

# **SIMPLE SECOND-ORDER FINITE DIFFERENCES FOR ELLIPTIC PDES WITH DISCONTINUOUS COEFFICIENTS AND INTERFACES**

CHUNG-NAN TZOU AND SAMUEL N. STECHMANN

In multiphase fluid flow, fluid-structure interaction, and other applications, partial differential equations (PDEs) often arise with discontinuous coefficients and singular sources (e.g., Dirac delta functions). These complexities arise due to changes in material properties at an immersed interface or embedded boundary, which may have an irregular shape. Consequently, the solution and its gradient can be discontinuous, and numerical methods can be difficult to design. Here a new method is presented and analyzed, using a simple formulation of one-dimensional finite differences on a Cartesian grid, allowing for a relatively easy setup for one-, two-, or three-dimensional problems. The derivation is relatively simple and mainly involves centered finite difference formulas, with less reliance on the Taylor series expansions of typical immersed interface method derivations. The method preserves a sharp interface with discontinuous solutions, obtained from a small number of iterations (approximately five) of solving a symmetric linear system with updates to the right-hand side. Second-order accuracy is rigorously proven in one spatial dimension and demonstrated through numerical examples in two and three spatial dimensions. The method is tested here on the variable-coefficient Poisson equation, and it could be extended for use on time-dependent problems of heat transfer, fluid dynamics, or other applications.

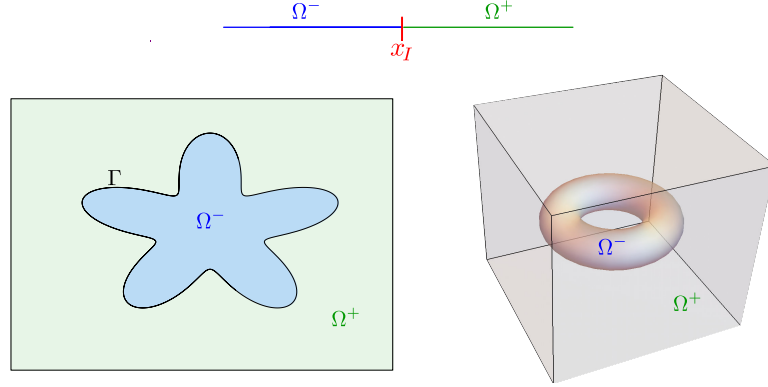
## **1. Introduction**

In many applications, partial differential equations (PDEs) arise with discontinuous coefficients and singular sources (e.g., Dirac delta functions). These complexities often arise due to changes in material properties at an interface or immersed boundary, which may have an irregular shape; see Figure 1. For example, the immersed boundary may be a rigid or flexible structure, such as a heart valve [10], or the immersed interface may separate two fluids as in gas bubbles or liquid droplets [34]. Our own interest was motivated by recently derived equations for atmospheric dynamics, in

---

*MSC2010:* 65M06, 76T99, 35J05.

*Keywords:* sharp interface, immersed boundary method, immersed interface method, ghost fluid method, jump conditions, phase changes.



**Figure 1.** Examples of interfaces separating two regions  $\Omega^-$  and  $\Omega^+$  in (top) 1D, (bottom left) 2D, and (bottom right) 3D.

the limit of rapid rotation and strong (moist) stratification, including phase changes of water and phase interfaces between cloudy and noncloudy regions [33].

For PDEs with such complexities, numerical methods can be challenging to design. Elliptic PDEs are a common test case, and they often form an important component of time-dependent systems. Many methods have been proposed using finite element methods [2; 12], finite volume methods [13; 4], and finite difference methods. Each of these approaches can be valuable in different situations, depending on priorities of computational efficiency, ease of implementation, etc. A primary goal of the present paper is simplicity, and finite difference methods, with Cartesian grids, are perhaps the simplest class of methods. Therefore, for comparison, we next describe some finite difference methods in more detail.

The immersed boundary method (IBM) was introduced in the pioneering work of Peskin [29; 30; 31]. The IBM is simple and efficient and has been applied to a variety of problems with three-dimensional fluid flow [10; 14]. In the IBM approach, the effect of the immersed boundary is represented as a forcing function applied to the fluid. Ideally, the forcing should be singular and the solution should have discontinuities. However, the IBM uses a smoothed version of a Dirac delta function, which introduces some smearing near the boundary or interface and causes the solution to be continuous. The method was originally designed with first-order accuracy, and it has been extended to be “formally” second-order accurate [16; 11; 28; 6; 7], although the “formal” second-order accuracy holds only in the case that the forcing is sufficiently smooth, not in the case of a nearly singular forcing.

