3.4, we have $I_{lip} = I_1 + I_3 + I_4 + I_{5,0} + I_{5,1} - I_6$ (4.24), (4.51) is Lipschitz in x_i in the $C_i^{1/2}$ estimate for fixed $x_{3-i}, i = 1, 2$.

For the case in section 4.3.6, $I_{lip} = I_1 + I_3 + I_4 + I_{5,0}^- + I_{5,1}^- + I_{5,1}^+ - I_6$ (4.24), (4.56), (4.57) and $I_{lip} = I_1 + I_3 + I_4 + I_{5,3}^- + I_{5,1}^+ - I_6$ (4.24), (4.56), (4.60) are Lipschitz, where $I_{5,1}^+$ associated with I_5^+ (4.56) is defined similarly to $I_{5,1}$ in (4.51).

For the case in section 4.3.7, $I_{lip} = I_1 + I_3 + I_4 + I_{5,0} + I_{5,1} + I_{5,2,1} - I_6$ (4.24), (4.63), and $I_{lip} = I_1 + I_3 + I_4 + I_{5,4} - I_6$ are Lipschitz, where $I_{5,2,1}$ associated with $I_{5,2}$ (4.63) is defined similar to $I_{5,1}$ in (4.51).

In summary, the sum of the terms in $f(x)$ (4.70) with piecewise derivative estimates is Lipschitz. Using the triangle inequality, we obtain the piecewise Lipschitz bound for I_{lip}. The remaining parts in $f(x)$ (4.70) are continuous and are estimated by the piecewise L^∞ bounds, e.g., $I_{5,2}^-$ (4.57), $I_{5,4}^-(a)$ (4.60), $I_{5,3}$ (4.63), and the improved Hölder estimates, e.g., $I_{5,2}$ (4.51)

By averaging the piecewise derivative bounds and using the estimates in Appendix E.2, for $x \in D_{i_1,j}, z \in D_{i_2,j}$, we can obtain

$$
|I_{lip}(x) - I_{lip}(z)| \le C_{lip}|x_1 - z_1| \cdot ||\omega\varphi||_\infty
$$

for constant C_{lip} depending only on $\{a_{kl,j}\}_{k,l\ge 1, j\le 3}$ and the mesh h_x explicitly. Hence, for the remaining terms in f not estimated using the seminorm $[\omega\psi]_{C_x^{1/2}}$, e.g., $I_1 + I_4 - I_6 + I_3 + I_{5,0} + I_{5,1}$ and J_2 related to $I_{5,2}$ (4.53), each term is continuous and they satisfy[3]

$$
f_R(x) = \sum_{1\le l\le N} f_l(x), \quad |f_l(x) - f_l(z)| \le \min(p_l|x_1 - z_1|, q_l) \cdot ||\omega\varphi||_\infty
$$

for some N, where we can choose $q_l = \infty$ if we do not have an L^∞ estimate for $f_l(x)$. A similar consideration applies to p_l. In our problem, there are only a few terms and $N < 10$. In the $C_{x_i}^{1/2}$ Hölder estimate of $P_{ij}Q_{ij}(x)$ (continuous in x_i) in I_{52} (4.53), we optimize two estimates (see the estimates between (4.53) and (4.54)), which is a nontrivial example of the above summand.

[3]In the previous version of this paper [11], some term $f_l(x)$ is not continuous when x crosses the grids. We have corrected this minor issue by reorganizing different terms so that each $f_l(x)$ is continuous. See the above paragraph "*Regularity of the combination.*" Related computer-assisted estimates have been updated and the full nonlinear stability estimates remain valid.

Now, for $x \in D_{i_1,j}, z \in D_{i_2,j}$, we have

$$\frac{|f_R(x)-f_R(z)|}{|z_1-x_1|^{1/2}} \le \sum_{1\le l\le N} \min(p_l\delta^{1/2}, q_l\delta^{-1/2})||\omega\varphi||_\infty, \tag{4.71}$$
$$\delta = z_1 - x_1 \in [\max(i_2-i_1-1,0)h_x, (i_2-i_1+1)h_x].$$

The upper bound can be obtained explicitly by partitioning the range of $z_1 - x_1$ into finite many subintervals M_l according to the threshold $\delta_l = q_l/p_l$. In each M_l, the bound reduces to

$$P\delta^{1/2} + Q\delta^{-1/2}$$

for some constants P, Q. It is convex in $\delta^{1/2}$ and can be optimized easily and explicitly in any interval $[\delta_l, \delta_u], \delta_l > 0$.

Remark 4.8. We combine the estimates of different parts in (4.70) using (4.71) to obtain a sharp estimate. If one estimates different parts separately, the distance $\delta = z_1 - x_1$ for the optimizer may not be achieved for the same value, which leads to an overestimate. We remark that for a small distance $|z_1 - x_1|$, such an overestimate can be significant since the ratio between the endpoints $|i_2 - i_1 + 1|/\max(i_2 - i_1 - 1, 0)$ varies a lot.

In some estimates, e.g., the $C_y^{1/2}$ estimate of u_x in section 4.3.6, we need to decompose I_5 using a different size of small singular region k_3. In such a case, we have a list of estimates associated to different k_3 for the part f_R not estimated by $[\omega\psi]_{C_x^{1/2}}$ or $[\omega\psi]_{C_y^{1/2}}$:

$$\frac{|f_R(x)-f_R(z)|}{|z_1-x_1|^{1/2}} \le \sum_{1\le l\le N} \min(p_{l,k_3}\delta^{1/2}, q_{l,k_3}\delta^{-1/2})||\omega\varphi||_\infty.$$

For $|x_1 - z_1|$ bounded away from 0, e.g., $|x_1 - z_1| \ge \frac{1}{10}h_x$, we can still partition the range of $|x_1 - z_1|$ and optimizing the above estimates first over δ and then k_3.

4.6.1. Hölder estimate for small distance. In some Hölder estimates, e.g., the $C_y^{1/2}$ estimate in sections 4.3.6, 4.3.7, when $|x-z|$ is very small, e.g., $|x-z| \le ch_x$ with $c < 1$, we need to choose a singular region with size a to be arbitrarily small. See also section 4.1.11 for the estimates of a log-Lipschitz function. In these estimates, we can decompose $f_R(x)$ that is not estimated using the Hölder norm of $\omega\psi$ as follows:

$$f_R(x) = f_1(x, a, b) + f_2(x, a)$$

for $a < b$ and b is fixed. We can estimate the derivative of f_1, and the L^∞ norm for f_2

$$|\partial_x f_1(x,a,b)| \le \Big(A_i + B_i \log\frac{b}{a}\Big)||\omega\varphi||_\infty, \quad |f_2| \le \frac{C_i a}{2}||\omega\varphi||_\infty$$

in each grid D_{ij} for any $a \le b$ (see, e.g., (4.61) and (4.70)). We drop j since we consider x, z with $x_2 = z_2$. For $t = |x - z| \le h_x$, we get

$$\frac{|f(x)-f(z)|}{|x-z|^{1/2}} \le \Big(A + B\log\frac{b}{a}\Big)\sqrt{t} + \frac{Ca}{\sqrt{t}} \triangleq F(a,t), \tag{4.72}$$

where $A = \max(A_i, A_{i+1}), B = \max(B_i, B_{i+1}), C = \max(C_i, C_{i+1})$. For each $t \le ch_x$, we can optimize the above estimate over $a \le b$ explicitly. Then we maximize the estimate over $t \le ch_x$ to obtain a uniform estimate for small $|x - z| \le ch_x$. We defer the derivations to Appendix B.5.2.

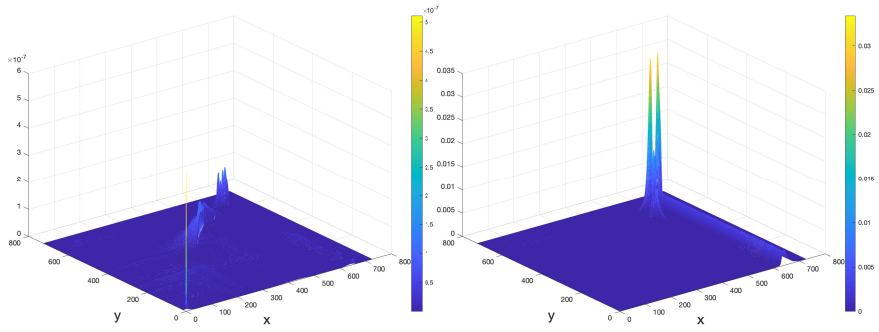


FIG. 4. *Piecewise $L^\infty(\varphi_{elli})$ bound of the error $\bar{\varepsilon}_1, \hat{\varepsilon}_1$ in solving the Poisson equations. Left: Error for the approximate steady steate. Right: Error for the approximate space-time solution $\hat{W}_2$.*

4.7. Improved estimate for the nonlocal error. In section 3.7, we discuss the estimates of the nonlocal error $\mathbf{u}(\bar{\varepsilon})$ based on the functional inequalities established in this section. Since the weight is singular $\varphi \sim |x|^{-2}|x_1|^{-1/2}, \varphi = \varphi_{elli}$ (A.2) near the origin, $\bar{\varepsilon}_1\varphi$ is much larger near $x = 0$. Due to the anisotropic mesh for large x and small y, or small x and large y, and the round-off error, $\bar{\varepsilon}_1$ is not very small in these far-field regions. On the other hand, these regions are small since either $|(x, y)|$ is very small or the ratio $x/y, y/x$ is very small, and the error is very small in the bulk, e.g., $x = O(1)$. See Figure 4 for the rigorous weighted bound of the error in the adaptive mesh. The weighted error of $\bar{\varepsilon}_1$ is larger near 0, while the error for $\hat{\varepsilon}_1$ is larger in the far-field. If we simply use the global norm $||\omega\varphi||_\infty, \omega = \bar{\varepsilon}, \hat{\varepsilon}$, and then apply the previous estimates to bound $\mathbf{u}(\bar{\varepsilon})$, we overestimate the nonlocal error significantly. For $x = O(1)$, where we have the smallest damping for the energy estimate, due to the decay of kernel and the smallness of these regions, the integral $\int K(x, y)\bar{\varepsilon}(y)dy$ near $y = 0$ or in the far-field is very small.

Note that we can obtain the piecewise derivative bounds for the error $\bar{\varepsilon}_1, \hat{\varepsilon}_1$ and we partition the domain of the integral into different regions (4.45). Instead of using the global norm to bound the integral, we use the localized norms $||W\varphi_{elli}||_{l^\infty(D)}$, $[W\psi_1]_{C^{1/2}_{x_i}(D)}$ (A.2), (A.1) to exploit the smallness of the error in most parts of the domains and improve the error estimate.

Recall the regions of rescaled $\hat{x}$ (4.4) and the mesh y_i partitioning the domain (4.11). We fix a scale λ and assume $\hat{\in}[x_c, 2x_c] \times [0, 2x_c]$. By definition, the singular region $R(\hat{x}, k)$ (4.18) satisfies

$$-R(\hat{x}, k) \cap \mathbb{R}^+_2, \ R(\hat{x}, k) \cap \mathbb{R}^+_2 \subset [x_c - kh, 2x_c + kh] \times [0, 2x_c + kh] \triangleq S_{kh}.$$

Thus, in the estimates of I_2, I_3, I_4 in (4.45), instead of using the global norm $||W\varphi||_{L^\infty}$, we use $||\omega_\lambda\varphi_\lambda||_{L^\infty(S_{kh})} = ||\omega\varphi||_{L^\infty(\lambda S_{kh})}$. For the error $\omega = \bar{\varepsilon}, \hat{\varepsilon}$, we can bound $||\omega\varphi||_{L^\infty(\lambda S_{kh})}$ by using the piecewise estimates of $\bar{\varepsilon}, \hat{\varepsilon}$ and covering the region λS_{kh}. Similarly, we use the localized bound $[\omega_\lambda\psi_\lambda]_{C^{1/2}_{x_i}(S_{kh})} = \lambda^{1/2}[\omega\psi]_{C^{1/2}_{x_i}(\lambda S_{kh})}$ for the Hölder seminorm in the estimate of I_2, I_3, I_4, and similar localized norms for I_5.

For the regular part I_1, we partition $[0, D]^2, \mathbb{R}^{++}_2$ into disjoint domains: near-field $D_{n,i}$, the bulk D_B, and the far-field $D_{f,i}$, e.g.,

$$D_{n,1} = [8h, 16h], \ D_B = [0, 2]^2 \backslash D_{n,1}, \ D_{f,1} = [0, D]^2 \backslash [0, 2]^2, \ D_{f,2} = \mathbb{R}^{++}_2 \backslash [0, D]^2,$$

where h is the mesh size in (4.11). Then we use the norm $||\omega_\lambda\varphi_\lambda||_{L^\infty(D)} = ||\omega\varphi||_{L^\infty(\lambda D)}$ for the estimate of the integral in region D.

In (4.8), we estimate the integral of $K_{00}(y)$ (4.5) for $|\hat{y}|_\infty \le k_{02}h$ and $|\hat{y}|_\infty \ge k_{02}h$ separately. Since the kernel is very singular near 0, the L^1 estimate of the integral in $|\hat{y}|_\infty \le k_{02}h$ in section 4.4.1 is not very small. Since we can evaluate $\omega = \bar{\varepsilon}, \hat{\varepsilon}$, we change the rescaling from $\hat{y}$ back to y by using $y = \lambda\hat{y}$ in (4.8),

$$J = \int_{|\hat{y}|_\infty \le k_{02}h} K_{00}(\hat{y})\omega(\lambda\hat{y})d\hat{y} = \lambda^2 \int_{|y|_\infty \le \lambda k_{02}h} K_{00}(y)\omega(y)dy,$$

where we get λ^2 since K_{00} is -4 homogeneous. For a list of dyadic scales $\lambda = 2^k$, we estimate the integral using Simpson's rule with very small mesh. This allows us to exploit the cancellation in the integral. For $|y|$ very close to 0, we use Taylor expansion. See section 6.4.1 in the supplementary material (supplement.pdf [local/web 1.43MB]) (attached to this paper) for more details.

In the estimate of the integral for very small x or large x in section 4.5 (see more details in section 7 in the supplementary material (supplement.pdf [local/web 1.43MB])), we estimate the rescaled integral for $\lambda \le \lambda_1$ and $\lambda \ge \lambda_n$ with small λ_1 and large λ_n uniformly. In the case of $\lambda \le \lambda_1$, we bound $||\omega_\lambda\varphi_\lambda||_{L^\infty([a,b]\times[c,d])} \le ||\omega\varphi||_{L^\infty(\lambda_1[0,b]\times[0,d])}$. Other norms in different cases are estimated similarly.

We do not track the bound $||\omega_\lambda\varphi_\lambda||_{L^\infty(Q_{ij})}$ in each small grid Q_{ij} for computational efficiency.

Appendix A. Weights and parameters.

A.1. Estimate of the weights. Recall the weights for the Hölder estimate of ω, η, ξ, and $\mathbf{u}$,

$$\begin{aligned} &\psi_1 = |x|^{-2} + 0.5|x|^{-1} + 0.2|x|^{-1/6}, \quad \psi_{du} = \psi_1, \quad \psi_u = |x|^{5/2} + 0.2|x|^{-7/6}, \\ &\psi_2 = p_{2,1}|x|^{-5/2} + p_{2,2}|x|^{-1} + p_{2,3}|x|^{-1/2} + p_{2,4}|x|^{1/6}, \\ &\psi_3 = \psi_2, \quad \vec{p}_{2,\cdot} = (0.46, 0.245, 0.3, 0.112), \end{aligned} \tag{A.1}$$

and the following weights for ω, ρ_i for $\mathbf{u}$ and the error,

$$\begin{aligned} &\varphi_1 = x^{-1/2}(|x|^{-2.4} + 0.6|x|^{-1/2}) + 0.3|x|^{-1/6}, \quad \varphi_{g1} = \varphi_1 + |x|^{1/16}, \\ &\varphi_{elli} = |x_1|^{-1/2}(|x|^{-2} + 0.6|x|^{-1/2}) + 0.3|x|^{-1/6}, \quad \rho_{10} = |x|^{-3} + |x|^{-7/6}, \quad \rho_{20} = \psi_1. \\ &\rho_3 = |x|^{-1} + |x|^{-1/6}, \quad \rho_4 = x^{-1/2}(|x|^{-2.5} + 0.6|x|^{-1/2}) + 0.3|x|^{-1/6}. \end{aligned} \tag{A.2}$$

To estimate the weighted L^∞ norm of the residual error in section 3, we use $\psi_i, \varphi_{evo,i}$,

$$\begin{aligned} &\varphi_{evo,1} = \varphi_1, \\ &\varphi_{evo,2} = x^{-1/2}(\tilde{p}_{5,1}|x|^{-5/2} + \tilde{p}_{5,2}|x|^{-3/2} + \tilde{p}_{5,3}|x|^{-1/6}) + \tilde{p}_{5,4}|x|^{-1/4} + \tilde{p}_{5,5}|x|^{1/7}, \\ &\varphi_{evo,3} = x^{-1/2}(\tilde{p}_{6,1}|x|^{-5/2} + \tilde{p}_{6,2}|x|^{-3/2} + \tilde{p}_{6,3}|x|^{-1/6}) + \tilde{p}_{6,4}|x|^{-1/4} + \tilde{p}_{6,5}|x|^{1/7}, \\ &\tilde{p}_{5,\cdot} = (0.42,\ 0.135,\ 0.216,\ 0.182,\ 0.0349) \cdot \mu_0, \quad \mu_0 = 0.917, \\ &\tilde{p}_{6,\cdot} = (2.5 \cdot \tilde{p}_{5,1}, 2.9 \cdot \tilde{p}_{5,2},\ 3.115 \cdot \tilde{p}_{5,3},\ 1.82 \cdot \tilde{p}_{5,4},\ 2.72 \cdot \tilde{p}_{5,5}), \end{aligned} \tag{A.3}$$

where φ_1 is defined in (A.2).

In our energy estimates and the estimates of the nonlocal terms, we need various estimates of the weights and their derivatives. From Appendix C.1 of Part I [13] and (A.2), (A.1), we have two types of weights. The first one is the radial weight

$$\rho(x,y)=\sum_i p_i r^{a_i},\quad r=(x^2+y^2)^{1/2},$$

where a_i is increasing and $p_i\geq 0$. We use these weights for the Hölder estimates. See, e.g., (A.1).

The second type of weight is the following:

$$\rho(x,y)=\rho_1(r)|x|^{-\alpha}+\rho_2(r),$$

where ρ_1,ρ_2 are the radial weights.

We use f_l, f_u to denote the lower and upper bound of f. We have the following simple inequalities:

$$\begin{aligned}&(f-g)_l=f_l-g_u,\quad (f-g)_u=f_u-g_l,\quad (f+g)_\gamma=f_\gamma+g_\gamma,\\ &(fg)_l=\min(f_lg_l,f_ug_l,f_lg_u,f_ug_u),\quad (fg)_u=\max(f_lg_l,f_ug_l,f_lg_u,f_ug_u),\end{aligned}\tag{A.4}$$

where $\gamma=l,u$. If $g\geq 0$, we can simplify the formula for the product

$$(fg)_l=\min(f_lg_l,f_lg_u),\quad (fg)_u=\max(f_ug_l,f_ug_u).\tag{A.5}$$

Given the piecewise bounds of $\partial^j f,\partial^j g, j\leq k$, we can estimate $\partial^k(fg)$ using the Leibniz rule

$$|\partial_x^i\partial_y^j(fg)|\leq\sum_{k\leq i,l\leq j}\binom{i}{k}\binom{j}{l}|\partial_x^k\partial_y^l f|\cdot|\partial_x^{i-k}\partial_y^{j-l}g|.\tag{A.6}$$

A.2. Radial weights. The advantage of radial weights ρ is that we can estimate them easily. Since $\rho(x,y)$ is even in x,y, we restrict the estimate of piecewise bounds to the case of $x\geq 0, y\geq 0$. The bound in general domain $D=[a,b]\times[c,d]$ can be obtained by decomposing D into four quadrants and then using the symmetry and combining the bounds from different quadrants.

A.2.1. Bounds for the derivatives. We can easily derive the derivatives and their upper and lower bound as follows. First, we have

$$(\partial_x^i\partial_y^j\rho(x,y))_\gamma=\sum_{1\leq k\leq n}p_k(\partial_x^i\partial_y^j r^{a_k})_\gamma,\tag{A.7}$$

where $\gamma=l,u$. Using induction, for any α,i,j, we can obtain

$$\begin{aligned}\partial_x^i\partial_y^j r^\alpha&=\sum_{k\leq i+j,l\leq\min(j,1)}C_{i,j,k,l}(\alpha)x^ky^lr^{\alpha-i-j-k-l}\\ &=\sum_{k\leq i+j,l\leq\min(j,1)}(C^+_{i,j,k,l}(\alpha)-C^-_{i,j,k,l}(\alpha))x^ky^lr^{\alpha-i-j-k-l}\end{aligned}$$

with $C^\pm_{i,j,k,l}(\alpha)\triangleq\max(0,C_{i,j,k,l}(\alpha))$. The bounds for $C^\pm_{i,j,k,l}(\alpha)x^ky^lr^{\alpha-i-j-k-l}$ are simple:

$$(C^\pm_{i,j,k,l}(\alpha)x^ky^lr^{\alpha-i-j-k-l})_\gamma=C^\pm_{i,j,k,l}(\alpha)x^k_\gamma y^l_\gamma r^{\alpha-i-j-k-l}_\gamma.\tag{A.8}$$

In particular, we use the derivatives bound for $i+j\le 4$ and we have

$$\begin{aligned}
\partial_x r^a &= axr^{a-2}, \quad \partial_x^2 r^a = ar^{a-2} + a(a-2)x^2 r^{a-4}, \quad \partial_{xy} r^a = a(a-2)xyr^{a-4},\\
\partial_x^3 r^a &= a(a-2)(a-4)x^3 r^{a-6} + 3a(a-2)xr^{a-4},\\
\partial_x^2\partial_y r^a &= a(a-2)yr^{a-4} + a(a-2)(a-4)x^2 yr^{a-6},\\
\partial_x^4 r^a &= 3a(a-2)r^{a-4} + 6a(a-2)(a-4)x^2 r^{a-6} + a(a-2)(a-4)(a-6)x^4 r^{a-8},\\
\partial_x^3\partial_y r^a &= a(a-2)(a-4)xyr^{a-6} + 2a(a-2)(a-4)xyr^{a-6}\\
&\quad + a(a-2)(a-4)(a-6)x^3 yr^{a-8},\\
\partial_x^2\partial_y^2 r^a &= a(a-2)(a-3)r^{a-4} + a(a-2)(a-4)(a-6)x^2 r^{a-6}\\
&\quad - x^4 a(a-2)(a-4)(a-6)r^{a-8}.
\end{aligned}$$

Using (A.4), the above identities, and linearity, we can obtain the upper and lower bounds for $\partial_x^i\partial_y^j\rho$. Since $\rho(x,y)$ is symmetric in x,y, we have $\partial_1^i\partial_2^j\rho(x,y) = (\partial_1^j\partial_2^i\rho)(y,x)$ and can obtain piecewise bounds of $\partial_1^i\partial_2^j\rho$ from that of $\partial_1^j\partial_2^i\rho$.

For the estimate in section 4.5, we need to use the estimates of $\partial_x^i\partial_y^j\rho(\lambda x)$ for very small $\lambda\le\lambda_*$ or very large $\lambda\ge\lambda_*$ uniformly. Obviously, the bounds are mainly determined by the leading order power of $p(\lambda x)$, i.e., $p_1|\lambda r|^{a_1}$ for small λ and $p_n|\lambda r|^{a_n}$ for large λ. We would like to estimate $(\partial_x^i\partial_y^j\rho(\lambda x))_\gamma\lambda^{-\beta}$ for $\lambda\le\lambda_*, \beta = a_1$ and $\lambda\ge\lambda_*, \beta = a_n$, $\gamma = l, u$. Using the above derivations (A.7), we have

$$\lambda^{-\beta}(\partial_x^i\partial_y^j\rho(x,y))_\gamma = \sum_{1\le k\le n} p_k(\partial_x^i\partial_y^j\lambda^{a_k-\beta}r^{a_k})_\gamma, \quad \gamma = l, u,$$

and we only need to derive the upper and the lower bounds for $C^{\pm}_{i,j,k,l}(a_m)x^k y^l$ $r^{\alpha-i-j-k-l}\lambda^{a_m-\beta}$ uniformly for $\lambda\le\lambda_*, \beta = a_1$ or $\lambda\ge\lambda_*, \beta = a_n$. Since a_i is increasing, in the first case, we have

$$\lambda^{a_1-a_1} = 1, \quad a_m - a_1 > 0, \quad (\lambda^{a_m-a_1})_l = 0, \ (\lambda^{a_m-a_1})_u = \lambda_*^{a_m-a_1}, m > 1.$$

In the second case, we get

$$\lambda^{a_n-a_n} = 1, \quad a_m - a_n < 0, \quad (\lambda^{a_m-a_n})_l = 0, \ (\lambda^{a_m-a_n})_u = \lambda_*^{a_m-a_n}, m > 1.$$

In both cases, if $a_m = \beta$, we get a trivial bound 1 for $\lambda^{a_m-\beta}$; if $a_m\ne\beta$, we get $0\le\lambda^{a_m-\beta}\le\lambda_*^{a_m-\beta}$. Using these bounds for $\lambda^{a_m-\beta}$, (A.8), (A.4), (A.5), we obtain the bounds for $\lambda^{-\beta}\partial_x^i\partial_y^j\psi(\lambda x)$ uniformly for small $\lambda, \beta = a_1$ and large $\lambda, \beta = a_n$.