The immersed interface method (IIM) was developed to produce improvements such as second-order accuracy and a solution with a sharp discontinuity and no smearing at the interface [19; 20]. The method is derived by allowing an extended stencil, beyond the standard stencil for the Laplacian operator, to be used at grid points near the interface; for the extended stencil, the finite-difference weights are

then found by the method of undetermined coefficients, with constraints on the coefficients being chosen to achieve the desired local truncation error based on Taylor series. The extended stencil of the IIM must be chosen with care in order to avoid instability [8; 21; 5], since the IIM linear operator is not symmetric. One approach is to carefully construct the IIM operator to satisfy a discrete maximum principle by using constrained quadratic optimization techniques [21; 5].

While the IIM has been implemented in multidimensional fluid flow problems, the formulation is complicated by the need for derivations of many spatial and temporal jump conditions, and also derivatives of jump conditions [22; 18; 38; 37]. Many other versions of the IIM with different derivations have been developed [36; 3; 32; 17], and some are discussed further in Section 5 below. In the present paper, one distinguishing feature is that the present derivation involves the relatively simple use of centered finite difference formulas, without the need for derivatives of jump conditions, and with less reliance on the Taylor series expansions of typical IIM derivations. Such simplifications to the derivations should contribute to enhanced ease of use on three-dimensional problems.

The ghost fluid method (GFM) is another method that produces a solution with a sharp discontinuity and no smearing at the interface [23]. While it is only first-order accurate, the GFM is simple to formulate and implement, and it is efficient for problems with three-dimensional multiphase fluid flow [15; 35]. Another advantageous property is that the GFM finite difference operator is symmetric, which allows the use of conjugate gradient algorithms and guarantees robustness of the method.

In the present paper, the goal is to design a method with the advantageous properties of the GFM—sharp interface, easy to formulate and implement, efficient for use on three-dimensional problems, and utilization of a symmetric matrix—while also achieving the possibility of second-order accuracy. The simple formulation here (Section 2) uses elementary finite differences along one-dimensional coordinates, and the resulting linear system can be written with the same symmetric matrix as the GFM but with corrections to the right-hand side that yield second-order accuracy. The right-hand-side corrections are determined iteratively, which is the main new computational expense beyond the GFM. Note that, while this interesting algorithmic connection exists with the GFM, the derivations of the GFM and the present method are quite different; the present method is derived using finite differences (with explicit estimates of local truncation error from finite difference formulas), whereas the GFM and its error and convergence are based on a weak formulation of the problem [24]. Example solutions with the present method are shown for one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D) problems (Section 3). A small, fixed number of iterations ( $\approx 5$ ) is shown to be sufficient for achieving a second-order accurate solution (Section 4), which suggests

the present methods may be efficient enough for use on complex three-dimensional fluid flow. Conclusions and further comparisons with the formulations of other methods [36; 3; 32; 17; 26; 27] are discussed in Sections 5 and 6.

Given that many previous methods have been proposed for this problem over many years, it is worthwhile to emphasize one of the main distinguishing features of the present method: a simple derivation and setup. The derivation here is mainly achieved using centered finite difference formulas, so it is relatively easy to formulate and set up the method, even in 3D. At the same time, the method does utilize a small number of iterations, so it may have a greater computational expense than some other methods (unless one could propose a more sophisticated and faster iterative procedure, a direction which we have not yet pursued exhaustively). In summary, in terms of practical use, the simple derivation and formulation should be useful for applications where one is less concerned with achieving the least possible expense of the computation itself and more concerned with minimizing the time and effort needed to initially design and code the method.

## 2. Numerical methods

In this section, the numerical methods are derived for 1D, 2D, and 3D equations in Sections 2.1, 2.2, and 2.3, respectively. A rigorous proof of second-order convergence is presented in Section 2.1.2 for the 1D case.

**2.1. One dimension.** Consider a one (spatial) dimensional domain  $\Omega$  divided into subdomains  $\Omega^+$  and  $\Omega^-$  by an interface  $\Gamma$ . The variable coefficient Poisson equation on each subdomain reads

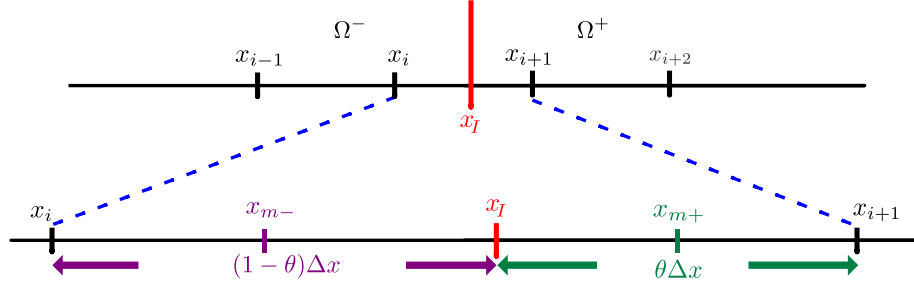
$$(\beta u_x)_x = f(x) \quad \text{for } x \in \Omega \setminus \Gamma, \quad (1)$$

where  $\beta = \beta(x)$  and  $f(x)$  can be discontinuous across interface points  $x_I \in \Gamma$ . The jump conditions across the interface are given as

$$\begin{aligned} [u] &= u^+ - u^- = a(x) \quad \text{for } x \in \Gamma, \\ [\beta u_x] &= \beta^+ u_x^+ - \beta^- u_x^- = b(x) \quad \text{for } x \in \Gamma. \end{aligned} \quad (2)$$

We focus here on the case of two subdomains and one interface point, as it is straightforward to extend the methods for cases with more subdomains and interface points. Here  $u^\pm = \lim_{x \rightarrow x_I^\pm} u(x)$  and  $\beta^\pm = \lim_{x \rightarrow x_I^\pm} \beta(x)$ .

As an alternative formulation of the problem, one could incorporate the jump conditions (2) into the differential equation itself by adding singular sources to the right-hand side of the equation. In such a formulation, the differential equation would take the form  $(\beta u_x)_x = f(x) + b_I \delta(x - x_I) + a_I \bar{\beta} \delta'(x - x_I)$ , where  $\bar{\beta} = (\beta^+ + \beta^-)/2$  and  $a_I = a(x_I)$  and  $b_I = b(x_I)$ , and where this differential equation is valid over the entire domain  $\Omega$ . On the other hand, the differential equation in (1) is



**Figure 2.** Cartesian grid points and an interfacial point in between.

valid only within each of the separate regions  $\Omega^+$  and  $\Omega^-$ , and the jump conditions in (2) are needed to connect the solutions in  $\Omega^+$  and  $\Omega^-$  and complete the problem specification. It will be convenient here to use the separate formulation in (1)–(2) throughout the paper.

**2.1.1. Finite differences.** A second-order finite-difference method can be derived on a Cartesian grid, with a symmetric operator, in the following way.

First, if the interface  $\Gamma = \{x_I\}$  does not intersect with the grid edges connecting the three points  $x_{i-1}$ ,  $x_i$ , and  $x_{i+1}$ , then we call  $x_i$  a standard Cartesian point. For all the standard Cartesian points we follow the standard second-order discretization for (1):

$$\frac{\beta_{i+1/2}((u_{i+1} - u_i)/\Delta x) - \beta_{i-1/2}((u_i - u_{i-1})/\Delta x)}{\Delta x} = f_i + O(\Delta x^2). \quad (3)$$

Next, consider nonstandard Cartesian points, such as  $x_i$  and  $x_{i+1}$  with an interfacial point  $x_I \in \Gamma$  in between and with  $x_i \in \Omega^-$  and  $x_{i+1} \in \Omega^+$ , as shown in Figure 2. Since the number of nonstandard points is assumed to be small, it should be possible to have an overall second-order accurate method that locally uses a first-order discretization at nonstandard points. Therefore, we use a first-order discretization of  $(\beta u_x)_x$ ,

$$(\beta u_x)_x(x_i) = \frac{\beta(x_{m-})u_x(x_{m-}) - \beta(x_{i-1/2})u_x(x_{i-1/2})}{x_{m-} - x_{i-1/2}} + O(\Delta x), \quad (4)$$

followed by second-order discretizations of the  $u_x$  terms, which lead to

$$\frac{\beta_{m-}(u_{I-} - u_i)/((1-\theta)\Delta x) - \beta_{i-1/2}(u_i - u_{i-1})/(\Delta x)}{((2-\theta)/2)\Delta x} = f_i + O(\Delta x), \quad (5)$$

where  $\theta = (x_{i+1} - x_I)/\Delta x$ . Note that the midpoints  $x_{m-} = (x_i + x_I)/2$  and  $x_{m+} = (x_I + x_{i+1})/2$ , illustrated in Figure 2, are useful here to allow second-order discretizations of  $u_x$ . Here  $\beta_{m-} = \beta(x_{m-})$  and  $\beta_{m+} = \beta(x_{m+})$ .

The final step is to replace in (5) the appearance of the interface value  $u_{I-}$  with Cartesian values and adjustments consisting of known quantities. To do this, we

obtain additional equations by discretizing (1) at  $x_{I-}$  and  $x_{I+}$ , the left and right limits of  $x_I$ , using a method similar to the one above:

$$\frac{\beta_{m+}(u_{i+1} - u_{I+})/(\theta \Delta x) - \beta_{I+}u_x(x_{I+})}{\theta \Delta x/2} = f_{I+} + O(\Delta x) \quad \text{at } x_{I+}, \quad (6)$$

$$\frac{\beta_{I-}u_x(x_{I-}) - \beta_{m-}(u_{I-} - u_i)/((1-\theta)\Delta x)}{(1-\theta)\Delta x/2} = f_{I-} + O(\Delta x) \quad \text{at } x_{I-}, \quad (7)$$

where  $\beta_{I\pm} = \beta(x_{I\pm})$ . The non-Cartesian unknowns  $u_x(x_{I\pm})$  and  $u_{I\pm}$  above can now be replaced by Cartesian unknowns by the following two steps. First, the weighted sum  $(\theta \Delta x/2) \cdot (6) + ((1-\theta)\Delta x/2) \cdot (7)$  is a combination that produces the jump  $[\beta u_x]$ :

$$\begin{aligned} \beta_{m+} \left( \frac{u_{i+1} - u_{I+}}{\theta \Delta x} \right) - \beta_{m-} \left( \frac{u_{I-} - u_i}{(1-\theta)\Delta x} \right) - [\beta u_x] \\ = (\theta \cdot f_{I+} + (1-\theta) \cdot f_{I-}) \frac{\Delta x}{2} + O(\Delta x^2). \end{aligned} \quad (8)$$