We also need to bound $M = \lambda^{-\beta}\rho_\lambda(x)\left|\frac{\rho_\lambda(y)}{\rho_\lambda(x)} - \frac{\rho_{lim}(y)}{\rho_{lim}(x)}\right|$ used in section 4.5, uniformly for $\lambda\le\lambda_*, \beta = a_1, \rho_{lim}(y) = p_1|y|^{a_1}$ or $\lambda\ge\lambda_*, \beta = a_n, \rho_{lim}(y) = p_n|y|^{a_n}$. Using the formula of ρ and a direct computation yield

$$\begin{aligned}
\frac{\rho_{lim}(y)}{\rho_{lim}(x)} = \frac{|y|^\beta}{|x|^\beta}, \quad M &\le \sum_{i\le n} p_i\lambda^{a_i-\beta}\left||y|^{a_i} - |x|^{a_i}\frac{|y|^\beta}{|x|^\beta}\right|\\
&\le \sum_{i\le n} p_i\lambda_*^{a_i-\beta}|y|^\beta\left||y|^{a_i-\beta} - |x|^{a_i-\beta}\right|.
\end{aligned}$$

We remark that the leading power $\lambda_*^{a_i-\beta}$ for $a_i = \beta$ is cancelled due to $|y|^0 = |x|^0 = 1$ in the above estimate and we gain the small factor $\lambda_*^{a_i-\beta}$ for $a_i\ne\beta$.

A.2.2. Leading order behavior of $\partial\rho/\rho$. In our verification, we need to bound $\partial\rho(\lambda x)/\rho(\lambda x)$ as $\lambda\to 0$ or $\lambda\to\infty$ uniformly. A direct calculation yields

$$\frac{\partial_{x_i}\rho}{\rho}=\frac{x_i}{|x|^2}\frac{\sum_i p_i a_i r^{a_i}}{\sum_i p_i r^{a_i}}\triangleq\frac{x_i}{|x|^2}S(x),\quad S(x)\triangleq\frac{\sum_i p_i a_i r^{a_i}}{\sum_i p_i r^{a_i}}.$$

For x close to 0, we introduce $b=a-a_1$. Clearly, we get $b_i\geq 0$ and

$$S(x)=a_1+\frac{\sum_i p_i b_i r^{a_i}}{\sum_i p_i r^{a_i}}=a_1+\frac{\sum_i p_i b_i r^{b_i}}{\sum_i p_i r^{b_i}}\triangleq a_1+\frac{A(r)}{B(r)}.$$

Using $b_i\geq 0$ and the Cauchy–Schwarz inequalities, we get

$$\begin{aligned}A'B-AB'&=r^{-1}\Big(\Big(\sum p_i b_i^2 r^{b_i}\Big)\Big(\sum p_i r^{b_i}\Big)-\Big(\sum p_i b_i r^{b_i}\Big)^2\Big)\\&=r^{-1}\frac{1}{2}\sum_{ij}p_i p_j(b_i-b_j)^2 r^{b_i+b_j}\geq 0,\end{aligned}$$

and thus A/B is increasing. For $\lambda\leq\lambda_*, r\in[r_l,r_u]$, we get the uniform bound for $S(\lambda x)$

$$a_1\leq S(\lambda x)\leq a_1+\frac{A(\lambda_* r_u)}{B(\lambda_* r_u)}.$$

For $\lambda=1$, we simply obtain

$$a_1+\frac{A(r_l)}{B(r_l)}\leq S(x)\leq a_1+\frac{A(r_u)}{B(r_u)}.$$

Similarly, for $\lambda\geq\lambda_*$, $r\in[r_l,r_u]$, we get

$$a_n+\frac{A(\lambda_* r_l)}{B(\lambda_* r_l)}\leq S(\lambda x)\leq a_n,\quad \frac{A(r)}{B(r)}=\frac{\sum_i p_i b_i r^{b_i}}{\sum_i p_i r^{b_i}},$$

where $b=a-a_n\leq 0$. Here, we have used that $A(r)/B(r)$ is increasing. Though b_i is negative, we still have $(A/B)'=\frac{A'B-AB'}{B^2}>0$. From the above estimates, we get

$$\begin{aligned}\lim_{\lambda\to 0}\lambda\frac{\partial_{x_i}\rho}{\rho}(\lambda x)&=\frac{x_i}{|x|^2}a_1\triangleq R_0(x),\\\left|\frac{\partial_{x_i}\rho}{\rho}(\lambda x)-R_0(\lambda x)\right|&\leq\lambda^{-1}\frac{x_i}{|x|^2}\frac{|A(\lambda_* x)|}{|B(\lambda_* x)|},\ \lambda\leq\lambda_*,\\\lim_{\lambda\to\infty}\lambda\frac{\partial_{x_i}\rho}{\rho}(\lambda x)&=\frac{x_i}{|x|^2}a_n\triangleq R_\infty(x),\\\left|\frac{\partial_{x_i}\rho}{\rho}(\lambda x)-R_\infty(\lambda x)\right|&\leq\lambda^{-1}\frac{x_i}{|x|^2}\frac{|A(\lambda_* x)|}{|B(\lambda_* x)|},\ \lambda\geq\lambda_*.\end{aligned}$$

A.2.3. Bounds for the derivatives of $1/\rho$. The bounds for $d_x^i d_y^j\rho^{-1}$ are more complicated since ρ^{-1} is not linear in the summand $p_i r^{a_i}$. We need such estimates in the estimate of the velocity. First, using the bounds in section A.2.1 and (A.5), we can obtain the upper and the lower bounds for R_{ij}:

$$R_{ij}=\frac{\partial_x^i\partial_y^j\rho}{\rho}.$$

For $i+j=1$ and $k=2,3$, we use the estimate in section A.2.1 to obtain the bounds for

$$R_{10}=\frac{x}{|x|^2}S(x),\quad R_{0,1}=\frac{y}{|x|^2}S(x),\quad (R_{ij})^k.$$

In our estimate, we need $\partial_x^i\partial_y^j\rho^{-1}$ for $i+j\le 3$. A direct calculation yields

$$\begin{aligned}
&\partial_x\rho^{-1}=-\frac{\rho_x}{\rho^2}=-\frac{R_{10}}{\rho},\quad \partial_{xx}\rho^{-1}=-\frac{\rho_{xx}}{\rho^2}+2\frac{\rho_x^2}{\rho^3}=\rho^{-1}(-R_{20}+2R_{10}^2),\\
&\partial_{xy}\rho^{-1}=-\frac{\rho_{xy}}{\rho}+\frac{2\rho_x\rho_y}{\rho^3}=\rho^{-1}(-R_{11}+2R_{10}R_{01}),\\
&\partial_{xxx}\rho^{-1}=-\frac{\rho_{xxx}}{\rho^2}+\frac{6\rho_{xx}\rho_x}{\rho^3}-\frac{6\rho_x^3}{\rho^4}=\rho^{-1}(-R_{30}+6R_{20}R_{10}-6R_{10}^3),\\
&\partial_{xxy}\rho^{-1}=-\frac{\rho_{xxy}}{\rho^2}+\frac{2\rho_{xx}\rho_y}{\rho^3}+\frac{4\rho_x\rho_{xy}}{\rho^3}-6\frac{\rho_x^2\rho_y}{\rho^4}\\
&\qquad=\rho^{-1}(-R_{21}+2R_{20}R_{01}+4R_{10}R_{11}-6R_{10}^2R_{01}).
\end{aligned}$$

Next, we estimate $\partial_x^i\partial_y^j(\partial_{x_l}\rho/\rho)$ for $i\le 2, j=0$ or $i=0, j\le 2$. Denote $f=\partial_{x_l}\rho$. Using a direct computation, for $D_2=\partial_x^{i_2}\partial_y^{j_2}$ with $i_2+j_2=1$, we get

$$D_2\frac{f}{\rho}=\frac{D_2f}{\rho}-\frac{fD_2\rho}{\rho^2}=\rho^{-1}(D_2f-fR_{i_2,j_2}).$$

For $(i_2,j_2)=(2,0),(0,2)$, denote $i_3=i_2/2$, $j_3=j_2/2$, $D_3=\partial_x^{i_3}\partial_y^{j_3}$. We get

$$\begin{aligned}
D_3^2\frac{f}{\rho}&=\frac{D_3^2f}{\rho}-\frac{2D_3f\cdot D_3\rho}{\rho^2}+fD_3^2\left(\frac{1}{\rho}\right)\\
&=\frac{D_3^2f}{\rho}-\frac{2D_3f\cdot D_3\rho}{\rho^2}+f\bigg(-\frac{D_3^2\rho}{\rho^2}+\frac{2(D_3\rho)^2}{\rho^3}\bigg)\\
&=\rho^{-1}(D_3^2f-2D_3fR_{i_3,j_3}-fR_{i_2,j_2}+2fR_{i_3,j_3}^2),
\end{aligned}$$

where we have used $D_3^2\frac{1}{\rho}=D_3(-\frac{D_3\rho}{\rho^2})=-\frac{D_3^2\rho}{\rho^2}+\frac{2(D_3\rho)^2}{\rho^3}$.

Since we have estimated $\partial_x^i\partial_y^j\rho$ and R_{ij}, we can bound these derivatives of $D_1\rho/\rho$ using (A.4).

We also need to obtain the uniform estimates of $\lambda^\beta\partial_x^i\partial_y^j(\rho^{-1}(\lambda x))$ for $\lambda\le\lambda_*, \beta=a_1$, and $\lambda\ge\lambda_*, \beta=a_n$. Denote $\rho_\lambda(x)=\rho(\lambda x)$. For example, for $D_1=\partial_{x_i}$, we have

$$\begin{aligned}
\lambda^\beta D_1(\rho_\lambda^{-1}(x))&=-\lambda^{1+\beta}\frac{(D_1\rho)(\lambda x)}{\rho_\lambda^2(x)}=-\lambda^{1+\beta}\rho_\lambda^{-1}(x)\lambda^{-1}\frac{x_i}{|x|^2}S(\lambda x)\\
&=-\lambda^\beta\rho_\lambda^{-1}(x)\frac{x_i}{|x|^2}S(\lambda x),
\end{aligned}$$

which can be estimated using the estimates in sections A.2.1, A.2.2. The power λ^β and the leading power $\lambda^{-\beta}$ in $\rho_\lambda^{-1}(x)$ cancel each other. The estimates of $\lambda^\beta\partial_x^i\partial_y^j(\rho^{-1}(\lambda x))$ with $i+j\ge 2$ and $\partial_x^i\partial_y^j\frac{\partial_{x_l}(\rho_\lambda)}{\rho_\lambda}$ are similar and follow from the above derivations for $\partial_x^i\partial_y^j\rho^{-1}, \partial_x^i\partial_y^j(\partial\rho/\rho)$ and the piecewise estimates for $\partial_x^i\partial_y^j\rho(\lambda x)$ in section A.2.1 and $\frac{\partial\rho}{\rho}(\lambda x)$ in section A.2.2, which are uniform in small $\lambda\le\lambda_*$ or large $\lambda\ge\lambda_*$. We remark that in all of these estimates for $\rho_\lambda(x)$, taking derivatives in x does not change the asymptotic power in λ.

A.2.4. Improved estimates for ρ^{-1} near $x = 0$. For the special case $a_1 = -2$, we can write

$$\rho(x) = r^{-2}\sum_i p_i r^{a_i+2} = r^{-2}\tilde{\rho}(x), \quad \rho^{-1} = (x^2+y^2)\tilde{\rho}(x)^{-1}.$$

To obtain a better estimate of ρ^{-1}, we use the fact that x^2+y^2 is a polynomial. First, we can obtain the bounds for $\partial_x^i\partial_y^j\tilde{\rho}^{-1}$. The bound for $S_0 = x^2+y^2$ is trivial, e.g.,

$$(\partial_x S_0)_\gamma = 2x_\gamma, (\partial_y S_0)_\gamma = 2y_\gamma,\ \gamma = u, l, \quad \partial_{xy}S_0 = 0, \quad \partial_{xx}S_0 = \partial_{yy}S_0 = 2.$$

Then using (A.4)–(A.5), we can bound ρ^{-1}.

A.3. The mixed weight. For the second type of weights $W = \rho_1(r)|x|^{-1/2} + \rho_2(r)$, we can compute its derivatives and its upper and lower bounds using linearity and the Leibniz rule (A.6). We consider $x, y \geq 0$. For example, we have

$$W_l = \rho_{1,l}x_u^{-1/2} + \rho_{2,l}, \quad (W^{-1})_u = (W_l)^{-1}, \quad W_x = \partial_x\rho_1 x^{-1/2} - \frac{1}{2}\rho_1 x^{-3/2} + \partial_x\rho_2.$$

To obtain the upper bound for $\partial_x^i\partial_y^j W$, we use the Leibniz rule (A.6):

$$|\partial_x^i\partial_y^j W| \leq \sum_{k\leq i}\binom{i}{k}|\partial_x^{i-k}\partial_y^j\rho_1|\frac{(2k-1)!!}{2^k}x^{-1/2-k} + |\partial_x^i\partial_y^j\rho_2|.$$

We need to bound $\rho(r)/W(x,y)$ in the estimate of the integrals. Suppose that the leading and the last powers of ρ are a_1, a_n. The leading and the last terms of W are given by $p_i r^{b_i}\cos(\beta)^{-\alpha_i}, \alpha_i \geq 0$,

$$W \geq p_1 r^{b_1}, \quad W \geq p_n r^{b_n}.$$

We estimate

$$\frac{\rho}{W} \leq C_1 r^{a_1-b_1}, \quad \frac{\rho}{W} \leq C_2 r^{a_n-b_n}$$

for all $x, y \in \mathbb{R}^+_2$. We apply the above estimates for x near 0 or x sufficiently large.

Using $W(\lambda x) \geq \rho_1(\lambda x)\lambda^{-1/2}|x_1|^{-1/2}$, $W(\lambda x) \geq \rho_2(\lambda x)$, and the uniform estimates of $\rho_i(\lambda x)$ in λ in section A.2.1, we can obtain the lower bound of $W(\lambda x)$ and the upper bound of $\frac{\rho(\lambda x)}{W(\lambda x)}$ uniformly in λ.

Appendix B. Estimate the derivatives of the velocity kernel and integrands. In this appendix, we estimate the derivatives of the kernel $-\frac{1}{2\pi}\log|x|$ associated to the velocity $\mathbf{u} = \nabla^\perp(-\Delta)^{-1}\omega$ and its symmetrization (4.25). These estimates are used to estimate the error terms in Lemmas 4.2, 4.4. We will perform an additional estimate for u with weight $\varphi(x)$ singular along $x_1 = 0$ in section B.4. Some additional derivations related to the estimate of the velocity are given in Appendix B.5.

B.1. Estimate the symmetrized kernel. In this section, we estimate the symmetrized kernel. We develop several symmetrized estimates for harmonic functions. Before we introduce the estimates, we have a simple 1D estimate, which is useful for later estimates.

LEMMA B.1. *We have*

$$|f(x)+f(-x)-2f(0)| \le x^2 ||f_{xx}||_{L^\infty[-x,x]},$$
$$|f(x)+f(-x)-2f(0)-x^2 f_{xx}(0)| \le \frac{x^4}{12} ||\partial_x^4 f||_{L^\infty[-x,x]}.$$

Proof. Denote $G(x) = f(x) + f(-x)$. Clearly, G is even and

$$G(0) = 2f(0), \quad G'(0) = 0, \quad \partial_x^2 G(0) = 2f_{xx}(0), \quad \partial_x^3 G(0) = 0. \tag{B.1}$$

Using the Taylor expansion, we obtain

$$G(x) = G(0) + G'(0)x + \frac{\partial_x^2 G(0) x^2}{2} + \frac{\partial_x^3 G(0) x^3}{6} + \frac{\partial_x^4 G(\xi) x^4}{24}$$

for some $\xi \in [0, x]$. Using (B.1), we get

$$|G(x) - G(0) - G''(x)\frac{x^2}{2}| \le ||\partial_x^4 G||_{L^\infty[0,x]} \frac{x^4}{24} \le ||\partial_x^2 f||_{L^\infty[-x,x]} \frac{x^4}{12}.$$

Plugging the identity (B.1) into the above estimate proves the second estimate in Lemma B.1. The first estimate is simpler. ∎

The following lemma is useful for estimating the symmetrized kernel (4.25) and its derivatives.

LEMMA B.2. *Suppose that* $Q_x = [-x_1, x_1] \times [-x_2, x_2]$ *and* $f \in C^4(Q_x)$ *is harmonic. Denote*

$$\begin{aligned} G_1(1,x) &\triangleq f(x_1,x_2) + f(-x_1,x_2) + f(x_1,-x_2) + f(-x_1,-x_2) - 4f(0,0), \\ G_2(1,x) &\triangleq f(x_1,x_2) - f(-x_1,x_2) - f(x_1,-x_2) + f(-x_1,-x_2), \\ \hat{G}_1(x) &\triangleq 2x_1^2 f_{xx}(0,0) + 2x_2^2 f_{yy}(0,0), \quad \hat{G}_2(x) \triangleq 4x_1 x_2 f_{xy}(0,0). \end{aligned} \tag{B.2}$$

We have

$$|G_1(1,x)| \le 2|x|^2 ||f_{xx}||_{L^\infty(Q_x)}, \quad |\partial_{x_i} G_1(1,x)| \le 4|x_i| \cdot ||f_{xx}||_{L^\infty(Q_x)}, \tag{B.3}$$

$$|G_1(1,x) - \hat{G}_1(x)| \le \frac{(x_1^4 + 6x_1^2 x_2^2 + x_2^4)}{6} ||\partial^4 f||_{L^\infty(Q_x)} \le \frac{|x|^4}{3} ||\partial^4 f||_{L^\infty(Q_x)}, \tag{B.4}$$

$$|G_1(1,x_1,0) - \hat{G}_1(x_1,0)| \le \frac{1}{6} x_1^4 ||\partial^4 f||_{L^\infty(Q_x)}, \tag{B.5}$$

$$|\partial_{x_i}(G_1(1,x) - \hat{G}_1(x))| \le \frac{2}{3}(3x_{3-i}^2 x_i + x_i^3) ||\partial^4 f||_{L^\infty(Q_x)} \le \frac{2\sqrt{2}}{3} |x|^3 ||\partial^4 f||_{L^\infty(Q_x)}, \tag{B.6}$$

where $||\partial^4 f||_{L^\infty} = \max_{0 \le i \le 4} ||\partial_x^i \partial_y^j f||_{L^\infty(Q_x)}$. *For* G_2, *we have the following estimate:*

$$|G_2(1,x)| \le 4x_1 x_2 ||f_{xy}||_{L^\infty(Q_x)}, \quad |\partial_{x_i} G_2(1,x)| \le 4|x_{3-i}| \cdot ||f_{xy}||_{L^\infty(Q_x)}, \tag{B.7}$$

(B.8)
$$|G_2(1,x)-\hat{G}_2(x)|\le\frac{2x_1x_2|x|^2}{3}||\partial^4 f||_{L^\infty(Q_x)},$$

(B.9)
$$|\partial_{x_i}(G_2(1,x)-\hat{G}_2(x))|\le\frac{2}{3}(3x_i^2x_{3-i}+x_{3-i}^3)||\partial^4 f||_{L^\infty(Q_x)}\le\frac{2\sqrt{2}}{3}|x|^3||\partial^4 f||_{L^\infty(Q_x)}.$$

Note that $G_1(\cdot,x)$ is even in x_i, and $G_2(\cdot,x)$ is odd in x_i. The polynomials of x_i in the upper bounds (without absolute value) have the same symmetries. Similar properties hold for $\partial G_1,\partial G_2$. Moreover the above bound satisfies the differential relation. These properties are useful for tracking different bounds for G_1,G_2.

Proof. Recall $Q_x=[-x_1,x_1]\times[-x_2,x_2]$. Denote
$$A_{ij}(x)=||\partial_x^i\partial_y^j f||_{L^\infty(Q_x)}.$$
Using Lemma B.1, for any $t\in[0,1]$, we obtain
$$\begin{aligned}|f(tx_1,x_2)+f(tx_1,-x_2)-2f(tx_1,0)|&\le A_{02}x_2^2,\\ |f(x_1,0)+f(-x_1,0)-2f(0,0)|&\le A_{20}x_1^2.\end{aligned}$$
Since f is a harmonic function, we have $\partial_x^{i+2}\partial_y^j f=-\partial_x^i\partial_y^{j+2}f$ and obtain $A_{i+2,j}=A_{i,j+2}$. Taking $t=\pm1$ in the above estimate and using the triangle inequality, we prove
$$|G(1,x)|\le 2A_{20}x_1^2+2A_{02}x_2^2=2A_{20}(x_1^2+x_2^2)=2A_{20}|x|^2,$$
which is the first estimate in (B.3).

The second estimate in (B.3) is simple. We consider $i=1$ without loss of generality. We get
$$\begin{aligned}|\partial_{x_1}G_1(1,x)|&=|(\partial_1 f)(x_1,x_2)-(\partial_1 f)(-x_1,x_2)+(\partial_1 f)(x_1,-x_2)-(\partial_1 f)(-x_1,-x_2))|\\ &\le 4x_1A_{20}(x).\end{aligned}$$

For (B.4), using Lemma B.1, we get
$$\begin{aligned}|f(tx_1,x_2)+f(tx_1,-x_2)-2f(tx_1,0)-x_2^2(\partial_2^2 f)(tx_1,0)|&\le A_{04}(x)\frac{x_2^4}{12},\\ |\partial_2^2 f(x_1,0)+\partial_2^2 f(-x_1,0)-2\partial_2^2 f(0,0)|&\le x_1^2A_{2,2}(x),\\ |f(x_1,0)+f(-x_1,0)-2f(0)-x_1^2\partial_1^2 f(0)|&\le A_{40}\frac{x_1^4}{12}\end{aligned}\tag{B.10}$$
for $t=\pm1$. Combining the above estimates and using the triangle inequality and $A_{40}=A_{22}=A_{04}$, we prove the first estimate in (B.4). The second estimate follows from $2|x|^4-x_1^4-6x_1^2x_2^2-x_2^4=(x_1^2-x_2^2)^2\ge0$.

Estimate (B.5) follows from (B.4) by taking $x_2=0$.

For (B.6), we consider the estimate of ∂_{x_1}. The other case is similar. Using
$$\partial_1 f(x_1,s)-(\partial_1 f)(-x_1,s)=\int_0^{x_1}(\partial_1^2 f)(t,s)+(\partial_1^2 f)(-t,s)dt,$$

we obtain

$$\begin{aligned}\partial_1(G(1,x)-\hat{G}_1(x)) &= (\partial_1 f)(x_1,x_2)-(\partial_1 f)(-x_1,x_2)+(\partial_1 f)(x_1,-x_2)\\ &\quad -(\partial_1 f)(-x_1,-x_2)-4x_1\partial_1^2 f(0)\\ &= \int_0^{x_1}\Big((\partial_1^2 f)(z,x_2)+(\partial_1^2 f)(-z,x_2)+(\partial_1^2 f)(z,-x_2)\\ &\quad +(\partial_1^2 f)(-z,x_2)-4\partial_1^2 f(0)\Big)dz.\end{aligned}$$

Applying (B.3), we get

$$|\partial_1(G(1,x)-\hat{G}_1(x))| \le \int_0^{x_1} 2(z^2+x_2^2)dz A_{4,0}(x) = \left(\frac{2}{3}x_1^3+2x_1x_2^2\right)A_{4,0}(x)$$

and complete the proof of the first estimate in (B.6). For the second estimate, we use the inequality of arithmetic and geometric means (AM-GM) inequality to yield

(B.11)
$$(3x_2^2x_1+x_1^3)^2 = (3x_2^2+x_1^2)^2x_1^2 = \frac{1}{4}(3x_2^2+x_1^2)^2 4x_1^2 \le \frac{1}{4}\left(\frac{2(3x_2^2+x_1^2)+4x_1^2}{3}\right)^3 = 2|x|^6.$$

Taking a square root completes the estimate.

To estimate G_2 in (B.2), we rewrite it as follows:

(B.12)
$$\begin{aligned}G_2(1,x)-c\hat{G}_2(x) &= \int_{-x_1}^{x_1}\int_{-x_2}^{x_2}\partial_{12}f(z_1,z_2)-c\partial_{12}f(0)dz\\ &= \int_0^{x_1}\int_0^{x_2}(\partial_{12}f)(z_1,z_2)+(\partial_{12}f)(-z_1,z_2)+(\partial_{12}f)(z_1,-z_2)\\ &\quad +(\partial_{12}f)(-z_1,z_2)-4c(\partial_{12}f)(0)dz\end{aligned}$$

for $c=0,1$. The integrand has the same form as G_1 in (B.2). For $c=0$, using the above decomposition, we prove

$$|G_2(1,x)| \le 4x_1x_2A_{11}.$$

When $c=1$, using (B.6), we get

$$|G_2(1,x)-\hat{G}_2(x)| \le A_{40}2\int_0^{x_1}\int_0^{x_2}|y|^2dy = A_{40}\frac{2}{3}(x_1^3x_2+x_1x_2^3) = A_{40}\frac{2}{3}x_1x_2|x|^2.$$

To estimate the derivatives, we focus on ∂_{x_1}. Using the above representation, we obtain

$$\begin{aligned}\partial_{x_1}(G_2(1,x)-c\hat{G}_2(x)) &= \int_0^{x_2}((\partial_{12}f)(x_1,y_2)+(\partial_{12}f)(-x_1,y_2))\\ &\quad +((\partial_{12}f)(x_1,-y_2)+(\partial_{12}f)(-x_1,-y_2)-4c(\partial_{12}f)(0))dy.\end{aligned}$$

We apply the same estimates to the integrands with $c=0,1$ and get

$$\begin{aligned}&|\partial_{x_1}G_2(1,x)| \le 4x_2A_{11},\\ &|\partial_{x_1}(G_2(1,x)-\hat{G}_2(x))| \le A_{31}2\int_0^{x_2}(x_1^2+y_2^2)dy_2 = A_{31}\left(2x_1^2x_2+\frac{2}{3}x_2^3\right).\end{aligned}$$

The second inequality in (B.9) follows from (B.11). The above estimates imply (B.7)–(B.9). ∎

Recall the kernels associated with $\nabla\mathbf{u}, \mathbf{u}$ in (4.1). These kernels are the derivatives of the Green function $-\frac{1}{2\pi}\log|x|$ and are harmonic away from 0. We have the following estimates for their derivatives.