Second, by using the jump conditions (2), we see that (8) can be rewritten as our desired formula for replacing  $u_{I-}$  by Cartesian  $u$  values:

$$\begin{aligned} u_{I-} = \frac{\hat{\beta}(1-\theta)}{\beta_{m-}} u_{i+1} + \frac{\hat{\beta}\theta}{\beta_{m+}} u_i \\ - \frac{\hat{\beta}\theta(1-\theta)\Delta x^2}{\beta_{m+}\beta_{m-}} \left( \frac{\beta_{m+}a_I}{\theta \Delta x^2} + \frac{b_I}{\Delta x} + \frac{1}{2}(\theta \cdot f_{I+} + (1-\theta) \cdot f_{I-}) \right), \end{aligned} \quad (9)$$

where

$$\hat{\beta} = \frac{\beta_{m+}\beta_{m-}}{(1-\theta) \cdot \beta_{m+} + \theta \cdot \beta_{m-}}. \quad (10)$$

Lastly, substituting (9) into (5) yields a first-order discretization of the differential equation at  $x_i$ , in terms of only Cartesian values of  $u$ :

$$\begin{aligned} \frac{1}{\Delta x^2} (\beta_{i-1/2} \cdot u_{i-1} - (\beta_{i-1/2} + \hat{\beta})u_i + \hat{\beta} \cdot u_{i+1}) = f_i \cdot \left( \frac{2-\theta}{2} \right) \\ + \frac{\hat{\beta}\theta}{\beta_{m+}} \left( \frac{\beta_{m+}}{\theta} \frac{a_I}{\Delta x^2} + \frac{b_I}{\Delta x} + \frac{1}{2}(\theta \cdot f_{I+} + (1-\theta) \cdot f_{I-}) \right) + O(\Delta x). \end{aligned} \quad (11)$$

For the neighboring nonstandard point at  $x_{i+1}$ , one can derive a similar finite difference formula:

$$\begin{aligned} \frac{1}{\Delta x^2} (\hat{\beta} \cdot u_i - (\hat{\beta} + \beta_{i+3/2})u_{i+1} + \beta_{i+3/2} \cdot u_{i+2}) = f_{i+1} \cdot \left( \frac{1+\theta}{2} \right) \\ + \frac{\hat{\beta}(1-\theta)}{\beta_{m-}} \left( -\frac{\beta_{m-}}{(1-\theta)} \frac{a_I}{\Delta x^2} + \frac{b_I}{\Delta x} + \frac{1}{2}(\theta \cdot f_{I+} + (1-\theta) \cdot f_{I-}) \right) + O(\Delta x). \end{aligned} \quad (12)$$

Comparing (11) and (12), it is clear that the difference operator acting on  $u$  is symmetric. The linear system can be solved using many standard efficient methods.

Note that this method in (11)–(12) looks similar to the GFM, which is first-order accurate [23; 24], but (11)–(12) include important differences that render this method second-order accurate. For instance, the right-hand-side terms in (11)–(12) have coefficients that are different from the GFM and that arise here as part of a systematic finite-differences derivation. Also, the values of  $\beta$  at the midpoints  $x_{m-}$  and  $x_{m+}$  were needed for the present method, whereas  $\beta$  values at the interface and Cartesian grid points and Cartesian midpoints are utilized in the GFM [23; 24].

In comparison to the IIM [19], notice that the present method has a symmetric operator, whereas the IIM operator is nonsymmetric. Also, the derivation of the IIM requires taking derivatives of jump conditions, whereas the present method is derived by simply applying finite difference formulas to the differential equation. It would be interesting to try to make a more firm connection between the present method and the IIM, which might also help tie together the GFM and IIM; however, we have not found any simple and clear connection beyond the comparisons described above.

To summarize, the basic idea in deriving (11)–(12) was to (i) start with midpoint-based finite differences using both Cartesian points and interface points, and then (ii) use the jump conditions to eliminate the interface values  $u_{I\pm}$  from the system.

### 2.1.2. Proof of second-order convergence.

**Theorem 1.** *The numerical solution in Section 2.1.1 converges to the exact solution in the  $L^2$  norm with second-order accuracy:  $\|U - U_{\text{ex}}\|_2 = O(\Delta x^2)$ .*

*Proof.* The setup of the proof is as follows. The numerical method in (3), (11), and (12) can be written in matrix-vector form as  $AU = F$ , and the exact solution satisfies  $AU_{\text{ex}} = F + \tau$ , where  $\tau$  is the local truncation error. The error  $e = U - U_{\text{ex}}$  then satisfies  $Ae = -\tau$ , and solving for  $e$  gives  $e = -A^{-1}\tau$ . The  $L^2$  norm of the error then satisfies

$$\|e\|_2 = \|A^{-1}\tau\|_2 \leq \|A^{-1}\|_2 \|\tau\|_2, \quad (13)$$

where the remaining task is to analyze  $\|A^{-1}\|_2$  and  $\|\tau\|_2$  for small  $\Delta x$ .

Consistency was established in Section 2.1.1. Specifically, the local truncation error can be written as

$$\tau = \tau_s + \tau_{\text{ns}} \quad \text{with } \|\tau_s\|_2 = O(\Delta x^2) \text{ and } \|\tau_{\text{ns}}\|_2 = O(\Delta x^2), \quad (14)$$

where we have split  $\tau$  so that the elements of  $\tau_s$  are nonzero only at standard points and the elements of  $\tau_{\text{ns}}$  are nonzero only at nonstandard points. The  $O(\Delta x^2)$  scaling in (14) is then true because each element of  $\tau_s$  is  $O(\Delta x^2)$ , based on the finite difference formulas at the standard points; and each element of  $\tau_{\text{ns}}$  is  $O(\Delta x)$ , but the fraction of nonstandard points is  $O(\Delta x)$ , so  $\|\tau_{\text{ns}}\|_2 = O(\Delta x^2)$ .

Stability is established by the bound

$$\|A^{-1}\|_2 \leq \frac{|\Omega|^2}{\beta_m}, \quad (15)$$

where  $|\Omega|$  is the total length of the domain and  $\beta_m = \min_{x \in \Omega} \beta(x)$  is a constant that is independent of  $\Delta x$ , and it is assumed that  $\beta(x) > 0$  for all  $x$ . The proof of this bound is well-known [25] and is based on summation by parts and discrete Poincaré–Friedrichs inequality.

The proof of the theorem is completed by combining the consistency and stability results in (14) and (15) to show that (13) is  $O(\Delta x^2)$ .  $\square$

Note that, in the  $L^\infty$  norm, the local truncation error can only be bounded as  $\|\tau_{\text{ns}}\|_\infty = O(\Delta x)$ , which would complicate the present proof technique for second-order convergence if attempted with the  $L^\infty$  norm instead of the  $L^2$  norm. Also, note that we have no such proof in two- or three-dimensional space, although proofs for 2D and 3D have been presented for similar methods [1], and numerical examples below demonstrate second-order convergence, in both the  $L^2$  and  $L^\infty$  norms.

**2.2. Two dimensions.** Now consider the two-dimensional Poisson equation

$$(\beta u_x)_x + (\beta u_y)_y = f(x, y) \quad \text{for } \Omega \setminus \Gamma, \quad (16)$$

where  $\Omega = \Omega^+ \cup \Omega^- \cup \Gamma$  and  $\Gamma$  is the interface between the sets  $\Omega^+$  and  $\Omega^-$ . With  $\mathbf{n} = (n^1(x, y), n^2(x, y))$  as the unit normal along  $\Gamma$ , the interface jump conditions are given as

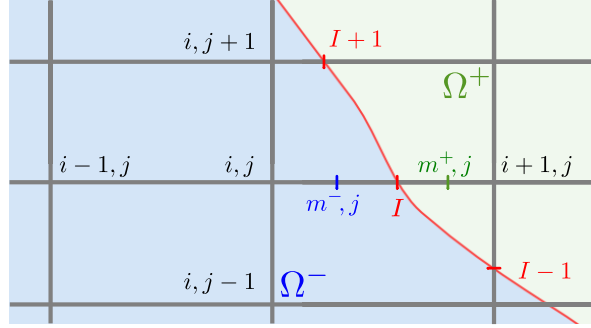
$$\begin{aligned} [u] &= u^+ - u^- = a(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Gamma, \\ [\beta u_n] &= \beta^+ u_n^+ - \beta^- u_n^- = b(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Gamma, \end{aligned} \quad (17)$$

where  $u_n = \mathbf{n} \cdot \nabla u$  is the derivative of  $u$  in the direction of the normal vector.

**2.2.1. Finite differences.** The goal of this section is to extend the ideas of the 1D case of Section 2.1 to the 2D case of (16)–(17) and arrive at a second-order finite-difference method. Similar to the 1D case, we call a Cartesian point  $(x_i, y_j)$  a standard point if this point and its nearest neighbors all lie within  $\Omega^+$  or all lie within  $\Omega^-$ . For standard points, (16) is discretized with the standard, second-order, five-point finite-difference formula. For nonstandard points, on the other hand, the interface must be taken into account.