LEMMA B.3. *Denote* $r = (x^2+y^2)^{\frac{1}{2}}$ *and* $f(x,y) = \log r$. *For any* $i, j \geq 0$ *with* $i+j \geq 1$, *we have*

$$|\partial_x^i \partial_y^j f(x,y)| \leq (i+j-1)! \cdot r^{-i-j}.$$

As a result, for $K_1(y) = -\frac{1}{2}\partial_{12} f(y), K_2(y) = -\frac{1}{2}\partial_1^2 f(y)$, *we have*

$$|K_i| \leq \frac{1}{2|y|^2}, \quad |\partial_{y_1}^j \partial_{y_2}^{2-j} K_i| \leq \frac{3}{|y|^4}, \quad |\partial_{y_1}^j \partial_{y_2}^{4-j} K_i| \leq \frac{60}{|y|^6},$$
$$|\partial_{y_1}^j \partial_{y_2}^{6-j} K_i| \leq \frac{2520}{|y|^8}, \quad i = 1, 2.$$

Proof. Consider the polar coordinate $\beta = \arctan(y/x), r = (x^2+y^2)^{1/2}$. We use induction on $n = i+j$ to prove

$$\partial_x^i \partial_y^j f = (n-1)! \cos(n\beta - \beta_{ij}) r^{-n} \tag{B.13}$$

for some constant β_{ij}. We have the formula

$$\partial_x g = \left(\cos\beta \partial_r - \frac{\sin\beta}{r}\partial_\beta\right) g, \quad \partial_y g = \left(\sin\beta\partial_r + \frac{\cos\beta}{r}\partial_\beta\right) g. \tag{B.14}$$

First, for $n = 1$, a direct calculation yields

$$\partial_x f = \frac{x}{r^2} = \frac{\cos\beta}{r}, \quad \partial_y f = \frac{y}{r^2} = \frac{\sin\beta}{r} = \frac{\cos(\beta - \pi/2)}{r}.$$

Suppose that (B.13) holds for any i, j with $i+j = n$ and $n \geq 1$. Now, since

$$\begin{aligned}\partial_x \partial_x^i \partial_y^j f &= (n-1)! \partial_x (\cos(n\beta - \beta_{ij}) r^{-n}) \\ &= (n-1)! (-n\cos\beta\cos(n\beta - \beta_{ij}) r^{-n-1} + n\sin\beta\sin(n\beta - \beta_{ij}) r^{-n-1}) \\ &= n! (-\cos(n\beta - \beta_{ij} + \beta) r^{-n-1}) = n! \cos((n+1)\beta - \beta_{ij} - \pi) r^{-n-1},\end{aligned}$$

using a similar computation and $\sin(x) = \cos(x - \pi/2)$, we can obtain that $\partial_y \partial_x^i \partial_y^j f$ has the form (B.13). Using induction, we prove (B.13). The desired estimate follows from (B.13). ∎

Using the above two lemmas, we can estimate the error in the discretization of the kernels $K(x,y)$ in both the x and y directions.

B.1.1. Estimate the kernels in the far-field. We apply Lemma B.2 to estimate the decay of F_1, F_2

$$\begin{aligned}&F_0 \triangleq G(y-x) - G(y_1 - x_1, y_2 + x_2) - G(y_1 + x_1, y_2 - x_2) + G(y+x), \\ &G(y) = -\log|y|/2, \\ &F_1 \triangleq F_0 - 4x_1 x_2 \partial_{12} G(y), \ F_2 \triangleq F_1 - \frac{2(x_1^2 - x_2^2)x_1 x_2}{3}\partial_1^3 \partial_2 G(y), \\ &I_{ijkl}(P) \triangleq \partial_{x_1}^i \partial_{x_2}^j \partial_{y_1}^k \partial_{y_2}^l P(x,y).\end{aligned} \tag{B.15}$$

Note that for stream function $\phi = (-\Delta)^{-1}\omega(y) = C \cdot G * W$, where W is the odd extension of ω from $\mathbb{R}^+_2$ to $\mathbb{R}^{++}_2$, since $G(z)$ is even in z_i, after symmetrization, we have

$$\tilde{\phi}(x) = \phi(x) - x_1x_2\phi_{12}(0) = C\int_{\mathbb{R}^2} G(y-x)W(y)dy = C\int F_1(x,y)W(y)dy,$$

where $\phi_{12}(0)$ is related to $C_{f0}K_{ux0}$ in (4.5). In the estimate of $\mathbf{u}, \nabla\mathbf{u}$ related to $\partial^i_{x_1}\partial^j_{x_2}\tilde{\phi}$, e.g., $(1,1)$ for $u_x = -\partial_{x_1x_2}\phi$, for $y \in Q$ away from the singularity, we get the symmetrized integrand

$$\partial^i_{x_1}\partial^j_{x_2}\int_Q F_1(x,y)W(y)dy = \int_Q \partial^i_{x_1}\partial^j_{x_2}F_1(x,y)W(y)dy.$$

In the error estimate of the trapezoidal rule Lemma 4.2, we estimate $\partial^i_{x_1}\partial^j_{x_2}\partial^2_{y_i}$ $F_1(x,y)$, which is $I_{ij20}(F_1)$ or $I_{ij02}(F_1)$ in (B.15). We apply the estimate of F_2 to $K_f - C_{f0}K_{ux0} - C_fK_{00}$ (4.5). Below, we show that $I_{ijkl}(F_i), i = 1,2$ has faster decay in $|y|$ than $\partial^i_{x_1}\partial^j_{x_2}\partial^k_{y_1}\partial^l_{y_2}G(y+x)$.

By definition, we get $i_1, j_1 \le 1$. Next, we fix y and introduce

$$\begin{aligned} &g_{pq}(z) \triangleq \partial^p_{y_1}\partial^q_{y_2}G(y+z), \quad M_{G,k} \triangleq \max_{a+b=k} ||(\partial^a_{y_1}\partial^b_{y_2}G)(y+\cdot)||_{L^\infty(Q_x)}, \\ &Q_x = [-x_1, x_1] \times [-x_2, x_2]. \end{aligned} \tag{B.16}$$

Since G is harmonic, we have

(B.17)

$$\begin{aligned} &\partial^k_{x_i}G(y_1 + s_1x_1, y_2 + s_2x_2) = s_i^k\partial^k_{y_i}G(x_1 + s_1y_1, x_2 + s_2y_2),\ s_l \in \{\pm 1\}, \\ &\partial^2_1 G(y) = -\partial^2_2 G(y), \quad \partial_{x_1x_2}g_{pq}(x)|_{x=0} = \partial^{p+1}_{y_1}\partial^{q+1}_{y_2}G(y), \quad \partial_{22}g_{rs}(0) = -\partial_{11}g_{rs}(0). \end{aligned}$$

Second approximation F_2. Note that taking ∂_{y_i} in F_i does not change the sign of the coefficient of the G term in (B.15). Applying (B.12) with $c = 1$ and $f(z) = g_{rs}(z)$ in G_2, we get

$$\begin{aligned} I_{pqrs}(F_2) &= \partial^p_{x_1}\partial^q_{x_2}\int_0^{x_1}\int_0^{x_2} g_{r+1,s+1,all}(z)dz, \\ g_{\alpha\beta,all}(z) &= g_{\alpha\beta}(z) + g_{\alpha\beta}(-z) + g_{\alpha\beta}(z_1,-z_2) + g_{\alpha\beta}(-z_1,z_2) \\ &\quad - 4g_{\alpha\beta}(0) - 2(z_1^2 - z_2^2)\partial_{11}g_{\alpha\beta}(0). \end{aligned}$$

If $\max(i,j) \le 1$, using the above notation to $I_{ijkl}(F_2)$ and the estimate of $G_1 - \hat{G}_1$ in Lemma B.2 with $f = g_{kl}$, and then integrating the bounds in z_2, we get

$$\begin{aligned} |I_{10kl}(F_2)| &= \left|\int_0^{x_2} g_{k+1,l+1,all}(x_1,z_2)dz_2\right| \le M_{G,d_2}\int_0^{x_2}\frac{x_1^4 + 6x_1^2z_2^2 + z_2^4}{6}dz_2 \\ &= \left(\frac{x_1^4x_2}{6} + \frac{x_1^2x_2^3}{3} + \frac{x_2^5}{30}\right)M_{G,d_2}, \end{aligned}$$

where $d_2 = k + l + 6$. Similarly, we get

$$|I_{01kl}(F_2)| \le \left(\frac{x_1^5}{30} + \frac{x_1^3x_2^2}{3} + \frac{x_1^4x_2}{6}\right)A_{G,d_2}, \quad I_{11kl}(F_2) \le \frac{x_1^4 + 6x_1^2x_2^2 + x_2^4}{6}A_{G,d_2}.$$

If $\max(i,j)\geq 2, i+j\leq 3$, without loss of generality, we consider $i\geq 2$. We choose $(i_1,j_1,k_1,l_1)=(i-2,j,k+2,l)$. From (B.17), we get

$$\partial_{x_1}^2(x_1x_2\partial_{12}G(y))=0,$$
$$\partial_{x_1}^2\partial_{y_1}^k\partial_{y_2}^l\left(\frac{2(x_1^2-x_2^2)x_1x_2}{3}\partial_1^3\partial_2 G(y)\right)=4x_1x_2\partial_{y_1}^{k_1+1}\partial_{y_2}^{l_1+1}G)(y)=4x_1x_2\partial_{12}g_{k_1l_1}(0).$$

Using (B.17) again, we rewrite $\partial_{x_1}^i\partial_{y_1}^k G(x+y)=\partial_{x_1}^{i_1}\partial_{y_1}^{k_1}G(x+y)$ and get

$$\begin{aligned}I_{ijkl}(F_2)=\partial_{x_1}^{i_1}\partial_{x_2}^{j_1}(&g_{k_1l_1}(x)-g_{k_1l_1}(x_1,-x_2)-g_{k_1l_1}(-x_1,x_2)\\&+g_{k_1l_1}(-x)-4x_1x_2\partial_{12}g_{k_1l_1}(0)).\end{aligned}\tag{B.18}$$

The same derivation applies to the case of $j\geq 2$, where we choose $(i_1,j_1,k_1,l_1)=(i,j-2,k,l+2)$. Since $i_1,j_1\leq 1$, using the estimate of $G_2-\hat{G}_2$ in Lemma B.2 with $f=g_{k_1l_1}$, we get

$$\begin{aligned}|I_{20kl}(F_2)|,|I_{02kl}(F_2)|&\leq\frac{2x_1x_2|x|^2}{3}M_{G,d_2},\ (i_1,j_1)=(0,0),\\|I_{30kl}(F_2)|,|I_{12kl}(F_2)|&\leq\frac{2}{3}(3x_1^2x_2+x_2^3)M_{G,d_2},(i_1,j_1)=(1,0),\\|I_{21kl}(F_2)|,|I_{03kl}(F_2)|&\leq\frac{2}{3}(x_1^3+3x_1x_2^2)M_{G,d_2},(i_1,j_1)=(0,1),\\d_2&=k_1+l_1+4=k+l+6.\end{aligned}$$

Note that the form (B.18) can be seen as the $\partial_{x_1}^{i_1}\partial_{x_2}^{i_2}F_1$. If $4\leq i+j\leq 5$, we still first perform (B.18) by choosing $(i_1,j_1,k_1,l_1)=(i-2,j,k+2,l)$ or $(i,j-2,k,l+2)$ and get

$$I_{ijkl}(F_2)=I_{i_1j_1k_1l_1}(\tilde{F}_1),$$

where $\tilde{F}_1$ is similar to F_1 in (B.15) with G replaced by $g_{i-i_1,j-j_1}=\partial_{y_1}^{i-i_1}\partial_{y_2}^{j-j_1}G(y)$. Then we apply the estimate for the first approximation below with $i_1+j_1\leq 3$.

First approximation. The estimate of $I_{ijkl}(F_1)$ is similar. Denote

$$i_2=i-2\left\lfloor\frac{i}{2}\right\rfloor,\ j_2=j-2\left\lfloor\frac{j}{2}\right\rfloor,\ k_2=k+2\left\lfloor\frac{i}{2}\right\rfloor,\quad l_2=l+2\left\lfloor\frac{j}{2}\right\rfloor.$$

If $\max(i,j)\leq 1$, we get $(i,j,k,l)=(i_2,j_2,k_2,l_2)$. Applying the estimate $G_2-\hat{G}_2$ in Lemma (B.2) with $f=g_{k_2l_2}$, we get

$$\begin{aligned}I_{10kl}(F_1)&\leq\frac{2}{3}x_2(x_2^2+3x_1^2)||\partial^d G(y+\cdot)||_{L^\infty(Q_x)}=\frac{2}{3}x_2(x_2^2+3x_1^2)M_{G,d},\\I_{01kl}(F_1)&\leq\frac{2}{3}x_1(x_1^2+3x_2^2)M_{G,d},\quad I_{00kl}(F_1)\leq\frac{2x_1x_2|x|^2}{3}M_{G,d},\\d&=k_2+l_2+4=k+l+4.\end{aligned}$$

If $(i,j)=(1,1)$, we apply the estimate of G_1 in Lemma (B.2) with $f=\partial_{x_1x_2}g_{kl}(x)$ (k,l are number of derivatives on $G(y+z)$) to get

$$|I_{11kl}(F_1)|\leq 2|x|^2M_{G,d},\quad d=k_2+l_2+4=k+l+4.$$

If $\max(i,j) \geq 2, i+j \leq 3$, $x_1x_2\partial_{12}G(0)$ vanishes in I_{ijkl}. We apply a derivation similar to (B.18) without $4x_1x_2\partial_{12}g_{k_2l_2}(0)$ and the estimate of G_2 in Lemma B.2 with $f = g_{k_2l_2}$ to get

$$|I_{ijkl}(F_1) \leq 4x_1^{1-i_2}x_2^{1-j_2}||\partial^2 g_{k_2l_2}||_{L^\infty(Q_x)} \leq 4x_1^{1-i_2}x_2^{1-j_2}M_{G,d}, \quad d = k_2+l_2+2 = k+l+4.$$

To bound $M_{G,k}$, we apply Lemma B.3 to get

(B.19)
$$M_{G,k} = \max_{a+b=k} ||(\partial_{y_1}^a\partial_{y_2}^b G)(y+\cdot)||_{L^\infty(Q_x)} \leq \frac{(k-1)!}{2\cdot \mathrm{Den}(x,y)^{k/2}}, \quad \mathrm{Den}(x,y) = \min_{z\in Q_x}|y-z|^2.$$

It is not difficult to obtain that for $x, y \in \mathbb{R}^{++}_2$, we have

$$\mathrm{Den}(x,y) = \sum_{i=1,2}\min_{|z_i|\leq x_i}|y_i - z_i|^2 = \sum_{i=1,2}(\max(y_i - x_i, 0))^2. \tag{B.20}$$

Using the above estimates, for $|y| \gg |x|$, we get $\mathrm{Den} \sim |y|^2$ and the decay estimate for $I_{ijkl}(F_1)$ (B.15) with a rate $|y|^{-k-l-4}$ and $I_{ijkl}(F_2)$ with a rate $|y|^{-k-l-6}$.

B.2. Piecewise L^∞ estimate of derivatives of the Green function. In this section, we develop sharp L^∞ estimates of the derivatives of the Green function $G(x) = -\frac{1}{2\pi}\log|x|$ and their linear combinations in a small domain $[a,b]\times[c,d]$. They will be used in Lemmas 4.2, 4.4 to estimate the error, especially near the singularity of the kernel. We remark that the linear combinations of $\partial_1^i\partial_2^j G$ can be quite complicated. If we simply use the triangle inequality to estimate it, we can overestimate some terms with cancellation significantly, especially near the singularity of G. These sharp estimates are useful for reducing the estimate of the error term in Lemmas 4.2, 4.4 without choosing a very small mesh, which can lead to large computational cost.

B.2.1. Coefficients of the derivatives of the Green function. To simplify the notation, we drop $\frac{1}{\pi}$ from G and denote $f_p = -\frac{1}{2}\log|x|$. First, we derive the formulas of $\partial_1^i\partial_2^j f_p$. Due to homogeneity, for $k+l \geq 1$, we assume

$$\partial_{x_1}^k\partial_{x_2}^l f_p = \frac{\sum_{i+j=k+l} c_{ij}x_1^i x_2^j}{|x|^{2(k+l)}}. \tag{B.21}$$

Next, we derive the recursive formula for c_{ij}. Using induction, we can obtain

$$\begin{aligned}\partial_{x_1}^{k+1}\partial_{x_2}^l f_p &= \frac{\sum_{i+j=k+l} c_{ij}ix_1^{i-1}x_2^j}{|x|^{2(k+l)}} - \frac{2(k+l)x_1}{|x|^{2(k+l+1)}}\sum_{i+j=k+l} c_{ij}x_1^i x_2^j \\ &= \frac{1}{|x|^{2(k+l+1)}}\left(\sum_{i+j=k+l} c_{ij}ix_1^{i+1}x_2^j + c_{ij}ix_1^{i-1}x_2^{j+2} - 2(k+l)c_{ij}x_1^{i+1}x_2^j\right) \\ &= \frac{1}{|x|^{2(k+l+1)}}\left(\sum_{i+j=k+l}(c_{ij}i + c_{i+2,j-2}(i+2) - 2(k+l)c_{ij})x_1^{i+1}x_2^j\right).\end{aligned}$$

Therefore, we obtain the recursive formula

$$c_{i+1,j} = ic_{ij} + (i+2)c_{i+2,j-2} - 2(k+l)c_{ij}$$

for all $i+j=k+l$, or equivalently,

$$c_{i,j}=(i-1)c_{i-1,j}-2(k+l)c_{i-1,j}+(i+1)c_{i+1,j-2}$$

for all $i+j=k+l+1$. Similarly, for ∂_{x_2}, we get

$$c_{i,j}=(j-1)c_{i,j-1}-2(k+l)c_{i,j-1}+(j+1)c_{i-2,j+1}$$

for all $i+j=k+l+1$.

B.2.2. Estimates of rational functions. We use the above formulas to develop sharp estimates of the derivatives of f_p and their linear combinations in a small grid cell $[y_{1l},y_{1u}]\times[y_{2l},y_{2u}]$. For $k<k_2$ and $S\subset\{(i,j):i+j=k\}$, we estimate

$$I_S\triangleq\frac{\sum_{(i,j)\in S}c_{ij}y_1^iy_2^j}{|y|^{k_2}}. \tag{B.22}$$

We assume that $I_S(x)$ is either odd in x_i or even in x_i for $i=1,2$. Clearly, this property holds for $\partial_{x_1}^k\partial_{x_2}^l f_p$ (B.21). Denote $i_1=\min_{i\in S}i, j_1=\min_{j\in S}j$. We get

$$I_S=\frac{y_1^{i_1}y_2^{j_1}}{|y|^{i_1+j_1}}\frac{\sum_{(i,j)\in S}c_{ij}y_1^{i-i_1}y_2^{j-j_1}}{|y|^{k_2-i_1-j_1}}.$$

We further introduce

$$P\triangleq\sum_{(i,j)\in S}c_{ij}^+y_1^{i-i_1}y_2^{j-j_1},\quad Q\triangleq\sum_{(i,j)\in S}c_{ij}^-y_1^{i-i_1}y_2^{j-j_1}.$$

We claim that $i-i_1,j-j_1$ are even for all $(i,j)\in S$. Since I_S is either odd or even in $x_i,i=1,2$, the numerator $\sum c_{ij}x_1^ix_2^j$ in (B.22) has the same symmetries in x_1,x_2. In particular, each monomial $c_{ij}x_1^ix_2^j$ in (B.22) also enjoys the same symmetries in x_1,x_2 as I_S. If $i-i_1$ is odd for some i, then $c_{ij}x_1^{i-i_1}x_2^{j-j_1}$ must be odd in x_1. It implies $i-i_1\ge1$ for any $(i,j)\in S$ and contradicts the minimality of i_1. The same argument applies to j_1.

As a result, P and Q are monotone increasing in $|y_1|,|y_2|\ge0$. For $|y_i|_l\le|y_i|\le|y_i|_u, i=1,2$, we can derive the upper and lower bounds for P,Q and get

$$\begin{aligned}|I|&\le\frac{\max(P_u-Q_l,Q_u-P_l)}{|y|_l^{k_2-i_1-j_1}}\max_{y\in\Omega}\frac{|y_1|^{i_1}|y_2|^{j_1}}{|y|^{i_1+j_1}}\\&\le\frac{\max(P_u-Q_l,Q_u-P_l)}{|y|_l^{k_2-i_1-j_1}}\left(\frac{|y_1|_u}{(|y_1|_u^2+|y_2|_l^2)^{1/2}}\right)^{i_1}\left(\frac{|y_2|_u}{(|y_1|_l^2+|y_2|_u^2)^{1/2}}\right)^{j_1},\end{aligned}$$

where $|y|_l$ is the lower bound of $|y|$ and we have used the fact that $z_i/|z|$ is increasing in z_i for $z_i\ge0$ to obtain its upper bound. Now, for $y_i\in[y_{il},y_{iu}]$, we estimate $|y_i|_l,|y_i|_u$ as follows:

$$\begin{aligned}&|y_i|\ge\max(0,|y_{il}+y_{iu}|/2-(y_{iu}-y_{il})/2)\triangleq|y_i|_l,\\&|y_i|\le\max(|y_{il}|,|y_{iu}|)\triangleq|y_i|_u,\ |y|_l\triangleq(|y_1|_l^2+|y_2|_l^2)^{1/2}.\end{aligned} \tag{B.23}$$

Note that for $y_i\in[y_{il},y_{iu}]$, y_i can change signs.

B.3. Improved estimate of the higher order derivatives of the integrands. In the Hölder estimate, we need to estimate the derivatives of the integrands (4.28), (4.29), (4.24), which take the form

$$K^C(x,y)(p(x)-p(y))+K^{NC}p(x)$$

for some weight p and kernels K^C, K^{NC}. Using the estimates of the kernels in Appendices B.1, B.2 and the weights in section A.1, the Leibniz rule (A.6), and the triangle inequality, we can estimate the derivative of the integrands. However, such an estimate can lead to significant overestimates near the singularity of the integrand. We use the estimates in Appendix B.2 to handle the cancellations among different terms and obtain improved estimates for the integrand and its derivatives near the singularity:

$$T_{00}(x,y)\triangleq K(y-x)(p(x)-p(y)),\quad \partial_{x_i}T_{00}(x,y). \tag{B.24}$$

We choose weight $p(x)$ that is even in x and y. The basic idea is to perform a Taylor expansion on $p(x)-p(y)$ and obtain the factor $|x-y|$, which cancels one order of singularity from $K(x,y)$. We use the formulas in Appendix B.2 to collect the terms with the same singularity and exploit the cancellation.

B.3.1. Y-discretization. In the Y-discretization of the integral, we need to estimate the y-derivatives of the integrand (B.24). For $a,b=1,2$, denote

$$D_1=\partial_a,\quad D_2=\partial_b,\quad x_m=\frac{x+y}{2}. \tag{B.25}$$

Next, we compute $\partial^j_{y_b}\partial^i_{x_a}T_{00}$. The reader should be careful about the sign. Note that

$$\partial_{x_a}(K(y-x))=-(\partial_a K)(y-x)=-D_1K(y-x).$$

Using the Leibniz rule, we get

$$\begin{aligned}\partial^2_{y_b}\partial_{x_a}T_{00}&=\partial^2_{y_b}(-D_1K(p(x)-p(y))+K\cdot D_1p(x))\\&=\partial^2_{y_b}(D_1K\cdot(p(y)-p(x))+K\cdot D_1p(x))\\&=D_2^2D_1K\cdot(p(y)-p(x))+2D_2D_1K\cdot D_2p(y)\\&\quad+D_1K\cdot D_2^2p(y)+D_2^2K\cdot D_1p(x).\end{aligned}$$

We use Taylor expansion at $x=x_m$ and write

$$\begin{aligned}&p(y)-p(x)=(y-x)\cdot\nabla p(x_m)+p_{m,2,err},\ \partial_i p(z)\\&\qquad=\partial_i p(x_m)+(\partial_i p(z)-\partial_i p(x_m)),\ z=x,y,\\&|f(z)-f(x_m)-(z-x_m)\cdot\nabla f(x_m)|\\&\qquad\le\frac{1}{2}\frac{d_1^2}{4}||f_{xx}||_{L^\infty(Q)}+\frac{d_1d_2}{4}||f_{xy}||_{L^\infty(Q)}+\frac{1}{2}\frac{d_1^2}{4}||f_{xx}||_{L^\infty(Q)}\triangleq I_f\end{aligned} \tag{B.26}$$

for $d = y - x, z = x, y$ and any f, where Q is the rectangle covering x, y. Then $p_{m,2,err}$ is bounded by $2I_p = O(|x-y|^2)$. Combining the terms involving ∇p, we get

(B.27)
$$\begin{aligned}\partial^2_{y_b}\partial_{x_a}T_{00} &= \sum_{i=1,2}\Big(D_2^2D_1K\cdot(y_i-x_i) + \mathbf{1}_{D_2=\partial_i}2D_2D_1K + \mathbf{1}_{D_1=\partial_i}D_2^2K\Big)\cdot\partial_{x_i}p(x_m)\\ &\quad + D_2^2D_1K\cdot p_{m,2,err} + 2D_2D_1K\cdot(D_2p(y)-D_2p(x_m))\\ &\quad + D_2^2K\cdot(D_1p(x)-D_1p(x_m)) + D_1K\cdot D_2^2p(y)\\ &\triangleq \left(\sum_{i=1,2} I_i\cdot\partial_{x_i}p(x_m)\right) + II_1 + II_2 + II_3 + II_4,\\ I_i &\triangleq D_2^2D_1K\cdot(y_i-x_i) + \mathbf{1}_{D_2=\partial_i}2D_2D_1K + \mathbf{1}_{D_1=\partial_i}D_2^2K,\end{aligned}$$

where $\partial_1^i\partial_2^jK$ is evaluated at $y-x$, and II_i denotes the last four terms in the second equation. The first term is the most singular term. We combine the most singular terms to exploit the cancellation and improve the estimates. We estimate the kernels

(B.28)
$$K_{mix}(D_1,D_2,i,s)(z_1,z_2) \triangleq D_2^2D_1K(z)z_i + \mathbf{1}_{D_2=\partial_i}2D_2D_1K(z) + s\mathbf{1}_{D_1=\partial_i}D_2^2K(z)$$

with $s=\pm1$ and $D_1, D_2 \in \{\partial_1, \partial_2\}$. Then we can bound $\partial^2_{y_b}\partial_{x_a}T_{00}$ using the triangle inequality. When $D_1 = D_2$, we have an improved estimate for II_2, II_3,

$$II_2 + II_3 = D_2^2K(D_2p(y) - D_2p(x_m) + (D_2p(y) + D_2p(x) - 2D_2p(x_m))). \tag{B.29}$$

We estimate $D_2p(y) + D_2p(x) - 2D_2p(x_m)$ using (B.26) with $f = D_2p$ and $z = x, y$.

B.3.2. The second singular term. For $x = (x_1, x_2)$ close to the y-axis or the x-axis, since we have symmetrized the integral (see (4.28) and section 4.1.5), we have another singular term in the integrand

$$T_{01} \triangleq K(y_1 - x_1, y_2 + x_2)(p(x) - p(y)), \text{ or } T_{10} \triangleq K(y_1 + x_1, y_2 - x_2)(p(x) - p(y)).$$

We have the first term if $x_2 < x_1$ and x_2 close to 0 and the second term if $x_1 < x_2$ and x_1 close to 0. We label the former case with $side = 1$ and the latter $side = 2$. See the right figure in Figure 1 for an illustration of the first case. The T_{01} term is supported in the blue region $R(x, k, S)$. Denote

$$(s_1, s_2) = (1, -1) \text{ if } side = 1, \quad (s_1, s_2) = (-1, 1) \text{ if } side = 2. \tag{B.30}$$

Case I. If $(D_1, side) = (\partial_1, 1)$ or $(\partial_2, 2)$, we obtain

$$\partial_{x_a}K(y_1 - s_1x_1, y_2 - s_2x_2) = -\partial_{y_a}K(y_1 - s_1x_1, y_2 - s_2x_2)$$

for $(a, s_1, s_2) = (1, 1, -1)$ or $(2, -1, 1)$. The computations for $\partial^2_{y_b}\partial_{x_1}T_{01}, \partial^2_{y_b}\partial_{x_2}T_{10}$ are the same as (B.27) with K and its derivatives evaluating at $z = (y_1 - s_1x_1, y_2 - s_2x_2)$.

We estimate II_i in (B.27) directly using the triangle inequality and the bounds for $\partial_1^i\partial_2^jK$ in sections B.1, B.2 and p in section A.1. For I_i in (B.27) in the most singular term, if $i = side$, from definition (B.30), we get

$$s_i = 1, \quad s_{3-i} = -1, \quad z_i = y_i - s_ix_i = y_i - x_i, \quad z_{3-i} = y_{3-i} + x_{3-i}.$$

Therefore, it follows that

$$\begin{aligned}I_i &= D_2^2D_1K(z)\cdot(y_i-x_i) + \mathbf{1}_{D_2=\partial_i}2D_2D_1K(z) + \mathbf{1}_{D_1=\partial_i}D_2^2K(z)\\ &= K_{mix}(D_1,D_2,i,1)(z),\end{aligned}$$

where K_{mix} is defined in (B.28). If $i \neq side$, we have $z_i = y_i + x_i \geq |y_i - x_i|, z_{3-i} = y_{3-i} - x_{3-i}$. We simply bound the summand using the triangle inequality

$$|I_i| \leq |D_2^2 D_1 K(z)| \cdot |y_i - x_i| + \mathbf{1}_{D_2=\partial_i} 2|D_2 D_1 K(z)| + \mathbf{1}_{D_1=\partial_i} |D_2^2 K(z)|.$$

Case II. If $(D_1, side) = (\partial_1, 2)$ or $(\partial_2, 1)$, we obtain

$$\partial_{x_a} K(y_1 - s_1 x_1, y_2 - s_2 x_2) = (\partial_{y_a} K)(y_1 - s_1 x_1, y_2 - s_2 x_2)$$

for $(a, s_1, s_2) = (1, -1, 1)$ or $(2, 1, -1)$. Recall the definitions of D_1, D_2 (B.25). Using the above identity, we get

$$\partial_{y_b}^2 \partial_{x_a} T = \partial_{y_b}^2 (D_1 K \cdot (p(x) - p(y)) + K \cdot D_1 p) = -(\partial_{y_b}^2 (D_1 K \cdot (p(y) - p(x)) - K \cdot D_1 p))$$

for $T = T_{01}$ or T_{10}. Using an expansion similar to that in (B.27), (B.26), we get

(B.31)

$$\begin{aligned}
-\partial_{y_b}^2 \partial_{x_a} T = &\sum_{i=1,2} \left(D_2^2 D_1 K \cdot (y_i - x_i) + \mathbf{1}_{D_2=\partial_i} 2 D_2 D_1 K - \mathbf{1}_{D_1=\partial_i} D_2^2 K \right) \cdot \partial_{x_i} p(x_m) \\
&+ D_2^2 D_1 K \cdot p_{m,2,err} + 2 D_2 D_1 K \cdot (D_2 p(y) - D_2 p(x_m)) \\
&- D_2^2 K \cdot (D_1 p(x) - D_1 p(x_m)) + D_1 K \cdot D_2^2 p(y) \\
\triangleq &\left(\sum_{i=1,2} I_i \cdot \partial_{x_i} p(x_m) \right) + II_1 + II_2 + II_3 + II_4, \\
I_i \triangleq &\, D_2^2 D_1 K \cdot (y_i - x_i) + \mathbf{1}_{D_2=\partial_i} 2 D_2 D_1 K - \mathbf{1}_{D_1=\partial_i} D_2^2 K,
\end{aligned}$$

where $\partial_1^i \partial_2^j K$ is evaluated at $z = (y_1 - s_1 x_1, y_2 - s_2 x_2)$. We bound II_i using the triangle inequality, the estimate (B.29), and the bounds for K, its derivatives, and p in sections B.1, B.2, and A.1.

For I_i, if $i = side$, from (B.30), we get $s_i = 1$ and $z_i = y_i - s_i x_i = y_i - x_i$. Hence, we get

$$I_i = D_2^2 D_1 K \cdot (y_i - x_i) + \mathbf{1}_{D_2=\partial_i} 2 D_2 D_1 K - \mathbf{1}_{D_1=\partial_i} D_2^2 K = K_{mix}(D_1, D_2, i, -1)(z),$$

where K_{mix} is defined in (B.28).

If $i \neq side$ and $D_1 = D_2 = \partial_i$, we have $z_i = y_i - s_i x_i = y_i + x_i$ and get a cancellation between $D_2 D_1 K$ and $D_2^2 K$, yielding

$$|I_i| = |D_2^2 D_1 K \cdot (y_i - x_i) + D_2 D_1 K| \leq |D_2^2 D_1 K| \cdot |y_i - x_i| + |D_2 D_1 K|.$$

Otherwise, we simply bound each term in I_i using the triangle inequality.

B.3.3. X-discretization. For $K(s) = \frac{s_1 s_2}{|s|^4}, \frac{1}{2} \frac{s_1^2 - s_2^2}{|s|^4}$, we have $K(s) = K(-s)$. Denote

$$T = K(y - x)(p(x) - p(y)) = K(x - y)(p(x) - p(y)).$$

In this section, we compute $\partial_{x_b}^i \partial_{x_a}^j T$. Using the Taylor expansion at x,

$$p(x) - p(y) = (x - y) \cdot \nabla p(x) + p_{x,2,err},$$

and calculations similar to those in section B.3.1, we get

(B.32)

$$\begin{aligned}
\partial_{x_b}^2\partial_{x_a}T &= \partial_{x_b}^2(D_1K\cdot(p(x)-p(y))+KD_1p(x))\\
&= D_2^2D_1K\cdot(p(x)-p(y))+2D_1D_2K\cdot D_2p(x)\\
&\quad + D_1K\cdot D_2^2p(x)+D_2^2K\cdot D_1p(x)+2D_2K\cdot D_1D_2p(x)+K\cdot D_1D_2^2p(x)\\
&= \sum_{i=1,2}(D_2^2D_1K\cdot(x_i-y_i)+\mathbf{1}_{D_2=\partial_i}2D_1D_2K+\mathbf{1}_{D_1=\partial_i}D_2^2K)\partial_ip(x)\\
&\quad + D_2^2D_1K\cdot p_{x,2,err}+D_1K\cdot D_2^2p(x)+2D_2K\cdot D_1D_2p(x)+K\cdot D_1D_2^2p(x)\\
&\triangleq \left(\sum_{i=1,2}I_i\cdot\partial_ip(x)\right)+II,\\
I_i &\triangleq D_2^2D_1K\cdot(x_i-y_i)+\mathbf{1}_{D_2=\partial_i}2D_1D_2K+\mathbf{1}_{D_1=\partial_i}D_2^2K,
\end{aligned}$$

where II consists of the last four terms in the third equation, K and its derivatives are evaluated at $x-y$. Since $D_1, D_2=\partial_{x_i}$, we get

$$I_i = D_2^2D_1K\cdot(x_i-y_i)+\mathbf{1}_{D_2=\partial_i}2D_1D_2K+\mathbf{1}_{D_1=\partial_i}D_2^2K = K_{mix}(D_1,D_2,i,1)(x-y),$$

where K_{mix} is defined in (B.28). We use the bound for K_{mix}, $\partial_1^i\partial_2^jK$ and p to estimate $D_2^2D_1T$.

B.3.4. The second singular term. Similar to section B.3.2, we have the second singular term for x close to the x-axis or y-axis,

$$T_{01}\triangleq K(x_1-y_1,x_2+y_2)(p(x)-p(y)),\quad T_{10}\triangleq K(x_1+y_1,x_2-y_2)(p(x)-p(y)).$$

We have the former if $x_2<x_1$ and x_2 close to 0, and the latter if $x_1<x_2$ and x_1 close to 0. Using the definition of $side, s_1, s_2$ from section B.3.2 and (B.30), we get

$$\partial_{x_a}K(x_1-y_1s_1,x_2-y_2s_2)=(D_1K)(x_1-y_1s_1,x_2-y_2s_2).$$

Then the computations of $D_2^2D_1T$ are the same as those in (B.32) with $\partial_1^i\partial_2^jK$ evaluated at $z=(x_1-s_1y_1,x_2-s_2y_2)$. We bound II in (B.32) directly using the triangle inequality and the bounds for $\partial_1^i\partial_2^jK$ and p. For I_i in (B.32), if $i=side$, from (B.30), we get s_i and $z_i=x_i-s_iy_i=x_i-y_i$. It follows that

$$I_i = D_2^2D_1K\cdot z_i+\mathbf{1}_{D_2=\partial_i}2D_1D_2K+\mathbf{1}_{D_1=\partial_i}D_2^2K = K_{mix}(D_1,D_2,i,1)(z).$$

If $i\neq side$, we have $z_i=x_i+y_i>|x_i-y_i|$. We bound each term in I_i separately by following the previous argument.

B.4. Estimate of $u(x)$ for small x_1. In the energy estimate, we need to estimate $(u(x)-\hat{u}(x))\varphi(x)$ with weight φ singular along the line $x_1=0$, e.g., φ_1 (A.2), where $\hat{u}(x)$ is a finite rank approximation of $u(x)$. We use the property that u vanishes on $x_1=0$ to establish such an estimate.

By definition and symmetrizing the kernel using the odd symmetry of ω, we have

$$\begin{aligned}
u(x,y) &= \frac{1}{2\pi}\int_{y_1\geq0}\Big(\frac{x_2-y_2}{|x-y|^2}-\frac{x_2-y_2}{(x_1+y_1)^2+(x_2-y_2)^2}\Big)\omega(y)dy\\
&= \frac{1}{\pi}\int_{y_1\geq0}K(x,y)W(y)dy,
\end{aligned}$$

where

$$
\begin{aligned}
K &= \frac{1}{2}\left(\frac{x_2-y_2}{|x-y|^2} - \frac{x_2-y_2}{(x_1+y_1)^2+(x_2-y_2)^2}\right)\\
&= x_1 \cdot \frac{2(x_2-y_2)y_1}{|x-y|^2((x_1+y_1)^2+(x_2-y_2)^2)}\\
&\triangleq x_1 K_{du}(x,y) = x_1 \tilde{K}_{du}(x_1,y_1,x_2-y_2),\\
\tilde{K}_{du}(x,y,z) &= \frac{2yz}{((x-y)^2+z^2)((x+y)^2+z^2)}.
\end{aligned} \tag{B.33}
$$

We define K_{app} as the symmetrized kernel in $\mathbb{R}^{++}_2$ for $\hat{u}$ similar to that in section 4.2. Since W is odd in y_2, we can symmetrize the integral in y_2 and obtain the full symmetrized integrand

$$x_1 K_{du}(x,y) - x_1 K_{du}(x_1,x_2,y_1,-y_2) = x_1(\tilde{K}_{du}(x_1,y_1,x_2-y_2) - \tilde{K}_{du}(x_1,y_1,x_2+y_2)).$$

Since K is -1 homogeneous, using a rescaling argument, for $x=\lambda\hat{x}, y=\lambda\hat{y}$, we have

$$u = \frac{\lambda}{\pi}\int_{\hat{y}_1\geq 0}\Big(\mathbf{1}_{S^c}(\hat{y})K(\hat{x},\hat{y}) - K_{app,\lambda}(\hat{x},\hat{y})\Big)\omega_\lambda(\hat{y}) + \mathbf{1}_S(\hat{y})K(\hat{x},\hat{y})\omega_\lambda(\hat{y})d\hat{y} \triangleq I + II \tag{B.34}$$

for some rescaled kernel $K_{app,\lambda}(\hat{x},\hat{y})$, where $S = R(\hat{x},k)$ is the singular region (4.18) adapted to $\hat{x}$. For I, we further rewrite it and estimate it as follows:

$$I = \frac{\lambda}{\pi}\hat{x}_1\int_{\hat{y}_1\geq 0,\hat{y}\notin S}\Big(\mathbf{1}_{S^c}(\hat{y})K_{du}(\hat{x},\hat{y}) - \frac{1}{\hat{x}_1}K_{app,\lambda}(\hat{x},\hat{y})\Big)\omega_\lambda(\hat{y})d\hat{y}.$$

Since the integrand is not singular, we further symmetrize the integrand in y_2 and then use the method in section 4.1.3 to discretize and estimate the integral to obtain its tight bound.

Derivative bounds. To estimate the error in the trapezoidal rule in Lemma 4.2, we need to bound $\partial^2_{x_i}K_{du}(x,y), \partial^2_{y_i}K_{du}(x,y)$. Since $\frac{1}{x}C_{u0}(x,y), \frac{1}{x}C_u(x,y)$ (4.5) are smooth, from the construction in section 4.3, the kernel $\frac{1}{x_1}K_{app}(x,y)$ and its rescaled version are regular in $\hat{x}$. We estimate its derivatives following section 4.1. Since $K_{du}(x,y) = \frac{1}{x_1}K(x,y)$ (B.33), $K(x,y)$ is harmonic in y, and $|\partial^2_{x_2}K(x,y)| = |\partial^2_{y_2}K(x,y)|$, we get

$$\partial^2_{y_1}K_{du}(x,y) = -\partial^2_{y_2}K_{du}(x,y), \quad |\partial^2_{y_2}K_{du}(x,y)| = |\partial^2_{x_2}K_{du}(x,y)|.$$

Thus, we only need to bound $|\partial^2_{x_1}K_{du}|$ and $|\partial^2_{y_1}K_{du}|$ or $\partial^2_x\tilde{K}_{du}$ and $\partial^2_y\tilde{K}_{du}$ using the relation (B.33). We derive the formulas of $\partial^2_x\tilde{K}_{du}$ and $\partial^2_y\tilde{K}_{du}$ and then estimate them using methods similar to that in Appendix B.2. We have an improved estimate for $\partial_y\tilde{K}_{du}$ in $\{x\}\times[y_l,y_u]\times[z_l,z_u]$ near the singularity. A direct computation yields

$$
\begin{aligned}
\partial^2_y\tilde{K}_{du}(x,y,z) &= 24yz\frac{(z^4-(x^2-y^2)^2)(x^2+y^2+z^2)}{T_-^3T_+^3} + 64\frac{x^2y^3z^3}{T_-^3T_+^3}\\
&= \frac{yz}{T_-^2T_+^2}\Big(12\Big(\frac{1}{T_-}+\frac{1}{T_+}\Big)(z^4-(x-y)^2(x+y)^2) + 64x^2\frac{y^2z^2}{T_-T_+}\Big), \quad T_\pm = (x\pm y)^2+z^2,
\end{aligned}
$$

where we have used $\frac{1}{T_-}+\frac{1}{T_+} = 2\frac{x^2+y^2+z^2}{T_-T_+}$. We apply the estimate of K_{du} to $x,y\geq 0$. Since $|\partial^2_y\tilde{K}_{du}|$ is even in z, without loss of generality, we consider $z\geq 0$. Then for

P_2, we have $z/T_-^{1/2}, y/T_+^{1/2}$ are increasing in z, y, respectively. To bound other terms, we simply use the monotonicity of the polynomials, (B.22), interval operation (A.4), (A.5), and follow section B.2.1. For example, we use (B.23) to bound $(x-y)^2, (x+y)^2$ and

$$0 \le \frac{y}{T_+^{1/2}} \le \frac{y_u}{((x+y_u)^2+z_l^2)^{1/2}}, \quad 0 \le \frac{z}{T_-^{1/2}} \le \frac{z_u}{(|x-y|_l^2+z_u^2)^{1/2}}.$$

$\hat{x}_1$ not small. For II in (B.34), if $\hat{x}_1 \ge x_l = 2h > 0$ away from 0, we have $|K_{du}(\hat{x},\hat{y})| \lesssim \frac{1}{x_l}\frac{1}{|\hat{x}-y|}$, which is integrable near the singularity $\hat{x}$. We estimate II using

$$|II| \le \frac{\lambda}{\pi}\hat{x}_1 \int_{\hat{y}_1\ge 0, \hat{y}\in S} |K_{du}(\hat{x},\hat{y})|\varphi_\lambda^{-1}(\hat{y})d\hat{y}||\omega\varphi||_\infty, \quad S = R(\hat{x},k).$$

We follow section 4.1.6 by introducing $\hat{y} = \hat{x} + s, s \in S - \hat{x}$, decomposing $S - \hat{x}$ into the symmetric part D_{sym} and nonsymmetric part D_{ns} and estimating the piecewise integral of $K_{du}(\hat{x},\hat{y})$,

$$D_{sym} = R_s(\hat{x},k) - \hat{x},\ D_{ns} = (R(\hat{x},k)\backslash R_s(\hat{x},k)) - \hat{x},$$
$$|K_{du}(\hat{x},\hat{y})|\mathbf{1}_{\hat{y}_1\ge 0} = |F|\mathbf{1}_{\hat{x}_1+s_1\ge 0},\ F = \frac{(\hat{x}_1+s_1)s_2}{|s|^2((s_1+2\hat{x}_1)^2+s_2^2)},$$

and piecewise bounds of $\varphi_\lambda^{-1}(y)$, where we have used (B.33) to obtain the above formula. We observe that $|F|$ is even in s_2 and $F \ge 0$ for $s \in Q = [a,b]\times[c,d]$ with $c, d \ge 0$. We estimate the piecewise integrals of F in Q in section 6.2 in the supplementary material (supplement.pdf [local/web 1.43MB]). Denote $X_1^+ \triangleq \{y : y_1 \ge 0\}$. If $\hat{x}_1 \ge kh$, we get $S \cap X_1^+ = R(\hat{x},k)$ and the regions D_{sym}, D_{ns} are the same as those in section 4.1.6. If $\hat{x}_1 \in [ih,(i+1)h), i < k$, the region S touches $\{y : y_1 = 0\}$ and we get

$$S \cap X_1^+ = [0,(i+k+1)h] \times [(j-k)h,(j+1+k)h] \ \text{ for } x_2 \in [jh,(j+1)h].$$

In this case, the symmetric and nonsymmetric region becomes smaller. We do not have the left edge in the middle figure in Figure 2, part of the upper and the lower edge due to the restriction $\hat{y}_1 = s_1 + \hat{x}_1 \ge 0$. The estimate of the integrals for $s \in S \cap X_1^+ - \hat{x}_1$ follows similar argument.

Small $\hat{x}_1$. The difficulty is to estimate II for small $\hat{x}_1 \le 2h$. It is not difficult to obtain that

$$|II| \lesssim \frac{\lambda}{\pi}||\omega_\lambda||_{L^\infty(S)}\hat{x}_1|\log(\hat{x}_1)|. \tag{B.35}$$

Thus we cannot bound II by $C\hat{x}_1$ for some constant C uniformly for small $\hat{x}_1$. Denote by

$$\begin{aligned} &S_{sym} = [0,\hat{x}_1+kh]\times[\hat{x}_2-kh,\hat{x}_2+kh], \quad S_{in,1} = [0,\hat{x}_1]\times[\hat{x}_2-kh,\hat{x}_2+kh], \\ &S_{in,2} = [\hat{x}_1,\hat{x}_1+h]\times[\hat{x}_2-h,\hat{x}_2+h], \quad S_{in} = S_{in,1}\cup S_{in,2}, \\ &S_{out} = [\hat{x}_1,\hat{x}_1+hk]\times[\hat{x}_2-kh,\hat{x}_2+kh]\backslash S_{in,2}, \quad \hat{y} = \hat{x}+\hat{x}_1 s. \end{aligned} \tag{B.36}$$

See the right figure in Figure 2 for an illustration of different regions. By definition, we have $S_{sym} = S_{out}\cup S_{in,1}\cup S_{in,2}$. Here S_{in} captures the most singular region. Then $\hat{y} \in S_{in}$ is equivalent to

(B.37)

$$s \in \hat{x}_1^{-1}(S_{in} - \hat{x}) = x_1^{-1}([-\hat{x}_1, 0] \times [-kh, kh] \cup [0, h] \times [-h, h]) \triangleq R_1(B_1) \cup R_2(B_2),$$
$$R_1(B_1) = [-1, 0] \times \Big[-\frac{1}{B_1}, \frac{1}{B_1} \Big],$$
$$R_2(B_2) = \Big[0, \frac{1}{B_2}\Big] \times \Big[-\frac{1}{B_2}, \frac{1}{B_2} \Big],\ B_1 = \frac{\hat{x}_1}{kh},\ B_2 = \frac{\hat{x}_1}{h}.$$

We further decompose II as follows:

$$II = \frac{\lambda}{\pi} \hat{x}_1 \int_{y_1 \geq 0} (\mathbf{1}_{S \setminus S_{sym}}(\hat{y}) + \mathbf{1}_{S_{out}}(\hat{y}) + \mathbf{1}_{S_{in,1}}(\hat{y}) + \mathbf{1}_{S_{in,2}}(\hat{y})) K_{du}(\hat{x}, \hat{y}) \omega_\lambda(\hat{y}) d\hat{y}$$
$$= \frac{\lambda \hat{x}_1}{\pi} (II_1 + II_2 + II_{in,1} + II_{in,2}).$$

The integrals II_1, II_2 capture the nonsymmetric part and the symmetric part away from the singularity. We apply L^∞ estimate and the method in sections 4.1.6, 4.1.9. For $II_{in,i}$, using a change of variables (B.36), (B.37), we derive

$$II_{in,i} = \int_{s \in R_i(B_i)} K_{du}(\hat{x}, \hat{x} + \hat{x}_1 s) \hat{x}_1^2 \omega_\lambda(\hat{x} + \hat{x}_1 s) ds.$$

Note that $\hat{y} - \hat{x} = \hat{x}_1 s,\ \hat{y}_1 + \hat{x}_1 = \hat{x}_1(2 + s_1),\ \hat{y}_2 - \hat{x}_2 = \hat{x}_1 s_2$. By definition (B.33), we get

$$K_{du}(\hat{x}, \hat{x} + \hat{x}_1 s) \hat{x}_1^2 = -\frac{2\hat{x}_1 s_2 \cdot (\hat{x}_1 + \hat{x}_1 s_1)}{\hat{x}_1^2 |s|^2 \cdot \hat{x}_1^2 ((s_1 + 2)^2 + s_2^2)} \hat{x}_1^2 = -\frac{2(s_1 + 1) s_2}{|s|^2 ((s_1 + 2)^2 + s_2^2)} \triangleq -K_s(s),$$
$$II_{in,i} = -\int_{R_i(B_i)} K_s(s) \omega_\lambda(\hat{x} + \hat{x}_1 s) ds.$$

Since $K_s(s)$ is symmetric in s_2, we derive

$$|II_{in,1}| \leq ||\omega\varphi||_\infty \left(\max_{z \in [-\hat{x}_1, 0] \times [0, kh]} \varphi_\lambda^{-1}(\hat{x} + z) + \max_{z \in [-\hat{x}_1, 0] \times [-kh, 0]} \varphi_\lambda^{-1} \right) J_1(B_1),$$
$$|II_{in,2}| \leq ||\omega\varphi||_\infty \left(\max_{z \in [0,h] \times [0,h]} \varphi_\lambda^{-1} + \max_{z \in [0,h] \times [-h,0]} \varphi_\lambda^{-1} \right) J_2(B_2),$$

where B_i is given in (B.37) and

$$J_1(B_1) = \left| \int_{[-1,0] \times [0, 1/B_1]} K_s(s) ds \right| = \int_{[-1,0] \times [0, 1/B_1]} K_s(s) ds,$$
$$J_2(B_2) = \int_{[0, 1/B_2]^2} K_s(s) ds.$$

The formula of J_i can be obtained using the analytic integral formula for K_s, and obviously J_i is decreasing in B. Note that $J_1(B)$ is bounded, but $J_2(B) \lesssim 1 + \log(B) \lesssim 1 + |\log \hat{x}_1|$, which relates to the estimate (B.35). We defer the formulas of J_i to section 6.2 in the supplementary material (supplement.pdf [local/web 1.43MB]).

B.5. Additional derivations.