For nonstandard points, such as point  $(x_i, y_j)$  illustrated in Figure 3, we obtain a first-order discretization by using similar ideas as in the 1D case. Following a derivation similar to (5)–(11), by essentially just replacing  $f$  by  $f - (\beta u_y)_y$ , we





**Figure 3.** Nonstandard grid point at  $(x_i, y_j)$ .

arrive at

$$\begin{aligned} & \frac{1}{\Delta x^2} (\beta_{i-1/2,j} \cdot u_{i-1,j} - (\beta_{i-1/2,j} + \hat{\beta}) u_{i,j} + \hat{\beta} \cdot u_{i+1,j}) \\ & + \frac{1}{\Delta y^2} (\beta_{i,j-1/2} \cdot u_{i,j-1} - (\beta_{i,j-1/2} + \beta_{i,j+1/2}) u_{i,j} + \beta_{i,j+1/2} \cdot u_{i,j+1}) \\ & = f_{i,j} \cdot \left( \frac{2-\theta}{2} \right) + (\beta u_y)_y(x_i, y_j) \cdot \frac{\theta}{2} + F_{\text{cor}}^x + O(\Delta x), \quad (18) \end{aligned}$$

where

$$\hat{\beta} = \frac{\beta(x_{m+}, y_j) \cdot \beta(x_{m-}, y_j)}{(1-\theta) \cdot \beta(x_{m+}, y_j) + \theta \cdot \beta(x_{m-}, y_j)}, \quad (19)$$

and

$$\begin{aligned} F_{\text{cor}}^x = & \frac{\hat{\beta} \theta}{\beta(x_{m+}, y_j)} \left\{ \frac{\beta(x_{m+}, y_j) a(x_I, y_j)}{\theta \Delta x^2} + \frac{[\beta u_x]}{\Delta x} \right. \\ & \left. + \frac{1}{2} (\theta \cdot (f - (\beta u_y)_y)(x_{I+}, y_j) + (1-\theta) \cdot (f - (\beta u_y)_y)(x_{I-}, y_j)) \right\}. \quad (20) \end{aligned}$$

This finite-difference formula has a left-hand side with the desirable property of a symmetric operator, as in the 1D case. However, the right-hand side of (18) now depends on the solution  $u$  itself, so an iterative method will be described below for finding a solution.

Also, a more general case would allow for other interface crossings, such as a crossing at point  $(x_i, y_j)$ , with  $y_j < y_I < y_{j+1}$ , which would generate some slight modifications to the derivation and finite-difference formula. Since the more general case is only slightly different from (18), it is relegated to Appendix A.

To estimate the derivatives on the right-hand side of (18), simple finite differences are used. For the term  $(\beta u_y)_y(x_i, y_j)$ , standard centered differences can be used with the points  $(x_i, y_{j-1})$ ,  $(x_i, y_j)$ , and  $(x_i, y_{j+1})$ . For the term  $(\beta u_y)_y(x_{I-}, y_j)$  at the interface, from (20), one can approximate it with the nearby Cartesian value

$(\beta u_y)_y(x_i, y_j)$  with an acceptable error of  $O(\Delta x)$ , and then one can use a standard centered discretization with the points  $(x_i, y_{j-1})$ ,  $(x_i, y_j)$ , and  $(x_i, y_{j+1})$ . The term  $(\beta u_y)_y(x_{I+}, y_j)$  can be handled similarly by using  $(\beta u_y)_y$  at the nearby Cartesian point  $(x_{i+1}, y_j)$ . Lastly, the jump  $[\beta u_x]$  from (20) can be written in terms of normal and tangential jumps as

$$\begin{aligned} [\beta u_x] &= [\beta u_n]n^1 - [\beta u_\tau]n^2 \\ &= b_I n^1 - [\beta u_\tau]n^2. \end{aligned} \quad (21)$$

The term  $[\beta u_\tau]$  can then be estimated using finite differences with  $u$  values from the interface points labeled  $I-1$ ,  $I$ , and  $I+1$  in Figure 3 (or possibly using another triplet, say  $I-2$ ,  $I$ , and  $I+1$ , if the two interface points  $I-1$  and  $I$  are located too close together, such as within  $O(h^2)$  distance; see Appendix B for details). Note that a second-order finite-difference formula is needed for  $[\beta u_\tau]$  in order for the term  $[\beta u_x]/\Delta x$  to have an error of  $O(\Delta x)$ . To determine the  $u$  values at the interface points, one can use the formula