B.5.1. Estimate of the log-Lipschitz integral. In this section, we derive the coefficient in the estimate of $\partial_{x_2} I_{5,4}(x)$ (4.64), (4.65). For $I_{5,4}$, we further decompose it as follows:

$$\begin{aligned} I_{5,4} &= \Big(\int_{R(k_2)\setminus R_s(k_2)} + \int_{R_s(k_2)\setminus R_s(b)} + \int_{R_s(b)\setminus R_s(a)} \Big) K(x-y)(\psi(x)-\psi(y)W(y)dy \\ &\triangleq I_{5,4,1} + I_{5,4,2} + I_{5,4,3}. \end{aligned}$$

In practice, we choose $b = 2$. The first two terms are nonsingular and their derivatives can be estimated using the method in sections 4.1.6–4.1.9. In the estimate of $\partial_{x_i} I_{5,4}$, we only need to estimate the boundary term on $\partial R_s(a)$ since the boundary terms on $\partial R_s(k_2), \partial R_s(b)$ are canceled in $\partial_{x_i} I_{5,j}, j = 1,2,3$. For $I_{5,4,3}$, using the second order Taylor expansion to $\psi(x) - \psi(y)$ centered at x, we have

$$\begin{aligned} &\partial_{x_2}(K(x-y)(\psi(x)-\psi(y))) = (\partial_2 K)(x-y)(\psi(x)-\psi(y)) + K(x-y)\partial_2\psi(x) \\ &= (\partial_2 K(x-y)(x_2-y_2) + K(x-y))\partial_2\psi(x) + \partial_2 K(x-y)(x_1-y_1)\partial_1\psi(x) + \mathcal{R}_K, \end{aligned}$$

where the remainder $\mathcal{R}_K$ coming from the higher order term in the Taylor expansion satisfies

$$|\mathcal{R}_K| \le \sum_{i+j=2} ||\partial_x^i \partial_y^j \psi||_{L^\infty(Q)} |x_1 - y_1|^i |x_2 - y_2|^j c_{ij},$$

where $Q = B_{i_1 j_1}(h_x) + [-bh, bh]^2$ and $c_{20} = c_{02} = \frac{1}{2}, c_{11} = 1$. It follows that

$$|\partial_{x_2} I_{5,4,3}| \le ||\omega\varphi||_\infty \sum_{0\le i\le 1, 0\le j\le i+1} Scoe_{ij}(x) \cdot f_{ij}(a,b),$$

where the coefficients $Scoe_{ij}(x)$ depend on the weight ψ, φ, and $f_{ij}(a,b)$ bounds the integral

$$\int_{[-b,b]^2\setminus[-a,a]^2} |\partial_2 K(y) \cdot y_1^i y_2^j + \mathbf{1}_{(i,j)=(0,1)} K(y)| dy \le f_{ij}(a,b). \tag{B.38}$$

For example, $Scoe_{01}$ comes from the following estimate for $I_{5,4,3}$

$$\begin{aligned} &\int_{R_s(b)\setminus R_s(a)} |(\partial_2 K(x-y)(x_2-y_2) + K(x-y))\partial_2\psi(x)|\omega(y)dy \\ &\le ||\omega\varphi||_\infty ||\varphi^{-1}||_{L^\infty(Q)} \cdot |\partial_2\psi(x)| \int_{[-b,b]^2\setminus[-a,a]^2} |\partial_2 K(s)s_2 + K(s)| ds. \end{aligned}$$

The function $f_{ij}(a,b)$ satisfies the following estimates for some constants $B_{1j} > 0$

$$f_{1j}(a,b) \le B_{1j} \log(b/a), \quad j = 1,2.$$

We defer the derivations to section 5.1.5 in the supplementary material (supplement.pdf [local/web 1.43MB]).

B.5.2. Optimization in the Hölder estimate. Consider

$$\max_{t\le t_u} \min_{a\le b} F(a,t), \quad F(a,t) = \Big(A + B\log\frac{b}{a}\Big)\sqrt{t} + \frac{Ca}{\sqrt{t}},$$

in the upper bound in (4.72). For each $t \le t_u$, we first optimize $F(a,t)$ over $a \le b$. We assume that A, B, C, b, c, h, h_x are given. Denote

$$t_u = ch_x, \quad t_1 = \frac{Cb}{B}.$$

For a fixed t, since $\partial_a^2 F > 0, \partial_a F(0,t) < 0$, and $\partial_a F(a,t) = 0$ if $a = \frac{Bt}{C}$, we choose $a = \min(b, \frac{Bt}{C})$. For $t \le \frac{Cb}{B} = t_1$, we get

$$\min_{a \le b} F(a,t) \le F\Big(\frac{Bt}{C}, t\Big) = \Big(A + B\log\frac{bC}{B} + B\Big)\sqrt{t} - B\sqrt{t}\log t.$$

The right hand side can be further estimated by studying the concave function on $s = t^{1/2} \le s_u$,

$$f(p,q,s) = (p - q\log s)s \le f(p,q,\min(s_u, s_*)), \quad s_* = \exp\Big(\frac{p-q}{q}\Big)$$

with $p = A + B\log(\frac{bC}{B}) + B, q = 2B, s_u = \min(t_u^{1/2}, t_1^{1/2})$. We get the above inequality since $f(p,q,s)$ is increasing for $s \le s_*$ and is decreasing for $s \ge s_*$.

If $\frac{Cb}{B} \le t \le t_u$, we choose $a = b$ and get

$$\min_{a \le b} F(a,t) \le F(b,t) = A\sqrt{t} + \frac{Cb}{\sqrt{t}},$$

which is convex in $t^{1/2}$. Thus its maximum is achieved at the endpoints.

Appendix C. Representations and estimates of the solutions. In section 7 of Part I [13], we represent the approximate steady state as follows:

$$\begin{aligned}
&\bar\omega = \bar\omega_1 + \bar\omega_2, \quad \bar\theta = \bar\theta_1 + \bar\theta_2, \quad \bar\omega_1 = \chi(r) r^{-\bar\alpha_1} g_1(\beta), \quad \bar\theta_1 = \chi(r) r^{1-2\bar\alpha_1} g_2(\beta),\\
&\bar\phi^N = \bar\phi_1^N + \bar\phi_2^N + \bar\phi_3^N + \bar\phi_{cor}^N, \quad \bar\phi_3^N = \bar a \chi_{\phi,2D}, \quad \chi_{\phi,2D} = -xy\chi_\phi(x)\chi_\phi(y),\\
&\bar\phi_{cor}^N = -c \cdot \frac{xy^2}{2}\kappa_*(x)\kappa_*(y) = c\phi_1, \quad c = \partial_x(\bar\omega + \Delta(\bar\phi_1^N + \bar\phi_2^N + \bar\phi_3^N)),\\
&|\bar\alpha_1 + \frac{\bar c_\omega}{\bar c_l}| \ll 1, \quad \bar\alpha_1 \approx \frac{1}{3},
\end{aligned} \tag{C.1}$$

where $\bar\omega_2, \bar\theta_2, \bar\phi_2^N$ have compact supports and are represented as piecewise polynomials, $\bar a \in \mathbb{R}$ is some coefficient, κ_* is given in (D.5), ϕ_1 is the same as (3.14), and χ_ϕ is given in (D.7). We choose a small correction $\bar\phi_{cor}$ similar to that in section 3.2 so that $\bar\omega + \Delta\bar\phi^N = O(|x|^2)$ near 0. We use upper script N to distinguish the numerical approximation $\bar\phi^N$ for the exact stream function $\bar\phi = (-\Delta)^{-1}\bar\omega$. The exponent $\bar\alpha_1$ and angular profiles $g_i(\beta)$ are obtained by fitting the far-field asymptotics of an approximate steady state with $\bar\omega_1 = 0, \bar\theta_1 = 0$. Then we construct $(\bar\omega_1, \bar\theta_1)$ using the above formulas. Afterward, we refine the construction of the near-field part $(\bar\omega_2, \bar\theta_2)$ and exponents $(\bar c_\omega, \bar c_l)$ by fixing $(\bar\omega_1, \bar\theta_1, \bar\alpha_1)$. See more details on how to find the semianalytic part in section 7 of Part I [13]. We will discuss how to estimate the semianalytic part in section C.3. In the following sections, we discuss more details about the representations and establish a rigorous estimate of the derivatives of $\bar\omega, \bar\theta$.

Note that we do not need an approximation term $\bar\phi_3$ for the stream function in solving the linearized equation in section 3 since we can allow a larger residual error in section 3.

C.1. Representations. In a large domain $[0,L]^2$, we use piecewise polynomials to represent the solution. First, we choose a large L of order 10^{15} and then design the adaptive mesh $y_{-5}<\cdots<y_0=0<y_1<\cdots<y_{N-1}=L, N=748$ to partition $[0,L]$.

Adaptive mesh. We design three parts of the mesh $y_i, i\in I_j \triangleq [a_j,b_j], a_0=0$ as follows:

$$
\begin{aligned}
&y_i=\frac{i}{256}, i=-5,-4,\dots,1,\dots,b_1, \quad y_{a_2+i}=y_{a_2}+F(ih_3), i=1,\dots,b_2-a_2,\\
&y_{a_3+i}=y_{a_3}\exp(ir_1), i=1,\dots,b_3-a_3, \quad r_0=1.025,\ r_1=1.15\\
&F(z)=\frac{h_2}{h_3}z\exp(rz^2),\ r=\log\Big(\frac{r_0}{1+h_3}\Big)\frac{1}{(1+h_3)^2-1},\ h_2=\frac{1}{128}, h_3=\frac{1}{b_2-a_2}.
\end{aligned}
\tag{C.2}
$$

Since we need to estimate the weighted L^∞ norm of the residual error with a singular weight of order $|x|^{-\beta}, \beta\approx 3$ near $x=0$, we use a uniformly dense mesh near 0 so that we have a very small residual error. We choose the parameters $\frac{1}{256}, h_2=\frac{1}{128}$ since they can be represented exactly as floating point numbers. Thus, we can reduce the round-off error in the computation. In the far-field, we use a mesh that grows exponentially fast in space. Note that the error estimate $f-I(f)$ for the kth order interpolation of f on $[y_i,y_{i+1}]$ reads

$$|f-I(f)|\le C(y_{i+1}-y_i)^k|\partial_x^k f|.$$

For large x, we expect that $\partial_x^k f$ has a decay rate $|y|^{-k-\alpha}$ if $|f|\lesssim |y|^{-\alpha}$ for $\alpha>0$. Thus, to get a uniformly small error in the far-field, we just require $\frac{y_{i+1}-y_i}{y_i}\le\varepsilon$ with $\varepsilon<1$. This allows us to choose an exponentially growing mesh in the far-field and cover a very large domain without using too many points. We use the second part of the mesh to glue the first part of the mesh, which grows linearly, and the third part of the mesh. The function $F(z)$ behaves linearly for z close to 0, and it grows exponentially fast with rate r_1 for z close to 1:

$$F(1+h_3)/F(1)=(1+h_3)\exp(r((1+h_3)^2-1))=(1+h_3)\exp(\log(r_0/(1+h_3)))=r_0.$$

Parameters h_2,h_3 control the mesh size $y_{a_2+1}-y_{a_2}=F(h_3)=h_2\exp(rh_3^2)\approx h_2$. One can design another $F(z)$ by gluing the first and the third part of the mesh. The above explicit and simple form of $F(z)$ serves our purpose. We further glue $y_i, i\in[b_j,a_{j+1}], j=1,2$, using the Lagrangian interpolation for $j=1$. For $j=2$, we interpolate the growth rate using $\exp(\log(r_0)l(i)+(1-l(i))\log(r_1))$ with $l(i)$ linear in $i\in[b_2,a_3]$. Note that we do not use the specific property of the profile to design the adaptive mesh (C.2).

In our numerical computation, we compute the derivatives of the solution using the B-spline basis (see, e.g., (C.6)) and do not use the Jacobian related to the adaptive mesh. In particular, we do not use derivatives of the map $f(i)=y_i$ and have more flexibility to design the mesh.

Let $n_1=720<N$. We solve the dynamic rescaling equation (2.10)–(2.11) on first $n_1\times n_1, (y_i,y_j), i,j\le n_1-1$ grids. We construct

$$\bar\omega_2(x,y)=\sum_{0\le i\le n_1+11,-2\le j\le n_1+1} a_{ij}B_{1,i}(x)B_j(y), \tag{C.3}$$

where $a_{ij}\in\mathbb{R}$ is the coefficient, and $B_i(x), B_j(y)$ are constructed from the sixth order B-spline

$$B_i(x)=C_iB_{i0}(x), \quad B_{i0}(x)=\sum_{0\le j\le k} k\frac{(s_{ij}-x)_+^{k-1}}{d_j}, \quad d_j=\prod_{0\le l\le k, l\ne j}(s_{ij}-s_{il}), \tag{C.4}$$

with $k = 6$. The constant C_i will be chosen in (C.10), (C.11) so that the stiffness matrix associated to the B-spine basis has a better condition number. We choose s_{ij} as follows:

$$s_{ij} = y_{i+j-3}, \quad 0 \le j \le k = 6.$$

Then the B-spline B_i is supported in $[y_{i-3}, y_{i+3}]$ and is centered around y_i. Since ω is odd in x, to impose this symmetry in the representation, we modify the first few bases

$$B_{1,i}(x) = B_i(x) - B_i(-x), \quad i \le 2. \tag{C.5}$$

Then B_i is odd. We remark that $B_{1,0}(x) \equiv 0$.

B-spline and the tensor structure. We also use the B-spline basis to represent the stream function (C.8) and solve the Poisson equation using the B-spline based finite element method. We use the B-spline basis since it is easy to design a high order numerical scheme to solve the Poisson equation. Each basis function in (C.3), (C.7), (C.8) has the form $f(x)g(y)$, which allows us to evaluate and estimate the 2D function very effectively using the method in Appendix C.2.2.

Remark C.1. While the method described below to obtain the coefficients a_i is technical, since we perform a posteriori estimates of the profiles and residual error using the given a_i, the method of deriving a_i is not involved in the a posteriori estimates and the verification process.

Extrapolation. Near the boundary $y = 0$, we need two extra basis functions $a_{i,-j}B_{-j}(y), j = 1, 2$, that are not zeros in $y_1 \ge 0$. Without these two functions, the representation (C.3) does not approximate $\bar{\omega}$ with a sixth order error. We use a seventh order extrapolation [41, 42] to determine $a_{i,-j}$:

$$a_{i,-j} = \sum_{0 \le l \le 6} C_{3-j,l+1} a_{i,l}, \; C_{1,\cdot} = (28, -112, 210, -224, 140, -48, 7),$$
$$C_{2,\cdot} = (7, -21, 35, -35, 21, -7, 1).$$

We choose $C_{j,l}$ such that the 7th difference of $a_{i,j}, -2 \le j \le 6$ is 0. Since $a_{i,-j}$ depends on $a_{i,l}$ linearly, we can combine $a_{i,-j}B_{i,-j}, j = 1, 2$ with $a_{i,l}B_{i,l}$ and modify (C.3) as follows:

$$\bar{\omega}_2(x, y) = \sum_{0 \le i,j \le n_1+1} a_{ij} B_{1,i}(x) B_{2,j}(y), \tag{C.6}$$
$$B_{2,j}(y) = B_j(y) + C_{2,j+1}B_{-1}(y) + C_{1,j+1}B_{-2}(y), \; 0 \le j \le 6, \quad B_{2,j}(y) = B_j(y), j \ge 7.$$

The modified basis functions $B_{1,i}, B_{2,j}$ are still piecewise polynomials in $[y_l, y_{l+1}]$.

Far-field extension. In (C.3), (C.6), we use B-spline $B_{1,i}(x), B_j(y)$ up to $i, j \le n_1 + 1$ rather than $n_1 - 1$ since the support of $B_{1,i}, B_j$ intersects $[0, y_{n-1}]^2$ for $i, j \le n_1 - 1$. To determine the extra coefficients, we first extend the grid point values of $\omega_2(x, y)$ from (y_i, y_j) with $i, j \le n_1 - 1$ to $i, j \le n_1 + l_0 - 1$ by $\omega_2(y_{n_1+l}, y_j) = P(y_{n_1+l}; y_j), l = 0, 1, .., l_0 - 3$, where P is the Lagrangian interpolation polynomials on $(y_{n_1-1}, \omega(y_{n_1-1}, y_j)), (y_{n_1+l_0-3}, 0), (y_{n_1+l_0-2}), 0)$. We impose $\omega_2(y_{n_1+l}, y_j) = 0, l =$

l_0-3, l_0-2, l_0-1. Similarly, we extend $\omega(y_i, y_{n+l})$. Note that ω_2 is odd and $B_{1,0}=0$. We solve the coefficients $a_{kl}, 1\le k\le M, 0\le l\le M$ from

$$\omega_2(y_p,y_q)=\sum_{1\le i\le M, 0\le j\le M} a_{ij}B_{1,i}(x)B_{2,j}(y),\ 1\le p\le M,\ 0\le q\le M,\ M=n_1+l_0-1\,.$$

The value a_{0j} is not used since $B_{1,0}\equiv 0$. To simplify the notation, we keep it. We only keep $a_{ij}, i,j\le n_1+1$ and obtain (C.6). In practice, we choose $l_0=8$, and the above construction provides a solution with tail decaying smoothly to 0 for $|y|_\infty\ge y_{n_1+l_0-1}$.

To solve the dynamic rescaling equations numerically (2.10)–(2.12) (see section 7 Part I), we update the grid point value of ω_{n+1} at time t_{n+1} and then use the above method to obtain a_{ij}.

For the density $\bar\theta_2$, the representation is similar

$$\bar\theta_2=x\sum_{0\le i,j\le n_1+1} a_{ij}B_{1,i}(x)B_{2,j}(y). \tag{C.7}$$

Here, we multiply x since $\bar\theta$ is even and vanishes $O(x^2)$ near $x=0$.

For the stream function $\bar\phi_2^N$ (C.1), we choose $n_2>n_1$ and represent it as follows:

$$\bar\phi_2^N=\sum_{0\le i,j\le n_2-1} a_{ij}\tilde B_{1,i}(x)\tilde B_{2,j}(y)\rho_p(y). \tag{C.8}$$

Instead of using the above extension to determine the extra coefficients, we perform an additional extrapolation for the basis in the far-field similarly to (C.6):

$$\tilde B_{l,j}(z)=B_{l,j}(z),\quad j\le n_2-8,\quad \tilde B_{l,j}(z)=B_j(z)+C_{2,n_2-j}B_{n_2}(z)+C_{1,n_2-j}B_{n_2+1}(z).$$

We multiply $\rho_p(y)$ given below to impose the Dirichlet boundary condition

$$\rho_p(y)=\arctan(1+y)-\arctan(1). \tag{C.9}$$

We can obtain the exact formulas of $\partial_x^i\rho_p$ using a symbolic computation. We use induction to obtain a rigorous estimate of $\partial_x^i\rho_p$. See section D.3.

We choose C_i in (C.4) of order $s_{i,j+1}-s_{i,j}$ as

$$C_i=y_1, i\le 9,\quad C_i=(s_{i,4}-s_{i,2})/2,\ i>9, \tag{C.10}$$

so that the summand in (C.4) has order 1 for x in the support $[y_{i-3}, y_{i+3}]$. When we need to perform extrapolation for $a_nB_n, a_{n+1}B_{n+1}$ from $a_iB_i, i\le n-1$, e.g., (C.8), we modify the last few terms as follows:

$$C_i=(y_n-y_{n-1})/100,\ n-9\le i. \tag{C.11}$$

We choose C_i to be constant for i close to 0 or i close to n_1 since we need to perform extrapolation, and the choice of the constant does not affect the extrapolation formula for a_{ij}.

Far-field angular profile. To represent the far-field angular profile of $\bar\omega_1, \bar\theta_1, \bar\phi_1^N$ (C.1), we design adaptive mesh $0=\beta_0<\beta_1<\cdots<\beta_m=\pi/2$, and use an eighth order B-spline to represent $\bar\omega, \bar\zeta=\frac{\bar\theta}{x_1}$,

$$g(\pi/2-\beta)=\sum_{i\ge 0} b_iB_{1,i}^{(8)}(\beta),\quad g_\phi(\pi/2-\beta)=((\pi/2)^2-\beta^2)\sum_i b_i\tilde B_i^{(8)}(\beta), \beta\in[0,\pi/2],$$

where $B_{1,i}^{(8)}$ is an eighth order B-spline (C.4) $k=8$ with odd modification (C.5). Since $\bar{\omega},\bar{\zeta}$ are odd in x, in the angular direction, this symmetry becomes odd in $\beta=\pi/2$. To impose it, we write g in terms of $\pi/2-\beta$ and modify the first few B-splines B_i (C.4) following (C.5) so that $\tilde{B}_{1,i}$ is odd at $\beta=0$. Then g is odd in $\beta=\pi/2$. The stream function $\bar{\phi}^N$ satisfies the boundary condition $\bar{\phi}^N(x,0)=0$. For the angular profile, we need $g_\phi(0)=0$, and use the weight $\pi/2-\beta$ to impose this condition. We further modify a few B-spline $B_{1,i}(\beta)$ supported near $\beta=\pi/2$ using a 9th order extrapolation similar to (C.6) near $\beta=\pi/2$ and get $\tilde{B}_{1,i}(\beta)$. We choose the mesh β_i to be equispaced near $\beta=\pi/2$ and determine the coefficients for extrapolation similar to (C.6). We remark that to evaluate the derivative $\partial_\beta^i g$ at $\pi/2-\beta$, we have the sign $(-1)^k$

$$(\partial_\beta^k g)(\pi/2-\beta)=(-1)^k\partial_\beta^k g(\pi/2-\beta)=(-1)^k\sum b_i\partial_\beta^k B_{1,i}^{(8)}(\beta).$$

We discuss how to obtain these angular profiles using the curve fitting in section 7 in [13].

C.2. Estimate of the derivatives of piecewise polynomials. Our approximate steady state in a very large domain is represented as piecewise polynomials. We discuss how to estimate its derivatives. Suppose that we can evaluate a function f on finite many points. For example, f is an explicit function or a polynomial. To obtain a piecewise sharp bound of f on $I=[x_l,x_u]$, we use the following standard error estimate:

$$\max_{x\in I}|f(x)|\le\max(|f(x_l)|,|f(x_u)|)+\frac{h^2}{8}||f_{xx}||_{L^\infty(I)},\quad h=x_u-x_l. \tag{C.12}$$

If we can obtain a rough bound for f_{xx}, as long as the interval I is small, i.e., h is small, the error part is small. Similarly, if we can obtain a rough bound for $\partial_x^{k+2}f$, using induction and the above estimate recursively,

$$\max_{x\in I}|\partial_x^i f(x)|\le\max(|\partial_x^i f(x_l)|,|\partial_x^i f(x_u)|)+\frac{h^2}{8}||\partial_x^{i+2}f||_{L^\infty(I)}$$

for $i=k,k-1,\dots,0$, we can obtain the sharp bound for $\partial_x^i f$ on I. We call the above method the second order method since the error term is second order in h.

C.2.1. Estimate a piecewise polynomial in one dimension. Suppose that $p(x)$ is a piecewise polynomial on $x_0<x_1<\cdots<x_n$ with degree d, e.g., Hermite spline. Denote $I_i=[x_i,x_{i+1}]$. Then $p(x)$ is a polynomial in each I_i with degree $\le d$. Our goal is to estimate $\partial_x^k p(x)$ in I_i for all k by only finite many evaluations of $p(x)$ and its derivatives. First, we have

$$\partial_x^k p(x)=0,\quad k>d,\quad \partial_x^d p(x)=c_p,$$

for some constant c_p in I_i. Using induction from $k=d-1,d-2,\dots,0$, we have

$$\max_{x\in I_i}|\partial_x^k p(x)|\le\max(|\partial_x^k p(x_i)|,|\partial_x^k p(x_{i+1})|)+\frac{h_i^2}{8}||\partial_x^{k+2}p||_{L^\infty(I_i)},\quad h_i=x_{i+1}-x_i.$$

Since we know $\partial_x^{d+1}p(x)=0$ on I_i, using the above method, we can obtain the sharp piecewise bounds for all derivatives of $p(x)$ on I_i. Using the above approach, we can estimate the derivatives of the angular profile defined in section 7.1 of Part I [13] rigorously.

C.2.2. Estimate a piecewise polynomial in two dimensions. Now, we generalize the above ideas to two dimensions so that we can estimate the approximate steady state (C.6). We assume that $p(x,y)$ is a piecewise polynomial in the mesh $Q_{ij}=[x_i,x_{i+1}]\times[y_j,y_{j+1}]$ with degree d. That is, in Q_{ij}, $p(x,y)$ can be written as a linear combination of

$$x^k y^l, \quad \max(k,l)\le d,$$

e.g., (C.6). For (C.6), we have $d=5$. Similar to the 1D case, we have

$$\partial_x^k\partial_y^l p(x,y)=0, \quad \max(k,l)>d.$$

Moreover, we know $\partial_x^{d-1}\partial_y^{d-1}$ is linear in x,y.

We use the following direct generalization of (C.12) to two dimensions:

(C.13)

$$\max_{(x,y)\in Q}|f(x,y|\le \max_{\alpha,\beta=l,u}|f(x_\alpha,y_\beta)|+\frac{||f_{xx}||_{L^\infty(Q)}(x_u-x_l)^2}{8}+\frac{||f_{yy}||_{L^\infty(Q)}(y_u-y_l)^2}{8},$$
$$Q=[x_l,x_u]\times[y_l,y_u].$$

Denote

$$A_{kl}\triangleq\max_{Q_{ij}}||\partial_x^k\partial_y^l p||_{L^\infty(Q_{ij})}, B_{kl}\triangleq\max_{\alpha,\beta=l,u}|\partial_x^k\partial_y^l p(x_\alpha,y_\beta)|,$$
$$h_1=x_{i+1}-x_i, \quad h_2=y_{j+1}-y_j.$$

Since p is given, we can evaluate B_{kl}. Clearly, we have $A_{kl}=0$ for $\max(k,l)>d$. For $k=d-1,d$, using (C.13) and induction on the order $l=d,d-1,d-2,\dots,0$, we can obtain

$$A_{kl}\le B_{kl}+\frac{1}{8}(h_1^2A_{k+2,l}+h_2^2A_{k,l+2}).$$

This allows us to bound A_{kl} for $k=d,d-1$, and all l. Similarly, we can bound A_{kl} for $l=d,d-1$, and all k.

For the remaining cases, we can use induction on $n=\max(k,l)=d-2,d-1,\dots,0$ to estimate

$$A_{kl}\le B_{kl}+\frac{1}{8}(h_1^2A_{k+2,l}+h_2^2A_{k,l+2}).$$

This allows us to estimate all derivatives of $p(x,y)$ in Q_{ij}.

C.2.3. Estimate a piecewise polynomial in two dimensions with weights. We consider how to estimate the derivatives of $f=\rho(y)p(x,y)$, where ρ is a given weight in y and $p(x,y)$ is the piecewise polynomials in two dimensions. For example, our construction of the stream function (C.8) has such a form. First, we can estimate the derivatives of $p(x,y)$ using the method in Appendix C.2.2. For the weight ρ, we estimate its derivatives in section D.3. Then, using the Leibniz rule (A.6) and the triangle inequality, we can estimate the derivatives f

$$|\partial_x^i\partial_y^j f|\le\sum_{l\le j}\binom{j}{l}|\partial_x^i\partial_y^l p(x,y)||\partial_y^{j-l}\rho(y)|$$

for high enough derivatives.

Now, we plug the above bounds for $\partial_x^{i+2}\partial_j^y,\partial_x^i\partial_y^{j+2}f$ in (C.13) and evaluate $\partial_x^i\partial_y^j f$ on the grid points to obtain the sharp estimate of $\partial_x^i\partial_y^j f$.

C.3. Estimate of the far-field approximation. We estimate the derivatives of

$$g(x,y)=g(r,\beta)=A(r)B(\beta),\quad r=(x^2+y^2)^{1/2},\quad \beta=\arctan(y/x),$$

where (r,β) is the polar coordinate. The semianalytic parts of $\bar{\omega},\bar{\theta}$ have the above forms.

C.3.1. Formulas of the derivatives of g. First, we use induction to establish

$$F_{i,j}\triangleq \partial_x^i\partial_y^j g(r,\beta)=\sum_{k+l\le i+j} C_{i,j,k,l}(\beta)r^{-i-j+k}\partial_r^k A\partial_\beta^l B \tag{C.14}$$

with $C_{i,j,k,l}=0$, for $k<0$, $l<0$, or $k+l>i+j$. Let us motivate the above ansatz. Recall from (B.14) that

$$\partial_x=\cos\beta\partial_r-\frac{\sin\beta}{r}\partial_\beta,\quad \partial_y=\sin\beta\partial_r+\frac{\cos\beta}{r}\partial_\beta.$$

For each derivative ∂_x or ∂_y, we get the factor $\frac{1}{r}$ or a derivative ∂_r, which leads to the form $r^{-i-j+k}\partial_r^k A$. Moreover, we get a derivative ∂_β and some functions depending on β, which leads to the form $C_{i,j,k,l}(\beta)\partial_\beta^l B$.

For $D=\partial_x$ or ∂_y, a direct calculation yields

$$\begin{aligned}DF_{i,j}=&\sum_{k+l\le i+j} D(C_{i,j,k,l}r^{-i-j+k})\cdot\partial_r^k A\partial_\beta^l B\\&+C_{i,j,k,l}r^{-i-j+k}(D\partial_r^k A\cdot\partial_\beta^l B+\partial_r^k A\cdot D\partial_\beta^l B).\end{aligned} \tag{C.15}$$

Using the formula of ∂_x,∂_y, we get

$$\partial_x(C_{i,j,k,l}(\beta)r^{-i-j+k})=-\sin\beta\partial_\beta C_{i,j,k,l}r^{-i-j-1+k}+(k-i-j)\cos\beta C_{i,j,k,l}r^{-i-j-1+k},$$
$$\partial_x\partial_r^k A=\cos\beta\partial_r^{k+1}A,\quad \partial_x\partial_\beta^l B=-\frac{\sin\beta}{r}\partial_\beta^{l+1}B.$$

Using $\partial_x F_{i,j}=F_{i+1,j}$ and comparing the above formulas and the ansatz (C.14), we get

$$C_{i+1,j,k,l}=(k-i-j)\cos\beta C_{i,j,k,l}-\sin\beta\partial_\beta C_{i,j,k,l}+\cos\beta C_{i,j,k-1,l}-\sin\beta C_{i,j,k,l-1} \tag{C.16}$$

for $k\le i+j$. Similarly, for $D=\partial_y$, plugging the identities

$$\partial_y(C_{i,j,k,l}(\beta)r^{-i-j+k})=\cos\beta\partial_\beta C_{i,j,k,l}r^{-i-j-1+k}+(k-i-j)\sin(\beta)C_{i,j,k,l}r^{-i-j-1+k},$$
$$\partial_y\partial_r^k A=\sin\beta\partial_r^{k+1}A,\quad \partial_y\partial_\beta^l B=\frac{\cos\beta}{r}\partial_\beta^{l+1}B$$

into (C.15) and then comparing (C.14) and (C.15), we get

$$C_{i,j+1,k,l}=(k-i-j)\sin\beta C_{i,j,k,l}+\cos\beta\partial_\beta C_{i,j,k,l}+\sin\beta C_{i,j,k-1,l}+\cos\beta C_{i,j,k,l-1}. \tag{C.17}$$

The based case is given by

$$F_{0,0}=A(r)g(\beta),\quad C_{0,0,0,0}=1.$$

Using induction and the above recursive formulas, we can derive $C_{i,j,k,l}(\beta)$ in (C.14).

C.3.2. Estimates of $F_{i,j}$. To estimate $F_{i,j}$, using (C.14) and the triangle inequality, we only need to estimate $\partial_r^k A, \partial_\beta^l B(\beta)$, and $C_{i,j,k,l}(\beta)$. In our case, $B(\beta)$ is piecewise polynomials, whose estimates follow the method in Appendix (C.2.1). Function $A(r)$ is some explicit function, which will be constructed and estimated in section D.1. To estimate $C_{i,j,k,l}(\beta)$ on $\beta \in [\beta_1, \beta_2]$, we use the second order estimate in (C.12) and the induction ideas in section C.2.1. We can evaluate $C_{i,j,k,l}$ using its exact formula. It remains to bound $\partial_\beta^2 C_{i,j,k,l}$.

An important observation from (C.16), (C.15) is that $C_{i,j,k,l}$ is a polynomial on $\sin\beta$ and $\cos\beta$ with degree less than $i+j$, which can be proved easily using induction. In particular, we can write $C_{i,j,k,l}$ as follows:

$$C_{i,j,k,l} = \sum_{0\le k\le n} a_k \sin(k\beta) + b_k \cos(k\beta),$$
$$f \triangleq \partial_\beta^2 C_{i,j,k,l} = \sum_{1\le k\le n} c_k \sin(k\beta) + d_k \cos(k\beta), \quad n = i+j,$$

for some $a_k, b_k, c_k, d_k \in \mathbb{R}$. It is easy to see that $C_{i,j,k,l}$ is either odd or even in β depending on $j-l$, which implies $c_k \equiv 0$ or $d_k \equiv 0$. Using the Cauchy–Schwarz inequality, we get

$$||f||_\infty \le \sum_{1\le k\le n} (|c_k| + |d_k|) \le \left(n \sum_{k\le n} (c_k^2 + d_k^2) \right)^{1/2} = \left(\frac{n}{\pi} \int_0^{2\pi} f^2 \right)^{1/2},$$

where we have used the orthogonality of $\sin kx, \cos kx$ and $||f||_{L^2}^2 = \pi \sum_{k\le n}(c_k^2 + d_k^2)$ in the last equality. It is easy to see that f^2 is again a polynomial in $\sin\beta, \cos\beta$ with degree $\le 2n$. We fix $M > 2n$. For any $0 \le k < M$, it is easy to obtain

$$\frac{1}{2\pi} \int_0^{2\pi} e^{ikx} = \frac{1}{M} \sum_{j=1}^{M} \exp\Big(i \frac{2kj}{M} \pi \Big) = \delta_{k0}.$$

Using the above identity, we establish

$$||g||_{L^2}^2 = \frac{2\pi}{M} \sum_{j=1}^{M} |g\Big(\frac{2j\pi}{M}\Big)|^2$$

for any polynomial g in $\sin\beta, \cos\beta$ with degree $< M/2$. Hence, we prove

$$||f||_\infty \le \left(\frac{2n}{M} \sum_{k=1}^{M} f^2 \left(\frac{2j\pi}{M} \right) \right)^{1/2}.$$

The advantage of the above estimate is that to obtain the sharp bound of $C_{i,j,k,l}$, we only need to evaluate $C_{i,j,k,l}, f = \partial_\beta^2 C_{i,j,k,l}$ on finite many points.

C.3.3. From polar coordinates to the Cartesian coordinate. We want to obtain the piecewise estimate of $F_{p,q} = \partial_x^p \partial_y^q (A(r)g(\beta))$ on $Q_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}], 1 \le i, j \le n$. First, we partition the (r,β) coordinate into $r_1 < r_2 < \cdots < r_{n_1}, 0 = \beta_0 < b_1 < \cdots < \beta_{n_2} = \frac{\pi}{2}$. Then we apply the methods in section C.3 to bound $F_{p,q}$ on $S_{ij} \triangleq [r_i, r_{i+1}] \times [\beta_j, \beta_{j+1}]$. We cover Q_{ij} by $S_{k,l}$ and transfer the bound from (r,β) coordinate to (x,y) coordinate

$$\max_{x\in Q_{ij}} |F_{p,q}(x)| \le \max_{S_{k,l}\cap Q_{ij}\neq\emptyset} ||F_{p,q}(r,\beta)||_{L^\infty(S_{k,l})}.$$

For $(r,\beta)\in Q_{i,j}$, we get

$$r\in[(x_i^2+y_j^2)^{1/2},\ (x_{i+1}^2+y_{j+1}^2)^{1/2}],\quad \beta\in\Big[\arctan\frac{y_j}{x_{i+1}},\ \arctan\frac{y_{j+1}}{x_i}\Big].$$

Therefore, we get the necessary conditions for $Q_{i,j}\cap S_{k,l}\neq\emptyset$:

$$x_{i+1}^2+y_{j+1}^2\geq r_k^2,\quad x_i^2+y_i^2\leq r_u^2,\quad \arctan\frac{y_{j+1}}{x_i}\geq\beta_l,\quad \arctan\frac{y_j}{x_{i+1}}\leq\beta_{l+1}.$$

Given $Q_{i,j}$, we maximize $||F_{p,q}||_{L^\infty(S_{k,l})}$ over (k,l) satisfying the above bounds to control $||F_{p,q}||_{L^\infty(Q_{i,j})}$.

C.4. Estimates of the residual error. Let $\chi_{\bar\varepsilon}=1+O(|x|^4)$ be the cutoff function in (D.6). First, we decompose the error of solving the Poisson equations $\bar\varepsilon=\bar\omega-(-\Delta)\bar\phi^N$ as follows:

(C.18)
$$\bar\varepsilon=\bar\varepsilon_1+\bar\varepsilon_2,\quad \bar\varepsilon_2=\bar\varepsilon_{xy}(0)\Delta\Big(\frac{x^3y}{2}\chi_{\bar\varepsilon}\Big),\quad \mathbf{u}(\bar\varepsilon_2)=\nabla^\perp(-\Delta)^{-1}\bar\varepsilon_2=\frac{1}{2}\bar\varepsilon_{xy}(0)\nabla^\perp(x^3y\chi_{\bar\varepsilon}),$$
$$\mathbf{u}(\bar\varepsilon)=\mathbf{u}(\bar\varepsilon_1)+\mathbf{u}(\bar\varepsilon_2)=\mathbf{u}_A(\bar\varepsilon_1)+(\hat{\mathbf{u}}(\bar\varepsilon_1)+\mathbf{u}(\bar\varepsilon_2))\triangleq\mathbf{u}_A(\bar\varepsilon_1)+\mathbf{u}_{loc}(\bar\varepsilon),$$

where $\hat{\mathbf{u}}$ is the approximation term for $\mathbf{u}$ defined in section 4.3 in Part I [13]. We perform the above correction near 0 so that $\bar\varepsilon_1=O(|x|^3)$ near 0. We perform a similar decomposition for $(\nabla\mathbf{u})_A$. Note that we do not have $\partial_{x_i}\mathbf{u}_A=(\partial_{x_i}\mathbf{u})_A$. Using the above decomposition and the notation (3.4), we can rewrite the residual error $\bar{\mathcal{F}}_i$ (2.14) with rank-one correction as follows:

$$\bar{\mathcal{F}}_i-D_i^2\bar{\mathcal{F}}_i(0)f_{\chi,i}=\bar{\mathcal{F}}_{loc,i}+\mathcal{B}_{op,i}((\mathbf{u}_A(\bar\varepsilon_1),(\nabla\mathbf{u})_A(\bar\varepsilon_1)),\bar W),$$

where $D^2=(\partial_{xy},\partial_{xy},\partial_x^2)$ is defined in (3.23) and $\bar{\mathcal{F}}_{loc,i}$ is defined below in (C.19). Since $\mathbf{u}_A(\bar\varepsilon_1)=O(|x|^3),(\nabla\mathbf{u})_A(\bar\varepsilon_1)=O(|x|^2)$ (see section 4.3 in Part I [13] for these properties of $\mathbf{u}_A=\mathbf{u}-\hat{\mathbf{u}}$), from (3.4) and (C.18), we get

$$\mathcal{B}_{op,i}((\mathbf{u}_A(\bar\varepsilon_1),(\nabla\mathbf{u})_A(\bar\varepsilon_1)),\bar W)=O(|x|^3),\quad u_x(\bar\varepsilon_2)(0)=0,\quad u_{x,A}(\bar\varepsilon_1)(0)=0.$$

Using these properties of $\mathcal{B}_{op,i}$, we define $\bar{\mathcal{F}}_{loc,i}$ as follows:

(C.19)
$$\bar{\mathcal{F}}_{loc,i}=II_i-D_i^2II_i(0)f_{\chi,i},\quad II_i=\bar{\mathcal{F}}_i-\mathcal{B}_{op,i}((\mathbf{u}_A(\bar\varepsilon_1),(\nabla\mathbf{u})_A(\bar\varepsilon_1)),\bar W),$$
$$\mathbf{u}(\bar\omega)=\bar{\mathbf{u}}=\bar{\mathbf{u}}^N+\mathbf{u}_{loc}(\bar\varepsilon)+\mathbf{u}_A(\bar\varepsilon_1),\ \bar c_\omega=\bar c_\omega^N+u_x(\bar\varepsilon_1)(0),\ \bar c_\omega^N\triangleq\frac{\bar c_l}{2}+\bar u_x^N(0),$$
$$c_\omega(\bar\varepsilon_1)\triangleq u_x(\bar\varepsilon_1)(0),$$
$$II_1=-(\bar c_l x+\bar{\mathbf{u}}^N+\mathbf{u}_{loc}(\bar\varepsilon))\cdot\nabla\bar\omega+\bar\theta_x+(\bar c_\omega^N+\bar c_\omega(\bar\varepsilon_1))\bar\omega,$$
$$II_2=-(\bar c_l x+\bar{\mathbf{u}}^N+\mathbf{u}_{loc}(\bar\varepsilon))\cdot\nabla\bar\theta_x+2(\bar c_\omega^N+\bar c_\omega(\bar\varepsilon_1))\bar\theta_x-(\bar u_x^N+u_{x,loc}(\bar\varepsilon))\bar\theta_x$$
$$-(\bar v_x^N+v_{x,loc}(\bar\varepsilon))\bar\theta_y,$$
$$II_2=-(\bar c_l x+\bar{\mathbf{u}}^N+\mathbf{u}_{loc}(\bar\varepsilon))\cdot\nabla\bar\theta_y+2(\bar c_\omega^N+c_\omega(\bar\varepsilon_1))\bar\theta_y-(\bar u_y^N+u_{y,loc}(\bar\varepsilon))\bar\theta_x$$
$$-(\bar v_y^N+v_{y,loc}(\bar\varepsilon))\bar\theta_y,$$

where $f_{\chi,i}$ is defined in (D.6), and we have used $\bar c_\theta=\bar c_l+2\bar c_\omega$ (2.14), (2.11) for $\bar c_\omega$. The above decomposition is essentially the same as (3.12). We apply the functional inequalities in section 4 to estimate the nonlocal terms $\mathbf{u}_A(\bar\varepsilon_1),(\nabla\mathbf{u})_A(\bar\varepsilon_1)$, and combine the estimate of $\mathcal{B}_{op,i}((\mathbf{u}_A,(\nabla\mathbf{u})_A),\bar W)$ with the energy estimate. See section 5.8

in Part I [13] for more details about the decompositions and estimates. The terms II_i depend on the profile $\bar{\omega}, \bar{\theta}, \bar{\varepsilon}$ locally. Using the decomposition (C.18), we can further decompose the above II_i as follows:

$$II_i = II_i^N + II_i(\bar{\varepsilon}_1) + II_i(\bar{\varepsilon}_2),\ II_i(\bar{\varepsilon}_1) = \mathcal{B}_{op,i}(\hat{\mathbf{u}}(\bar{\varepsilon}_1), \widehat{\nabla \mathbf{u}}(\bar{\varepsilon}_1), \bar{W}),$$
$$II_i(\bar{\varepsilon}_2) = \mathcal{B}_{op,i}(\mathbf{u}(\bar{\varepsilon}_2), \nabla \mathbf{u}(\bar{\varepsilon}_2), \bar{W}),$$

where II_i^N contain the terms in II_i except the $u_{loc}, u(\bar{\varepsilon}_1)$ terms.

For $\hat{\mathbf{u}}(\bar{\varepsilon}_1)$, it is a finite rank operator on $\bar{\varepsilon}_1$, and we can write it as

$$\hat{\mathbf{u}}(\bar{\varepsilon}_1) = \sum_{i=1}^{n} a_i(\bar{\varepsilon}_1)\bar{g}_i(x) \triangleq C_{\mathbf{u}0}(x)u_x(\bar{\varepsilon}_1)(0) + \tilde{\hat{\mathbf{u}}}(\bar{\varepsilon}_1), \quad a_i(\bar{\varepsilon}_1) = \int_{\mathbb{R}_2^{++}} \bar{\varepsilon}_1(y)q_i(y)dy,$$

for some functions $\bar{g}_i(x)$ and $q_i(y)$, where $C_{\mathbf{u}0}(x)$ is given in (4.5), and $\tilde{\hat{\mathbf{u}}}(\bar{\varepsilon}_1)$ denotes other modes with $O(|x|^3)$ vanishing order near 0. See section 4.3 in [13] for a definition. We can obtain more regular estimates, e.g., C^3 estimates, of $\hat{\mathbf{u}}(\varepsilon_1)$ since $\bar{g}_1(x)$ is smooth. Similarly, we decompose $\widehat{\nabla \mathbf{u}}(\bar{\varepsilon}_1)$. We obtain piecewise estimates of $\partial_x^i \partial_y^j \bar{\varepsilon}_1, i+j \le 1$ following the methods in section 3.6 and section 8 in the supplementary material (supplement.pdf [local/web 1.43MB]) and then the above integrals on $\bar{\varepsilon}_1$. The main term in $\hat{\mathbf{u}}(\bar{\varepsilon}_1)$ is $C_{u0}u_x(0)$ with

$$\begin{aligned} u_x(\bar{\varepsilon}_1)(0) &= u_x(\bar{\varepsilon})(0) = -\frac{4}{\pi}\int_{\mathbb{R}_2^{++}} \bar{\varepsilon}(y)\frac{y_1y_2}{|y|^4}dy, \\ u_x(\bar{\varepsilon}_2)(0) &= -\varepsilon_{xy}(0)/2 \cdot \partial_y(x^3 y \chi_{\bar{\varepsilon}})|_{(0,0)} = 0. \end{aligned} \tag{C.20}$$

Since the kernel $\frac{y_1y_2}{|y|^4}$ has a slow decay for large $|y|$ (not in L^1), we need to estimate $u_x(\bar{\varepsilon})(0)$ carefully, using Simpson's rule. See section 6.4.2 in the supplementary material (supplement.pdf [local/web 1.43MB]).

Using the above decomposition, we further decompose $\hat{\mathbf{u}}(\bar{\varepsilon}_1)$

$$II_i(\bar{\varepsilon}_1) = u_x(\bar{\varepsilon})(0)\mathcal{B}_{op,i}\Big(C_{\mathbf{u}0}(x), C_{\nabla \mathbf{u}0}(x), \bar{W}\Big) + \mathcal{B}_{op,i}(\tilde{\hat{\mathbf{u}}}, \widetilde{\widehat{\nabla \mathbf{u}}}, \bar{W}) \triangleq II_{i,M}(\bar{\varepsilon}_1) + II_{i,R}(\bar{\varepsilon}_1).$$

Since D_i^2 is linear, we estimate each term $g_i - D_i^2 g_i(0) f_{\chi,i}$ for $g_i = II_{i,M}(\bar{\varepsilon}_1), II_{i,R}(\bar{\varepsilon}_1)$, $II_i^N, II_i(\bar{\varepsilon}_2)$ to bound $\mathcal{F}_{loc,i}$. To estimate $II_{i,R}$, since $\tilde{\hat{\mathbf{u}}}(\bar{\varepsilon}_1) = O(|x|^3)$ near 0, (see section 4.3 in [13]), we get $D_i^2 II_{i,R}(\bar{\varepsilon}_1) = O(|x|^3)$ and estimate

$$\tilde{\hat{\mathbf{u}}}(\bar{\varepsilon}_1)\rho_{10},\ \partial_i\tilde{\hat{\mathbf{u}}}(\bar{\varepsilon}_1)\rho_{20},\ \widetilde{\widehat{\nabla \mathbf{u}}}(\bar{\varepsilon}_1)\rho_{20},\ \partial_i\widetilde{\widehat{\nabla \mathbf{u}}}(\bar{\varepsilon}_1)\rho_3,\ \rho_4\tilde{\hat{\mathbf{u}}}(\bar{\varepsilon}_1)$$

for ρ_{i0} (A.2) with $\rho_{i0} \sim |x|^{-4+i}, i \le 3$, near 0 using the C^3 bounds of $\tilde{\hat{\mathbf{u}}}, \widetilde{\widehat{\nabla \mathbf{u}}}$. Note that $\partial_i\tilde{\hat{\mathbf{u}}} \neq \widetilde{\widehat{\partial_i \mathbf{u}}}$. The former is the derivative of $\tilde{\hat{\mathbf{u}}}$, and the latter is the approximation term for $\partial_i \mathbf{u}$. With the above weighted estimate, we can bound a typical term, e.g., $\widetilde{\widehat{u_x}}\bar{\theta}_x\varphi_2$ in $II_{i,R}(\bar{\varepsilon}_1)\varphi_2$, as follows:

$$\begin{aligned} \widetilde{\widehat{u_x}}\bar{\theta}_x\varphi_2 &= \widetilde{\widehat{u_x}}\rho_{20} \cdot \Big(\bar{\theta}_x\frac{\varphi_2}{\rho_{20}}\Big),\ \partial_x(\widetilde{\widehat{u_x}}\bar{\theta}_x)\rho = (\partial_x\widetilde{\widehat{u_x}}\bar{\theta}_x + \widetilde{\widehat{u_x}}\partial_x\bar{\theta}_x)\rho \\ &= \partial_x\widetilde{\widehat{u_x}}\rho_3 \cdot \frac{\bar{\theta}_x\rho}{\rho_3} + \widetilde{\widehat{u_x}}\rho_{20} \cdot \frac{\partial_x\bar{\theta}_x\rho}{\rho_{20}}, \end{aligned}$$

where φ_2 is given in (A.2). Each term A, B in the above products $A \cdot B$ is regular and we estimate each term and then the product to bound weighted L^∞ and C^1 norm of $II_{i,R}(\bar{\varepsilon}_1)$.

The remaining part in $II_i^N, II_{i,M}(\bar{\varepsilon}_1), II_i(\bar{\varepsilon}_2)$ depends on $(\bar{\phi}^N, \bar{\omega}, \bar{\theta})$ locally and they are given functions. To estimate the weighted L^∞ and $C^{1/2}$ norms of $g_i - D_i^2 g_i(0) f_{\chi,i} = O(|x|^3)$ with $g = II_{i,M}(\bar{\varepsilon}_1), II_i(\bar{\varepsilon})$, we follow the methods in sections 3.6, 3.7 with $\partial_t \bar{\omega} = \partial_t \bar{\theta} = 0$.