$$\begin{aligned} u(x_{I-}, y_j) &= \frac{(1-\theta)\hat{\beta}}{\beta(x_{m-}, y_j)} u_{i+1,j} + \frac{\theta\hat{\beta}}{\beta(x_{m+}, y_j)} u_{i,j} - \frac{\hat{\beta}(1-\theta)\theta\Delta x^2}{\beta(x_{m+}, y_j)\beta(x_{m-}, y_j)} \\ &\quad \cdot \left( \frac{\beta(x_{m+}, y_j)a_I}{\theta\Delta x^2} + \frac{[\beta u_x]}{\Delta x} + \frac{\theta}{2} \cdot ((\beta u_x)_x)_{i+1,j} + \frac{(1-\theta)}{2} \cdot ((\beta u_x)_x)_{i,j} \right), \end{aligned} \quad (22)$$

and  $u(x_{I+}, y_j) = u(x_{I-}, y_j) + a(x_I, y_j)$  by the jump condition (17). This formula arises as part of the derivation of (18) and is similar to the 1D case, and formulas for  $u(x_i, y_{J\pm})$  can be obtained similarly if the crossing is in the  $y$ -direction. Note that this formula in 2D does not actually provide the desired result of the interface  $u$  value in terms of the Cartesian  $u$  values, since the right-hand side depends on interface  $u$  values via the  $[\beta u_x]$  term. Nevertheless, this formula can be used as part of an iterative procedure to complete the specification of the numerical methods.

**2.2.2. Iterative methods.** In this section, a simple iterative method is proposed here for solving the linear system from Section 2.2.1.

Before describing the standard iterative method of the present paper, consider first a type of Picard iteration:

$$A\mathbf{u}^{[k+1]} = \mathbf{F}^{[k]}. \quad (23)$$

This is an iterative version of the matrix-vector form of the finite difference method, one row of which is described in (18):  $A$  is the symmetric matrix from the left-hand side,  $\mathbf{u}^{[k+1]}$  is the vector of all Cartesian  $u$  values (from iteration  $k+1$ ), and  $\mathbf{F}^{[k]}$  is the vector from the right-hand-side terms. The basic idea is to iteratively update  $\mathbf{F}^{[k]}$  on the right-hand side as new, more accurate information about  $\mathbf{u}^{[k]}$  is obtained. As an initial condition,  $\mathbf{F}^{[0]}$  is defined as the right-hand side of (18) with all instances of  $u$

ignored (i.e., with  $[\beta u_\tau]$ ,  $(\beta u_x)_x$ , and  $(\beta u_y)_y$  all set to zero, essentially equivalent to setting  $u^{[0]} = 0$  as an initial guess), and the first solution  $\mathbf{u}^{[1]}$  is found by solving  $A\mathbf{u}^{[1]} = \mathbf{F}^{[0]}$ . As a result, *the solution  $\mathbf{u}^{[1]}$  at the first iteration is essentially the same as the GFM solution [23; 24] and is therefore a first-order accurate solution.* It can be used to estimate the interface  $u$  values, which we assemble abstractly into a vector  $\mathbf{u}_I^{[k]}$  and update iteratively as  $\mathbf{u}_I^{[k+1]} = B\mathbf{u}_I^{[k]} + C\mathbf{u}^{[k+1]} + \mathbf{G}$ , one row of which is described by (22): the  $B\mathbf{u}_I^{[k]}$  corresponds to the  $[\beta u_\tau]$  term, the  $C\mathbf{u}^{[k+1]}$  corresponds to all terms with Cartesian  $u$  values, and the  $\mathbf{G}$  corresponds to the jump terms involving  $a_I$  and  $b_I$ . An initial interface value of  $\mathbf{u}_I^{[0]} = \mathbf{\Gamma}$  is used, consistent with the idea of ignoring all instances of  $u$  in the initial condition  $\mathbf{F}^{[0]}$ . The second iteration then proceeds by defining  $\mathbf{F}^{[1]}$  based on the right-hand side of (18) and now using  $\mathbf{u}^{[1]}$  and  $\mathbf{u}_I^{[1]}$  to provide a more accurate estimate of the true  $\mathbf{F}$  value. The solution  $\mathbf{u}^{[2]}$  at the second iteration is then found from solving the symmetric system  $A\mathbf{u}^{[2]} = \mathbf{F}^{[1]}$ . This procedure can be repeated to iteratively estimate the solution of the finite-difference method.

For the stopping criterion for the iterative procedure, the differences  $u_d^{[k]} = \|\mathbf{u}^{[k+1]} - \mathbf{u}^{[k]}\|_\infty$  and  $F_d^{[k]} = \|\mathbf{F}^{[k+1]} - \mathbf{F}^{[k]}\|_\infty$  are monitored. When  $k$  is large enough so that  $u_d^{[k]} < C_u h^2$ , where  $h = \Delta x = \Delta y$ , one can presumably stop iterating since the iterations are producing only small corrections that are within the desired  $O(h^2)$  accuracy of the numerical solution. As our standard stopping criterion, in addition to  $u_d^{[k]} < C_u h^2$  we also require  $F_d^{[k]} < C_F h$  in order to ensure that the estimated right-hand-side terms are not significantly changing at any location. The constants  $C_u$  and  $C_F$  will be set equal to 1 here for simplicity, but in the future it would be interesting to tailor the choices of  $C_u$  and  $C_F$  to the particular problem under consideration; for instance, they could be chosen based on the expected error, which could be estimated based on, e.g., expected local truncation error and/or smoothness of the solution. Also note that, while this standard stopping criterion was chosen with solution accuracy as the main consideration, one could also imagine other stopping criteria that consider computational efficiency or other factors; some other stopping criteria are explored in Section 4.