Estimate in the far-field. Since $\bar{\omega}, \bar{\theta}$ are supported globally, we need to estimate the error in the far-field. Recall the formulas of $\bar{\omega}, \bar{\omega}_1, \bar{\theta}, \bar{\theta}_1$ from (C.1). We consider $|x|_\infty \geq R_1 \geq 10^{12} > 10a_2$ beyond the support of $\bar{\omega}_2, \bar{\theta}_2, \bar{\phi}_2^N, \bar{\phi}_3^N, \bar{\phi}_{cor}^N$ (C.1) so that $\chi(r) = 1$ (D.4) and

$$\bar{\omega} = \bar{\omega}_1 = \bar{g}_1(\beta) r^{\bar{\alpha}_1}, \quad \bar{\theta} = \bar{\theta}_1 = r^{1+2\bar{\alpha}_1} \bar{g}_2(\beta), \quad \bar{\phi}^N = \bar{\phi}_1^N = r^{2+\bar{\alpha}_1} \bar{f}(\beta).$$

We estimate the angular derivatives of $f(\beta), g_i(\beta)$ using the methods in section C.2.1. Using the above representation, $x \cdot \nabla r^\beta = r \partial_r r^\beta = \beta r^\beta$, $x \cdot \nabla(\partial \bar{\theta}_1) = 2\bar{\alpha}_1 (\partial \bar{\theta}_1), x \cdot \nabla \bar{\omega}_1 = \bar{\alpha}_1 \bar{\omega}_1$, $\bar{c}_\omega = \bar{c}_\omega^N + \bar{c}_\omega^{\bar{\varepsilon}}$ (3.11), and separating $\mathbf{u}^N$ and $\mathbf{u}_{loc}$ in (C.19), for $|x|_\infty \geq 10^{12}$, we obtain

$$\begin{aligned}
\bar{\mathcal{F}}_{loc,1} &= \Big((\bar{c}_\omega^N - \bar{c}_l \bar{\alpha}_1) \bar{\omega}_1 - \bar{\mathbf{u}}^N \cdot \nabla \bar{\omega}_1 + \bar{\theta}_{1,x} \Big) + \bar{c}_\omega^{\varepsilon} \bar{\omega}_1 - \mathbf{u}_{loc} \cdot \nabla \bar{\omega}_1 \triangleq I_{11} + I_{12}, \\
\bar{\mathcal{F}}_{loc,2} &= \Big((2\bar{c}_\omega^N - 2\bar{c}_l \bar{\alpha}_1) \bar{\theta}_{1,x} - \partial_x (\bar{\mathbf{u}}^N \cdot \nabla \bar{\theta}_1) \Big) \\
&\quad + 2\bar{c}_\omega^{\varepsilon} \bar{\theta}_{1,x} - \mathbf{u}_{loc} \cdot \nabla \bar{\theta}_{1,x} - u_{x,loc} \bar{\theta}_x - v_{x,loc} \bar{\theta}_y \triangleq I_{21} + I_{22}, \\
\bar{\mathcal{F}}_{loc,3} &= \Big((2\bar{c}_\omega^N - 2\bar{c}_l \bar{\alpha}_1) \bar{\theta}_{1,y} - \partial_x (\bar{\mathbf{u}}^N \cdot \nabla \bar{\theta}_1) \Big) \\
&\quad + 2\bar{c}_\omega^{\varepsilon} \bar{\theta}_{1,y} - \mathbf{u}_{loc} \cdot \nabla \bar{\theta}_{1,y} - u_{y,loc} \bar{\theta}_x - v_{y,loc} \bar{\theta}_y \triangleq I_{31} + I_{32},
\end{aligned}$$

where we have simplified $\mathbf{u}_{loc}(\bar{\varepsilon})$ as $\mathbf{u}_{loc}$ and used $f_{\chi,i} = 0$ (D.6), $\bar{\mathcal{F}}_{loc,i} = II_i$ (C.19) since $f_{\chi,j}$ is supported near 0. The terms I_{11}, I_{21}, I_{31} are local with the form $r^\gamma q(\beta)$ for some angular function q and decay rate γ. We estimate their piecewise L^∞ and derivative bounds using (B.14). From our choice of $\bar{\alpha}_1$ (C.1), $\bar{c}_\omega^N - \bar{c}_l \bar{\alpha}_1$ is very small. Thus the first term in I_{11}, I_{21}, I_{31} is small. The second term in I_{11}, I_{21}, I_{31} has faster decay rates $r^{2\bar{\alpha}_1}, r^{3\bar{\alpha}_1}$ and is also very small.

Estimate of the velocity approximation. From (C.18), since $\bar{\varepsilon}_2$ is supported near 0, we get $\mathbf{u}_{loc} = \hat{\mathbf{u}}(\varepsilon_1)$. For I_{j2} in the above decomposition in the far-field, it remains to estimate

(C.21)
$$\bar{c}_\omega^e \bar{\omega} - \hat{\mathbf{u}}(\bar{\varepsilon}_1) \cdot \nabla \bar{\omega}, \ 2\bar{c}_\omega^e \bar{\theta}_x - \hat{\mathbf{u}}_x(\bar{\varepsilon}_1) \cdot \nabla \bar{\theta} - \hat{\mathbf{u}}(\bar{\varepsilon}_1) \cdot \nabla \bar{\theta}_x, \ 2\bar{c}_\omega^e \bar{\theta}_y - \hat{\mathbf{u}}_y(\bar{\varepsilon}_1) \cdot \nabla \bar{\theta} - \hat{\mathbf{u}}(\bar{\varepsilon}_1) \cdot \nabla \bar{\theta}_x.$$

Note that $c_\omega(\bar{\varepsilon}_1) = c_\omega(\bar{\varepsilon})$ (C.20) and $c_\omega(\bar{\varepsilon}) = \bar{c}_\omega^e$ in our notation. For any $a \in \mathbb{R}$, we estimate

$$A(f,g) = ag - \hat{\mathbf{u}}(f) \cdot \nabla g, \quad B_i(f,g) = 2a\partial_i g - \hat{\mathbf{u}}(f) \cdot \nabla \partial_i g - \widehat{\partial_i \mathbf{u}}(f) \cdot \nabla g, \ i = 1, 2,$$

for $|x|_\infty \geq R_1$. From sections 4.3.2–4.3.3 in Part I [13], for $|x|_\infty \geq R_1$, $\hat{\mathbf{u}}, \widehat{\nabla \mathbf{u}}$ reduce to

$$\begin{aligned}
&\hat{u}(f) = x_1 I_{far}(f), \quad \hat{v}(f) = -x_2 I_{far}(f), \quad \widehat{\partial_1 u(f)} = I_{far}(f), \quad \widehat{\partial_2 v(f)} = -I_{far}(f), \\
&\widehat{\partial_2 u(f)} = \widehat{\partial_1 v(f)} = 0, \quad I_{far}(f) \triangleq -\frac{4}{\pi} \int_{\max(y_1, y_2) \geq R_n} \frac{y_1 y_2}{|y|^4} \omega(y) dy,
\end{aligned}$$

where $R_n = 1024 \cdot 64 h_x$ is the largest threshold. Denote $b = I_{far}(f)$. A direct calculation yields

$$\begin{aligned} A(f,g) &= (a-b)g + b(g - x_1\partial_1 g + x_2\partial_2 g), \\ B_1(f,g) &= 2a\partial_1 g - b\partial_1 g - bx_1\partial_{11} g + bx_2\partial_{12} g \\ &= (2a-2b)\partial_1 g + b(\partial_1 g - x_1\partial_{11} g + x_2\partial_{12} g), \\ B_2(f,g) &= 2a\partial_2 g + b\partial_2 g - bx_1\partial_{12} g + bx_2\partial_{22} g \\ &= (2a-2b)\partial_2 g + b(3\partial_1 g - x_1\partial_{11} g + x_2\partial_{12} g). \end{aligned} \tag{C.22}$$

Therefore, we only need to bound the functions following section C.2, e.g., $g - x_1\partial_1 g + x_2\partial_2 g$ and g, and the functional $b(f)$ and a. We apply these estimates for (C.21) with $a = \bar{c}_\omega^e, f = \bar{\varepsilon}_1, g = \bar{\omega}, \bar{\theta}$.

Appendix D. Estimate of explicit functions. In this section, we estimate the derivatives of several explicit or semiexplicit functions using induction, including several cutoff functions used in the estimates and the weight in the stream function (C.8).

D.1. Estimate of the radial functions.

D.1.1. Estimate of the cutoff function. We estimate the derivatives of the cutoff function

$$\chi_e(x) = \Big(1 + \exp\Big(\frac{1}{x} + \frac{1}{x-1}\Big)\Big)^{-1}, \tag{D.1}$$

where e is short for *exponential.* In our verification, it involves high order derivatives of χ_e. Although χ_e is explicit, its formula is complicated and is difficult to estimate. Instead, we use the structure of $\partial_x^i \chi_e$ and induction to estimate $\partial_x^i \chi_e$. Denote

$$p(x) = \frac{1}{x} + \frac{1}{x-1}, \quad f = \frac{1}{1+x}, \quad \chi_e = f(e^p).$$

First, we use induction to derive

$$d_x^k \chi_e = \sum_{i=1}^{k} (\partial^i f)(e^p) e^{ip} Q_{k,i}(x),$$

where $Q_{k,i} = 0$ for $i > k, i < 0$. A direct calculation yields

$$\begin{aligned} \partial \sum_{i=1}^{k} \partial^i f e^{ip} Q_{k,i}(x) &= \sum_{i=1}^{k} (\partial^{i+1} f)(e^p) \cdot p' e^p e^{ip} Q_{k,i} + (\partial^i f)\partial_x (e^{ip} Q_{k,i}) \\ &= \sum_{i=1}^{k} (\partial^{i+1} f)(e^p) \cdot e^{(i+1)p} p' Q_{k,i} + (\partial^i f) e^{ip} (ip' Q_{k,i} + Q'_{k,i}). \end{aligned}$$

Comparing the above two equations, we derive

$$Q_{k+1,i} = p' Q_{k,i-1} + ip' Q_{k,i} + Q'_{k,i}.$$

The first few terms in $Q_{k,i}$ are given by

$$Q_{0,0} = 1, \quad Q_{1,1} = p', \quad Q_{1,0} = 0.$$

It is not difficult to see that $Q_{k,i}$ is a polynomial of $\partial_x^j p, j \le k$ with nonnegative coefficients. We derive the expression of $Q_{k,i}$ in terms of $\partial_x^j p, j \le k$ symbolically. Thus, using the triangle inequality, we only need to bound $\partial_x^j p$. We have

$$|\partial_x^n p(x)| = n!|x^{-n-1} + (x-1)^{-n-1}| \le n!(|z|^{-n-1} + 2^{n+1}), \quad z = \min(|x|, |1-x|).$$

If n is even, x^{-n-1} and $(x-1)^{-n-1}$ have a different sign, and we get a better estimate,

$$|\partial_x^n p(x)| \le n! \max(|x|^{-n-1}, |x-1|^{-n-1}) = n! \cdot z^{-n-1}.$$

Substituting the above bounds into the formula of $Q_{k,i}$, we can obtain the upper bound $Q_{k,i}^u(x)$ for $Q_{k,i}(x)$, which is a polynomial of z^{-1} with positive coefficient. Since each term in $Q_{k,i}$ is given by $c_{i_1,i_2,..,i_m} \prod_{j=1}^m \partial_x^{i_j} p$ with $\sum i_j = k$, the above estimate implies

$$\left| c_{i_1,i_2,..,i_m} \prod_{j=1}^m \partial_x^{i_j} p \right| \le c_{i_1,i_2,..,i_m} \prod_{j=1}^m i_j!(|z|^{-i_j-1} + 2^{i_j+1}).$$

Since $m \le k$, the highest order of z^{-1} in the upper bound is bounded by $2k$. Thus, we obtain that $Q_{k,i}^u$ is a polynomial in z^{-1} with $\deg Q_{k,i}^u \le 2k$. Next, we bound

$$|e^{ip} Q_{k,i}| \le e^{ip} Q_{k,i}^u.$$

For $k \le 20, x \ge 1 - \frac{1}{2k} \ge \frac{1}{2}, z^{-1} = |x-1|^{-1} \ge 2k$, a direct calculation implies that $e^{ip(x)} Q_{k,i}^u(x)$ is decreasing. In fact, for $l \le 2k$, we have $z = |x-1| = 1-x$ and

$$\begin{aligned} \partial_x(\exp(ip(x))(1-x)^{-l}) &= \exp(ip(x))(ip'(1-x)^{-l} + l(1-x)^{-l-1}) \\ &= \exp(ip(x))\Big(-\frac{i}{x^2} - \frac{i}{(x-1)^2} + l(1-x)^{-1} \Big)(1-x)^{-l} \le 0. \end{aligned}$$

In the last inequality, we have used $-\frac{i}{1-x} + l \le -2ki + 2k \le 0$.

Note that $|(\partial_x^i f)(e^p)| = i!|(1+e^p)^{-i-1}| \le i!$. Thus, for $x \in [x_l, x_u]$ with x_l close to 1, we get

$$\begin{aligned} |\partial_x^k \chi_e(x)| &\le \sum_{i=1}^k |(\partial^i f)(e^p)| e^{ip(x)} Q_{k,i}^u(x) \le \sum_{i=1}^k i! \frac{e^{ip(x)}}{(1+e^p)^{i+1}} Q_{k,i}^u(x) \\ &\le \sum_{i=1}^k i! e^{ip(x_l)} Q_{k,i}^u(x_l). \end{aligned}$$

For x away from 1, we use monotonicities of p, Q^u and the above estimate to estimate piecewise bounds of $\partial_x^k \chi_e(x)$. Using the above derivatives bound, the symbolic formula of $\partial_x^k \chi_e$, and the refined second order estimate in section C.2.1, we can obtain sharp bounds for $\partial_x^k \chi_e$. Note that we only apply the above estimate to $k \le 15$.

D.1.2. Estimate of polynomial decay functions. For cutoff function $\chi_e(\frac{|x|-a}{b})$ based on the exponential cutoff function (D.1), it has rapid change from $|x| \le a$ to $|x| \ge a+b$, which is not very smooth in the computational domain if there is not enough mesh for x with $a \le |x| \le b$. We apply these cutoff functions to the far-field, e.g., $|x| \ge 10$, where the mesh is relatively sparse. Thus, we need another function similar to a cutoff function that has a slower change than the exponential cutoff function. We consider

$$\chi(x) = \frac{x^7}{(1+x^2)^{7/2}}, \quad x \in \mathbb{R}_+, \tag{D.2}$$

and will use its rescaled version, e.g., $\chi(\frac{x-a}{b})$, in our verification.

First, we use induction to derive

$$\partial_x^k \chi = \frac{p_k(x)}{(1+x^2)^{7/2+k}}, \quad p_0 = x^7,$$

where $p_k(x)$ is a polynomial. A direct calculation yields

$$\partial_x^{k+1} \chi = \frac{p_k'(x)(1+x^2) - (\frac{7}{2}+k)\cdot 2x p_k(x)}{(1+x^2)^{7/2+k+1}}.$$

Comparing the above two formulas, we get

$$p_{k+1} = p_k'(1+x^2) - (7+2k)x p_k(x).$$

The first few terms are given by $p_0 = x^7, p_1 = 7x^6$. Using the recursive formula and $\deg p_1 = 6$, we get

$$\deg p_{k+1} \le \deg p_k + 1, \quad \deg p_k \le k+5, \quad k \ge 1. \tag{D.3}$$

Since p_k is a polynomial, the above recursive formula shows that p_{k+1} is also a polynomial.

To estimate $\partial_x^k \chi$, we decompose p_k into the positive and the negative parts. Suppose that $p_k = \sum_i a_i x^i$. We have

$$p_k = p_k^+ - p_k^-, \quad p_k^+ = \sum a_i^+ x^i, \quad p_k^- = \sum a_i^- x^i.$$

For $x \ge 0$, p_k^+, p_k^- are increasing. Thus, for $x \in [x_l, x_u]$, we get

$$|\partial_x^k \chi| \le \frac{\max(p_k^+(x_u) - p_k^-(x_l), p_k^-(x_u) - p_k^+(x_l))}{(1+x_l^2)^{7/2+k}}.$$

Next, we estimate $\partial_x^k \chi$ for large x. For $x \ge 2, k \ge 1$ and any polynomial $q(x)$ with nonnegative coefficients and $\deg q \le k+5$, we get

$$xq' \le (k+5)q, \quad \frac{q'(1+x^2)}{(7+2k)xq} \le \frac{(1+x^2)(k+5)}{(7+2k)x^2} \le \frac{5(k+5)}{4(7+2k)} < 1.$$

The first inequality follows by comparing the coefficients of xq' and $(k+5)q$, which are nonnegative. It follows that

$$\partial_x \frac{q}{(1+x^2)^{7/2+k}} = \frac{q'(1+x^2) - (7/2+k)2xq}{(1+x^2)^{7/2+k+1}} \le 0, \quad k \ge 1,\ x \ge 2.$$

Thus $\frac{q}{(1+x^2)^{7/2+k}}$ is decreasing. For $k \ge 1$ and $x \ge x_l \ge 2$, using (D.3) and the monotonicity, we get

$$|\partial_x^k(x)| \le \frac{p_k^+(x) + p_k^-(x)}{(1+x^2)^{7/2+k}} \le \frac{p_k^+(x_l) + p_k^-(x_l)}{(1+x_l^2)^{7/2+k}}.$$

For $k = 0$, the estimate is trivial: $\chi(x) \le 1$. Using these higher order derivative bounds, we can use the discrete values of $\partial_x^k \chi$ and the bound for $\partial_x^{k+2} \chi$ to obtain sharp bounds of $\partial_x^k \chi$.

Note that $\chi_1(x-a) = \frac{(x-a)_+^7}{(1+(x-a)^2)^{7/2}}$ is only $C^{6,1}$. Suppose that $a \in [x_l, x_u]$. Since χ_1 is smooth on $x \le a$ and on $x \ge a$, we can still use a first order estimate to estimate $\partial_x^k \chi_1$ as follows:

$$|\partial_x^k \chi_1(x)| \le \max_{\alpha \in \{l,u\}} |\partial_x^k \chi_1(x_\alpha)| + \max(||\partial_x^{k+1}\chi_1||_{L^\infty[x_l,a]}||\partial_x^{k+1}\chi_1||_{L^\infty[a,x_u]})|x_u - x_l|.$$

D.1.3. Radial cutoff function. Now, we construct the radial cutoff functions for the far-field approximation terms of ω and ϕ as follows:

$$\begin{aligned}(\text{D.4})\quad &\chi(r)=\chi_1(1-\chi_2)+\chi_2,\quad \chi_1(r)=\chi_{rati}\left(\frac{r-a_1}{l_1^{1/2}}\right),\quad \chi_2(r)=\chi_{exp}\Big(\frac{r-a_2}{9a_2}\Big),\\ &a_1=10,\quad l_1=50000,\quad a_2=10^5,\end{aligned}$$

where χ_{exp} and χ_{rati} are defined in (D.1) and (D.2), respectively. Using the estimates of χ_{rati},χ_{exp} established in the last two sections, the Leibniz rule (A.6), and (C.12), we can evaluate χ on the grid points and estimate its derivative bounds.

D.2. Cutoff function near the origin. For the cutoff function $\kappa(x)$ used in section 3, we choose it as follows:

$$(\text{D.5})\qquad \kappa(x;a,b)=\kappa_1\Big(\frac{x}{a}\Big)\Big(1-\chi_e\Big(\frac{x}{b}\Big)\Big),\quad \kappa_1(x)=\frac{1}{1+x^4},\quad \kappa_*(x)=\kappa\Big(x;\frac{1}{3},\frac{3}{2}\Big),$$

where χ_e is the cutoff function chosen in (D.1). We mostly use the cutoff κ_*. Since $\chi_e(y)=1$ for $y\geq 1$ and $\chi_e(y)=0$ for $y\leq 0$, the above cutoff function is supported in $x\leq a_2$. Using Taylor expansion, we have the following properties for κ:

$$\kappa_1(x/a_1)=1+O(x^4),\quad \kappa(x)=1+O(x^4).$$

For the cutoff functions χ_{NF} in section 4.2.1 in Part I [13], $\chi_{\bar\varepsilon}$ in (C.18), and $\chi_{\hat\varepsilon}$ in (3.42), we choose

$$\begin{aligned}(\text{D.6})\quad &\chi_{\bar\varepsilon}(x,y)=\kappa(x;\nu_{\bar\varepsilon,1},\nu_{\bar\varepsilon,2})\kappa(y;\nu_{\bar\varepsilon,1},\nu_{\bar\varepsilon,2}),\quad \nu_{\bar\varepsilon,1}=1/192,\quad \nu_{\bar\varepsilon,2}=3/2,\\ &\chi_{\hat\varepsilon}(x,y)=\kappa_*(x)\kappa_*(y),\quad \chi_{NF}(x,y)=\kappa(x;2,10)\kappa(y;2,10),\\ &f_{\chi,1}=\Delta\Big(\frac{xy^3}{6}\chi_{NF}(x,y)\Big),\quad f_{\chi,2}=xy\chi_{NF}(x,y),\quad f_{\chi,3}=\frac{x^2}{2}\chi_{NF}(x,y).\end{aligned}$$

For the cutoff function in the stream function (C.1), we choose

$$(\text{D.7})\qquad \chi_\phi=\kappa_2\Big(\frac{x}{\nu_{4,1}}\Big)\Big(1-\chi_e\Big(\frac{x}{\nu_{4,2}}\Big)\Big),\quad \kappa_2(x)=\frac{1}{1+x^2},\quad \nu_{4,1}=2,\quad \nu_{4,2}=128.$$

For $\kappa_1(x),\kappa_2(x)$, we use induction to obtain

$$\partial_x^k\kappa_1(x)=\frac{P_k^+(x)-P_k^-(x)}{(1+x^4)^{k+1}},\quad \partial_x^k\kappa_2(x)=\frac{R_k^+(x)-R_k^-(x)}{(1+x^2)^{k+1}}$$

for some polynomials $P_k^\pm,R_k^\pm$ with nonnegative coefficients, and the same method as that in section D.1.2 to estimate the derivatives of $\partial_x^i\kappa_1(x)$. The estimate of κ_1 is simpler since κ_1 has a simpler form. Using the Leibniz rule (A.6) and the triangle inequality, we can obtain estimate $\partial_x^l\kappa_1(x)$ in $[a,b]$. Then we use these derivative estimates for $\partial_x^{l+2}\kappa_1(x)$, evaluate $\kappa(x;a_1,a_2)$ on the grid points, and then use (C.12) to obtain a sharp estimate of $\partial_x^l\kappa_1(x)$ on $[a,b]$. The same method applies to estimate κ_2,χ_ϕ.

For large x, e.g., $x\geq 100$, the above estimates can lead to a very large round off error. Instead, for $a\geq 2, a\in\mathbb{Z}_+$, we use the Taylor expansion

$$\begin{aligned}&F_a=\frac{1}{1+x^a}=\sum_{k\geq 0}(-1)^k x^{-a(k+1)},\quad \partial_x^i F_a=\sum_{k\geq 0}(-1)^{k+i}C_{i,k}x^{-a(k+1)-i},\\ &C_{i,k}=\prod_{0\leq j\leq i-1}(a(k+1)+j).\end{aligned}$$

We want to bound $|\partial_x^i F_a| \le C_{i,0}(1+C_\varepsilon)x^{-a-i}$ for $x \ge x_l = 100, i \le 20$. For $k \le 20$, we bound

$$C_{i,k}x^{-a(k+1)-i} \le C_{i,k}x_l^{-(a-1)k}x^{-a-i-k} \le C_{i,0}\varepsilon_1^{-a-i-k},\ \varepsilon_1 \triangleq \max_{i\le 20,k\le 20} x_l^{-(a-1)k}C_{i,k}C_{i,0}^{-1}.$$

For the tail part $k > 20$, we consider $G(k) = k\log x - i\log(1+k)$. Since $x > 21, i \le 20$, we get

$$\partial_k G = \log x - \frac{i}{1+k} \ge \log x - 1 > \log 4 - 1 > 0, \quad G(k) \ge G(21) = 21\log x - i\log 21 > 0.$$

It follows that $x^k > (1+k)^i$. Using $\frac{a(k+1)+j}{a+j} \le 1+k, C_{i,k} \le C_{i,0}(1+k)^i$, and $a \ge 2$, we further get

$$C_{i,k}x^{-a(k+1)-i} \le x^{-k-a-i}C_{i,k}x^{-k} \le x^{-k-a-i}C_{i,0}(1+k)^i x^{-k} \le C_{i,0}x^{-k-a-i},\ k > 20.$$

Combining the above estimates and $x \ge x_l > 10$, we obtain

$$|\partial_x^i F_a| \le C_{i,0}x^{-a-i}C_a,$$
$$C_a \le 1 + \varepsilon_1\sum_{k=1}^{20} x^{-k} + \sum_{k\ge 21} x^{-k} \le 1 + \frac{\varepsilon_1 x^{-1}}{1-x^{-1}} + \frac{x^{-21}}{1-x^{-1}} \le 1 + \frac{\varepsilon_1}{x_l - 1} + x_l^{-20}.$$

D.3. Estimate of $\rho_p(y)$. We estimate the weight $\rho_p(y)$ (C.9) in the representation of the stream function. Using symbolic computation, e.g., MATLAB or Mathematica, we get

$$\partial_x^9\rho_p(y) = \frac{f_2(y) - f_1(y)}{(g(y))^9}, \quad g(y) = 2 + 2y + y^2,$$
$$f_1 = 288y^2 + 672y^3 + 504y^4, \quad f_2 = 16 + 168y^6 + 72y^7 + 9y^8.$$

Since $f_1, f_2, g \ge 0$ are increasing in $y \ge 0$, for $y \in [y_l, y_u]$, we get

$$|\partial_x^9\rho_p(y)| \le \frac{\max(f_2(y_u) - f_1(y_l), f_1(y_u) - f_2(y_l))}{(g(y_l))^9}.$$

We have a trivial estimate similar to (C.12)

$$\max_{x\in I}|f(x)| \le \max(|f(x_l)|, |f(x_u)|) + \frac{h}{2}||f_x||_{L^\infty(I)}, \tag{D.8}$$

which is useful if we do not have a bound for f_{xx}.

Based on the above estimates, using the estimates (C.12), (D.8), ideas in section C.2.1, and evaluating ρ_p on some grid points, we can obtain piecewise sharp bounds for $\partial_x^i\rho_p$ for $i \le 8$.

Appendix E. Piecewise $C^{1/2}$ and Lipschitz estimates. In this section, we estimate the piecewise $C^{1/2}$ bound and Lipschitz bound for a function.

E.1. Hölder estimate of the functions. In the following two sections, we estimate the Hölder seminorms $[f]_{C_x^{1/2}}$ or $[f]_{C_y^{1/2}}$ of some function f, e.g., $f = (\partial_t - \mathcal{L})\widehat{W}$ in (3.28), based on the previous L^∞ estimates. We will develop two approaches.

Below, we will assume $x, y \in \mathbb{R}_2^{++}$ since our function $f(x)$ defined on $x \in \mathbb{R}_2^+ (x_2 \ge 0)$ is either even or odd in x_1 and we can reduce essentially all estimates to the case of

$\mathbb{R}^2_{++}$ using symmetry. Suppose that we have bounds for $\partial_x f, \partial_y f$, and f. First, we consider the $C_x^{1/2}$ estimate. For $x_1 < y_1$ and $x_2 = y_2$, we have

$$I = \frac{|f(x) - f(y)|}{|x-y|^{1/2}} \le |x-y|^{1/2} \frac{1}{|x-y|} \int_{x_1}^{y_1} |f_x(z_1, x_2)| dz_1.$$

We further bound the average of f_x piecewisely using the method in Appendix E.2 to obtain the first estimate. We have a second estimate

$$|I| = \Big| \int_{x_1}^{y_1} f_x(z_1, x_2) dz \Big| \cdot \frac{1}{|x-y|^{1/2}} \le ||f_x x^{1/2}||_\infty \int_{x_1}^{y_1} z_1^{-1/2} dz_1 \cdot \frac{1}{|x-y|^{1/2}}$$
$$\le ||f_x x^{1/2}||_\infty 2 \frac{y_1^{1/2} - x_1^{1/2}}{|x-y|^{1/2}} = ||f_x x^{1/2}||_\infty \frac{2\sqrt{y_1 - x_1}}{\sqrt{x_1} + \sqrt{y_1}}.$$

We also have a trivial L^∞ estimate

$$|I| \le ||f x_1^{-1/2}||_\infty \frac{x_1^{1/2} + y_1^{1/2}}{|x-y|^{1/2}}, \quad |I| \le ||f||_\infty \frac{2}{|x-y|^{1/2}}.$$

Similar L^∞ and Lipschitz estimates apply to $||f||_{C_y^{1/2}}$.

Near the origin, optimizing the above estimates, for $x_2 = y_2$, we obtain

$$\Big| \frac{f(x) - f(y)}{|x-y|^{1/2}} \Big| \le \min(||f_x x^{1/2}||_\infty 2t,\ ||f x_1^{-1/2}||_\infty t^{-1}), \quad t = \frac{\sqrt{y_1 - x_1}}{\sqrt{x_1} + \sqrt{y_1}}.$$

In the Y-direction, $x_1 = y_1, x_2 \le y_2$, and we use

$$I_Y = \Big| \frac{f(x) - f(y)}{|x-y|^{1/2}} \Big| \le \frac{1}{|x_2 - y_2|^{1/2}} \int_{x_2}^{y_2} |f_y(x_1, z_2)| |z|^{1/2} \cdot |z|^{-1/2} dz_2$$
$$\le ||f_y |x|^{1/2}||_\infty \frac{|x_2 - y_2|^{1/2}}{|x|^{1/2}} \triangleq At,$$
$$I_Y \le (|f(x)x_1| + |f(y)x_1|^{1/2}) \Big(\frac{x_1}{|x|} \Big)^{1/2} \cdot \frac{|x|^{1/2}}{|x_2 - y_2|^{1/2}} \triangleq Bt^{-1},$$
$$t \triangleq \frac{|x_2 - y_2|^{1/2}}{|x|^{1/2}},\ I_Y \le \min(At, Bt^{-1}).$$

Since $x_1 \le |x|$, A, B are not singular near $x = 0$. We derive the piecewise bounds for A, B and then optimize two estimates to estimate I_Y.

From the above estimates, to obtain a sharp Hölder estimate of f, we estimate the piecewise bounds of $f, f x_1^{-1/2}, f|x|^{-1/2}, f_x, f_y,\ f_x |x_1|^{1/2}, f_y |x|^{1/2}$, which are local quantities. These estimates can be established using the piecewise bounds of $\partial_x^i \partial_y^j f$ and the methods in section 8 in the supplementary material (supplement.pdf [local/web 1.43MB]).

E.1.1. The second approach of Hölder estimate. We develop an additional approach to estimate $I(f) = \frac{|f(x) - f(z)|}{|x-z|^{1/2}}$ that is sharper if $|x - z|$ is not small and f is smooth. We need the grid point values and derivative bounds of f.

We estimate $I(f) = \frac{|f(x) - f(z)|}{|x-z|^{1/2}}$ for $x \in [x_l, x_u], z \in [z_l, z_u]$. Denote by $\hat{f}$ the linear approximation of f with $\hat{f}(x_i) = f(x_i)$ on the grid point x_i. We have the following lemma.

LEMMA E.1. *Suppose that f is linear on $[x_l, x_u], [z_l, z_u]$ and $x_l \le x_u \le z_l \le z_u$. Then we have*

$$\max_{x\in[x_l,x_u],z\in[z_l,z_u]} \frac{|f(x)-f(z)|}{|x-z|^{1/2}} = \max_{\alpha,\beta\in\{l,u\}} \frac{|f(x_\alpha)-f(z_\beta)|}{|x_\alpha - z_\beta|^{1/2}}.$$

The above lemma shows that for the linear interpolation of f, the maximum of the Holder norm is achieved at the grid point.

Proof. Denote by M the right hand side in the lemma. Clearly, it suffices to prove that the left hand side is bounded by M. We fix $x \in [x_l, x_u], z \in [z_l, z_u]$. Suppose that

$$x = a_l x_l + a_u x_u, \quad z = b_l z_l + b_u z_u, a_u + a_l = 1, \quad b_l + b_u = 1$$

for $a_l, b_l \in [0,1]$. Denote

$$m_{\alpha\beta} = a_\alpha b_\beta, \quad \alpha,\beta \in \{l,u\}.$$

Since $f(x)$ is linear on $[x_l, x_u]$ and $[z_l, z_u]$, we get

$$f(x) = a_l f(x_l) + a_u f(x_u), \quad f(z) = b_l f(z_l) + b_u f(z_u).$$

For any function g linear on $[x_l, x_u], [z_l, z_u]$, e.g., $g(x) = 1, g(x) = x, g(x) = f(x)$, we have

$$g(x) = \sum_{\alpha,\beta\in\{l,u\}} m_{\alpha\beta} g(x_\alpha), \quad g(z) - g(x) = \sum_{\alpha,\beta\in\{l,u\}} m_{\alpha\beta}(g(z_\beta) - g(x_\alpha)). \tag{E.1}$$

Using the above identities and the triangle inequality and the definition of M, we get

$$|f(x)-f(z)| = \left| \sum_{\alpha,\beta\in\{l,u\}} m_{\alpha\beta}(f(x_\alpha) - f(z_\beta)) \right| \le \sum_{\alpha,\beta\in\{l,u\}} m_{\alpha\beta} M |x_\alpha - z_\beta|^{1/2}.$$

Using the Cauchy–Schwarz inequality, $|x_\alpha - z_\beta| = z_\beta - x_\alpha$, and (E.1), we establish

$$\begin{aligned} |f(x)-f(z)| &\le \sum_{\alpha,\beta\in\{l,u\}} m_{\alpha\beta} \sum_{\alpha,\beta\in\{l,u\}} m_{\alpha\beta} M |x_\alpha - z_\beta|^{1/2} = \sum_{\alpha,\beta\in\{l,u\}} m_{\alpha\beta} M |x_\alpha - z_\beta|^{1/2} \\ &= M \left(\sum_{\alpha,\beta\in\{l,u\}} m_{\alpha\beta}(z_\beta - x_\alpha) \right)^{1/2} = M(z-x)^{1/2}. \end{aligned}$$

The desired result follows. ∎

We generalize Lemma E.1 to two dimensions as follows.

LEMMA E.2. *Let $I_x = [x_l, x_u], I_z = [z_l, z_u], I_y = [y_l, y_u]$ with $x_l \le x_u \le z_l \le z_u$. Suppose that f is linear on $I_x \times I_y$ and $I_z \times I_y$. Then we have*

$$\max_{x\in I_x, z\in I_z, y\in I_y} \frac{|f(x,y)-f(z,y)|}{|x-z|^{1/2}} = \max_{\alpha,\beta,\gamma\in\{l,u\}} \frac{|f(x_\alpha,y_\gamma)-f(z_\beta,y_\gamma)|}{|x_\alpha - z_\beta|^{1/2}}.$$

Proof. Note that the function $I(x,z,y)=\frac{f(x,y)-f(z,y)}{|x-z|^{1/2}}$ is linear in y. We get

$$|I(x,z,y|=\max(|I(x,z,y_l)|,|I(x,z,y_u)|).$$

Applying Lemma E.1 completes the proof. ∎

Let $\hat{f}$ be the linear interpolation of f. Suppose that $x\in I_x, z\in I_z, y\in I_y$ with $x_u\le z_l$. Using the above estimates and notation, we can bound $I(f)$ as follows:

$$\begin{aligned}I(f)&=\frac{|f(z,y)-f(x,y)|}{|x-z|^{1/2}}\\&\le\frac{|\hat{f}(x,y)-f(x,y)|+|\hat{f}(z,y)-f(z,y)|}{|x-z|^{1/2}}+\max_{\alpha,\beta,\gamma\in\{l,u\}}\frac{|f(x_\alpha,y_\gamma)-f(z_\beta,y_\gamma)|}{|x_\alpha-z_\beta|^{1/2}}\\&\le\left(\frac{h_x^2}{8}||f_{xx}||_{I_x\times I_y}+\frac{h_y^2}{8}(||f_{yy}||_{I_x\times I_y}+||f_{yy}||_{I_z\times I_y})+\frac{h_z^2}{8}||f_{xx}||_{I_z\times I_y}\right)\\&\quad\times|x-z|^{-1/2}+M.\end{aligned}$$

E.2. Piecewise derivative bounds. In this section, we discuss how to obtain the sharp bound of $\frac{p(b)-p(a)}{b-a}$ using piecewise derivative bounds of p.

Suppose that $|p'(y)|\le C_i, y\in I_i=[y_i,y_{i+1}]$. For any $a\in I_k, b\in I_l, a<b$, we have the bound

$$\begin{aligned}|p(b)-p(a)|&\le\int_a^b|p'(y)|dy\le|y_{k+1}-a|C_k+|b-y_l|C_l+\sum_{k+1\le m\le l-1}C_m(y_{m+1}-y_m)\\&=(y_{k+1}-a)C_k+(b-y_l)C_l+M_{kl}(y_l-y_{k+1})\mathbf{1}_{l\ge k+1},\end{aligned}$$

where M_{kl} is defined below:

$$M_{kl}=|y_l-y_{k+1}|^{-1}\left(\sum_{k+1\le m\le l-1}C_m|y_{m+1}-y_m|\right). \tag{E.2}$$

Next, we want to bound $\frac{|p(b)-p(a)|}{|b-a|}$. If $l-k\le 1$, we get

$$|p(b)-p(a)|\le(b-a)\max(C_k,C_l).$$

Otherwise, if $l\ge k+2$, we have

$$|p(b)-p(a)|\le(y_{k+1}-a)(C_k-M_{kl})+(b-y_l)(C_l-M_{kl})+M_{kl}(b-a).$$

Since $\frac{y_{k+1}-a}{b-a}$ is decreasing in a and b, $\frac{b-y_l}{b-a}$ is increasing in b and a, we get

$$0\le\frac{y_{k+1}-a}{b-a}\le\frac{y_{k+1}-y_k}{y_l-y_k},\quad 0\le\frac{b-y_l}{b-a}\le\frac{y_{l+1}-y_l}{y_{l+1}-y_{k+1}}.$$

Using the above estimates, for $a\in I_k, b\in I_l$, we obtain

$$\frac{|p(b)-p(a)|}{|b-a|}\le\max(C_k-M_{kl},0)\frac{y_{k+1}-y_k}{y_l-y_k}+\max(C_l-M_{kl},0)\frac{y_{l+1}-y_l}{y_{l+1}-y_{k+1}}+M_{kl}. \tag{E.3}$$

For uniform mesh, i.e., $y_{i+1}-y_i=h$, we can simplify the above estimate as follows:

$$\frac{|p(b)-p(a)|}{|b-a|} \le \frac{(\max(C_k-M_{kl},0)+\max(C_l-M_{kl},0))}{l-k}+M_{kl},$$
$$M_{kl}=\frac{1}{l-k-1}\sum_{k+1\le m\le l-1} C_m.$$

The same argument applies to obtain piecewise bounds of $J(a,b)=\frac{p(b)-p(a)}{b-a}$. We use piecewise upper bounds $p'(y)\le C_i, y\in I_i=[y_i,y_{i+1}]$ and obtain the same upper bounds as (E.3). To get lower bounds of $J(a,b)$, we use piecewise lower bounds $p'(y)\ge C_i$ and (E.2) to get

$$\frac{p(b)-p(a)}{b-a}\ge \min(C_k-M_{kl},0)\frac{y_{k+1}-y_k}{y_l-y_k}+\min(C_l-M_{kl},0)\frac{y_{l+1}-y_l}{y_{l+1}-y_{k+1}}+M_{kl}.$$

Appendix F. Notation. For the reader's convenience, we collect the main notation used in this paper.

Weights. We use the following weights defined in (A.1), (A.2), (A.3) for the estimates

$$\psi_1,\psi_2,\psi_3,\psi_{du},\psi_u,\quad \varphi_1,\varphi_{g1},\varphi_{elli},\varphi_{evo,i}, i=1,2,3,\quad \rho_{10},\rho_{20},\rho_3,\rho_4.$$

We use $f_\lambda(x)=f(\lambda x)$ for rescaled function (4.2).

Cutoff functions. We use various cutoff functions to construct the approximate solutions.

$\chi_{ij}, i=1,2,3, j=1,2$ are defined in (3.17).

$\chi_{\bar\varepsilon},\chi_{\hat\varepsilon},f_{\chi,i}, i=1,2,3$ are defined in (D.6), (D.7).

Operators. We use $\mathcal{L}.$ to denote various linear operators. $\mathcal{L}_i$ is the full linearized operator around the approximate steady state. We decompose $\mathcal{L}_i$ into $\mathcal{L}_i^e,\mathcal{L}_i^{\bar e},\mathcal{L}_i^N$ (3.12).

$\mathcal{B}_{op,i}$ (3.4), (3.5) denotes bilinear operators related to the linearized operators.

$\mathcal{R}.$ (3.7), (3.8) denotes residual error in the construction of the approximate solution to the linearized equations.

Velocity and kernels. We use K_i to denote the kernels of the velocity, e.g., $K_1,K_2,K_f, f=u,v,u_x,v_x,u_y$ (4.1). We use K^{sym} for the symmetrized kernel (4.25) and K_{ux0},K_{00} (4.5) for the kernel of the approximation terms near $x=0$.

We use $f=u,v,u_x,u_y,v_x,v_y$ to denote the original velocity and its derivatives, $\hat f$ for its finite rank approximation, and $f_A=f-\hat f$. See the beginning of section 4.1.

Regions for integrals. We use $B_{lm}(r)$ (4.17) to denote different grids and $R.(\cdot)$ to denote various singular regions: $R(x,k)$ (4.18), $R_s(x,k), R_{s,i}(x,k)$ (4.19), $R^{\pm}(x,k)$ (4.20), $R(x,k,\alpha), \alpha=N,E,S,W$ (4.21)

Approximate profiles and solutions. We use ω,η,ξ,ϕ to denote the vorticity, θ_x,θ_y (θ is the density (2.3)), and the stream functions, respectively. We use $\bar f$ to denote the approximate profile for f, e.g., $\bar\omega,\bar\theta$, and use $\hat f$ to denote the numeric solution, e.g., $\widehat{W}$ (3.34) and $\hat G$ (3.6).

We use $\bar F_\omega,\bar F_\theta,\bar{\mathcal{F}}_i$ (2.14) to denote the residual error of the profile.

Mesh. To construct the approximate profile, we use the adaptive mesh y_i (C.2). To estimate the integrals $\int f(x,y)dy$ in section 4, we use mesh y_i (4.11) with mesh size h_x,h (4.14).

Differential operators. We denote (3.23) $D^2 = (D_1^2, D_2^2, D_3^2) = (\partial_{xy}, \partial_{xy}, \partial_x^2)^T$.

Acknowledgments. We are grateful to Drs. Pengfei Liu and De Huang for a number of stimulating discussions in the early stage of this project. The first author is grateful to Mr. Xiaoqi Chen for several suggestions on coding and the use of High Performance Computing. Part of the computation in this paper was performed using the Caltech IMSS High Performance Computing. The support from its staff is greatly appreciated.

REFERENCES

[1] T. Buckmaster, S. Shkoller, and V. Vicol, *Formation of shocks for 2D isentropic compressible Euler*, Comm. Pure Appl. Math., 75 (2020), pp. 2069–2120.
[2] T. Buckmaster, S. Shkoller, and V. Vicol, *Formation of point shocks for 3D compressible Euler*, Comm. Pure Appl. Math., 76 (2023), pp. 2073–2191.
[3] R. Caflisch, *Singularity formation for complex solutions of the 3D incompressible Euler equations*, Phys. D, 67 (1993), pp. 1–18.
[4] R. Caflisch and O. F. Orrelana, *Singular solutions and ill-posedness for the evolution of vortex sheets*, SIAM J. Math. Anal., 20 (1989), pp. 249–510.
[5] A. Castro and D. Córdoba, *Infinite energy solutions of the surface quasi-geostrophic equation*, Adv. Math., 225 (2010), pp. 1820–1829.
[6] J. Chen, *Singularity formation and global well-posedness for the generalized Constantin–Lax–Majda equation with dissipation*, Nonlinearity, 33 (2020), 2502.
[7] J. Chen, *On the regularity of the De Gregorio model for the 3D Euler equations*, J. Eur. Math. Soc. (JEMS), (2023), https://doi.org/10.4171/JEMS/1399.
[8] J. Chen, *On the slightly perturbed De Gregorio model on S^1*, Arch. Ration. Mech. Anal., 241 (2021), pp. 1843–1869, https://doi.org/10.1007/s00205-021-01685-w.
[9] J. Chen, *Remarks on the smoothness of the $C^{1,\alpha}$ asymptotically self-similar singularity in the 3D Euler and 2D Boussinesq equations*, Nonlinearity, 37 (2024), 065018.
[10] J. Chen and T. Y. Hou, *MATLAB Codes for Computer-Assisted Proofs in the Paper "Stable Nearly Self-Similar Blowup of the 2D Boussinesq and 3D Euler Equations with Smooth Data"*, 2023, https://jiajiechen94.github.io/codes.
[11] J. Chen and T. Y. Hou, *Stable Nearly Self-Similar Blowup of the 2D Boussinesq and 3D Euler Equations with Smooth Data* II*: Rigorous Numerics*, preprint, arXiv:2305.05660, 2023.
[12] J. Chen and T. Y. Hou, *Finite time blowup of 2D Boussinesq and 3D Euler equations with $C^{1,\alpha}$ velocity and boundary*, Comm. Math. Phys., 383 (2021), pp. 1559–1667.
[13] J. Chen and T. Y. Hou, *Stable Nearly Self-Similar Blowup of the 2D Boussinesq and 3D Euler Equations with Smooth Data* I*: Analysis*, preprint, arXiv:2210.07191, 2022.
[14] J. Chen and T. Y. Hou, *Supplementary Material for "Stable Nearly Self-Similar Blowup of the 2D Boussinesq and 3D Euler Equations with Smooth Data* I*: Analysis*, preprint, arXiv:2210.07191, 2022.
[15] J. Chen, T. Y. Hou, and D. Huang, *On the finite time blowup of the De Gregorio model for the 3D Euler equations*, Comm. Pure Appl. Math., 74 (2021), pp. 1282–1350.
[16] J. Chen, T. Y. Hou, and D. Huang, *Asymptotically self-similar blowup of the Hou–Luo model for the 3D Euler equations*, Ann. PDE, 8 (2022), 24.
[17] K. Choi, T. Hou, A. Kiselev, G. Luo, V. Sverak, and Y. Yao, *On the finite-time blowup of a 1D model for the 3D axisymmetric Euler equations*, Comm. Pure Appl. Math., 70 (2017), pp. 2218–2243.
[18] K. Choi, A. Kiselev, and Y. Yao, *Finite time blow up for a 1D model of 2D Boussinesq system*, Comm. Math. Phys., 334 (2015), pp. 1667–1679.
[19] P. Constantin, *On the Euler equations of incompressible fluids*, Bull. Amer. Math. Soc., 44 (2007), pp. 603–621.
[20] P. Constantin, C. Fefferman, and A. Majda, *Geometric constraints on potentially singular solutions for the 3D Euler equations*, Comm. Partial Differential Equations, 21 (1996).
[21] P. Constantin, P. D. Lax, and A. Majda, *A simple one-dimensional model for the three-dimensional vorticity equation*, Comm. Pure Appl. Math., 38 (1985), pp. 715–724.
[22] S. De Gregorio, *On a one-dimensional model for the three-dimensional vorticity equation*, J. Stat. Phys., 59 (1990), pp. 1251–1263.
[23] J. Deng, T. Hou, and X. Yu, *Geometric properties and nonblowup of 3D incompressible Euler flow*, Comm. Partial Difference Equations, 30 (2005), pp. 225–243.

[24] T. M. Elgindi, *Finite-time singularity formation for $C^{1,\alpha}$ solutions to the incompressible Euler equations on $\mathbb{R}^3$*, Ann. of Math. (2), 194 (2021), pp. 647–727.
[25] T. M. Elgindi, T.-E. Ghoul, and N. Masmoudi, *On the stability of self-similar blow-up for $C^{1,\alpha}$ solutions to the incompressible Euler equations on $\mathbb{R}^3$*, Camb. J. Math., 9 (2021).
[26] T. M. Elgindi, T.-E. Ghoul, and N. Masmoudi, *Stable self-similar blow-up for a family of nonlocal transport equations*, Anal. PDE, 14 (2021), pp. 891–908.
[27] T. M. Elgindi and I.-J. Jeong, *Finite-time singularity formation for strong solutions to the axi-symmetric 3D Euler equations*, Ann. PDE, 5 (2019), pp. 1–51.
[28] T. M. Elgindi and I.-J. Jeong, *On the effects of advection and vortex stretching*, Arch. Ration. Mech. Anal., 235 (2019), pp. 1763–1817, https://doi.org/10.1007/s00205-019-01455-9.
[29] T. M. Elgindi and I.-J. Jeong, *Finite-time singularity formation for strong solutions to the Boussinesq system*, Ann. PDE, 6 (2020), pp. 1–50.
[30] J. Gibbon, *The three-dimensional Euler equations: Where do we stand?*, Phys. D, 237 (2008), pp. 1894–1904.
[31] S. He and A. Kiselev, *Boundary layer models of the Hou-Luo scenario*, J. Differential Equations, 298 (2021), pp. 182–204.
[32] V. Hoang, B. Orcan-Ekmekci, M. Radosz, and H. Yang, *Blowup with vorticity control for a 2D model of the Boussinesq equations*, J. Differential Equations, 264 (2018), pp. 7328–7356.
[33] V. Hoang and M. Radosz, *Singular solutions for nonlocal systems of evolution equations with vorticity stretching*, SIAM J. Math. Anal., 52 (2020), pp. 2158–2178.
[34] T. Hou, *Blow-up or no blow-up? A unified computational and analytic approach to 3D incompressible Euler and Navier-Stokes equations*, Acta Numer., 18 (2009), pp. 277–346.
[35] T. Hou and C. Li, *Dynamic stability of the three-dimensional axisymmetric Navier-Stokes equations with swirl*, Comm. Pure Appl. Math., 61 (2008), pp. 661–697.
[36] T. Hou and R. Li, *Dynamic depletion of vortex stretching and non-blowup of the 3D incompressible Euler equations*, J. Nonlinear Sci., 16 (2006), pp. 639–664.
[37] C. E. Kenig and F. Merle, *Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case*, Invent. Math., 166 (2006), pp. 645–675.
[38] A. Kiselev, *Small scales and singularity formation in fluid dynamics*, in Proceedings of the International Congress of Mathematicians, Vol. 3, 2018.
[39] A. Kiselev and C. Tan, *Finite time blow up in the hyperbolic Boussinesq system*, Adv. Math., 325 (2018), pp. 34–55.
[40] M. J. Landman, G. C. Papanicolaou, C. Sulem, and P.-L. Sulem, *Rate of blowup for solutions of the nonlinear Schrödinger equation at critical dimension*, Phys. Rev. A (3), 38 (1988), pp. 3837–3843.
[41] G. Luo and T. Hou, *Toward the finite-time blowup of the 3D incompressible Euler equations: A numerical investigation*, Multiscale Model. Simul., 12 (2014), pp. 1722–1776.
[42] G. Luo and T. Y. Hou, *Potentially singular solutions of the 3D axisymmetric Euler equations*, Proc. Natl. Acad. Sci. USA, 111 (2014), pp. 12968–12973.
[43] A. Majda and A. Bertozzi, *Vorticity and Incompressible Flow*, Cambridge Texts Appl. Math. 27, Cambridge University Press, Cambridge, 2002.
[44] Y. Martel, F. Merle, and P. Raphaël, *Blow up for the critical generalized Korteweg–de Vries equation. I: Dynamics near the soliton*, Acta Math., 212 (2014), pp. 59–140.
[45] D. W. McLaughlin, G. C. Papanicolaou, C. Sulem, and P.-L. Sulem, *Focusing singularity of the cubic Schrödinger equation*, Phys. Rev. A, 34 (1986), 1200.
[46] F. Merle and P. Raphael, *The blow-up dynamic and upper bound on the blow-up rate for critical nonlinear Schrödinger equation*, Ann. of Math., 161 (2005), pp. 157–222.
[47] F. Merle and H. Zaag, *Stability of the blow-up profile for equations of the type $u_t = \Delta u + |u|^{p-1}u$*, Duke Math. J., 86 (1997), pp. 143–195.
[48] F. Merle and H. Zaag, *On the stability of the notion of non-characteristic point and blow-up profile for semilinear wave equations*, Comm. Math. Phys., 333 (2015), pp. 1529–1562.
[49] R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.
[50] H. Okamoto, T. Sakajo, and M. Wunsch, *On a generalization of the Constantin-Lax-Majda equation*, Nonlinearity, 21 (2008), pp. 2447–2461.
[51] W. Pauls, T. Matsumoto, U. Frisch, and J. Bec, *Nature of complex singularities for the 2D euler equation*, Phys. D, 219 (2006), pp. 40–59.

[52] S. Rump, *INTLAB - INTerval LABoratory*, in Developments in Reliable Computing, T. Csendes, ed., Kluwer Academic Publishers, Dordrecht, the Netherlands, 1999, pp. 77–104, http://www.ti3.tuhh.de/rump/.
[53] S. M. Rump, *Verification methods: Rigorous results using floating-point arithmetic*, Acta Numer., 19 (2010), pp. 287–449.
[54] M. Siegel and R. Caflisch, *Calculation of complex singular solutions to the* 3*D incompressible euler equations*, Phys. D, 238 (2009), pp. 2368–2379.
[55] A. Soffer, *Soliton dynamics and scattering*, in Proceedings of the International Congress of Mathematicians, Vol. 3, 2006, pp. 459–471.
[56] A. Soffer and M. I. Weinstein, *Multichannel nonlinear scattering for nonintegrable equations*, Comm. Math. Phys., 133 (1990), pp. 119–146.
[57] M. I. Weinstein, *Modulational stability of ground states of nonlinear schrödinger equations*, SIAM J. Math. Anal., 16 (1985), pp. 472–491.