As the standard iterative method used here, a modification of Picard iteration is actually used. While Picard iteration does work well in many cases, we found that it diverges in some cases. Nevertheless, by making some slight modifications, a robust method can be designed. Our standard iterative method here uses a simple relaxation procedure to extend Picard iteration; it is described in Appendix C, and it is shown below to provide robust results.

**2.3. Three dimensions.** The three-dimensional Poisson equation is

$$(\beta u_x)_x + (\beta u_y)_y + (\beta u_z)_z = f(x, y, z), \quad \text{for } \Omega \setminus \Gamma, \quad (24)$$

where  $\Omega = \Omega^+ \cup \Omega^- \cup \Gamma$  and  $\Gamma$  is a surface that marks the interface between the sets  $\Omega^+$  and  $\Omega^-$ . The interface jump conditions are given as in the 2D case in (17).

The 3D discretization is essentially the same as in the 2D case in Section 2.2. We note one difference that arises: in 3D, the jump  $[\beta u_x]$  from (21) takes the form

$$\begin{aligned} [\beta u_x] &= [\beta u_n]c^0 + [\beta u_{\tau_1}]c^1 + [\beta u_{\tau_2}]c^2 \\ &= b_I c^0 + [\beta u_{\tau_1}]c^1 + [\beta u_{\tau_2}]c^2, \end{aligned} \quad (25)$$

where  $\hat{\mathbf{x}} = c^0 \hat{\mathbf{n}} + c^1 \hat{\mathbf{\tau}}_1 + c^2 \hat{\mathbf{\tau}}_2$  was used to write the unit coordinate vector  $\hat{\mathbf{x}}$  in terms of the interface normal vector  $\hat{\mathbf{n}}$  and two unit vectors  $\hat{\mathbf{\tau}}_1$  and  $\hat{\mathbf{\tau}}_2$  from the 2D tangent plane of the interface. Here, in 3D, note that tangential derivatives are needed in two independent directions in the 2D tangent plane. The two directions can be conveniently chosen by using the Cartesian coordinate planes. For example, if  $(x_I, y_j, z_k) \in \Gamma$ , where  $x_I$  is not a Cartesian grid point, then the intersection of surface  $\Gamma$  and the plane  $z = z_k$  can be used to define one direction in the 2D tangent plane, and the intersection of surface  $\Gamma$  and the plane  $y = y_j$  can be used to define the other direction. In this way, computation of the tangential derivatives in 3D can be reduced to essentially the same form as in 2D.

### 3. Examples

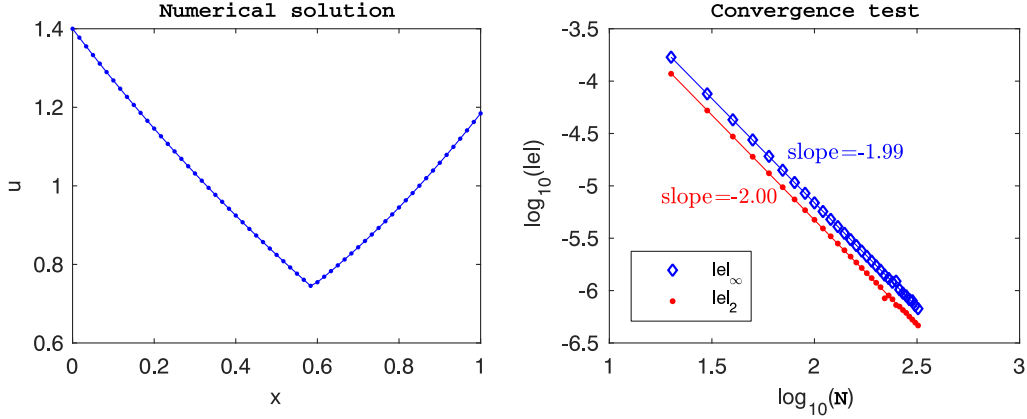
In this section, second-order convergence is demonstrated through numerical examples. In all examples, the same grid spacing is used in each coordinate direction ( $\Delta x = \Delta y = \Delta z$ ), and the number of grid points in each coordinate direction is  $N$ , so the total number of grid points is  $N$ ,  $N^2$ , or  $N^3$  for the 1D, 2D, or 3D cases, respectively.

#### 3.1. One dimension.

**Example 1D-1.** Consider a domain  $\Omega = [0, 1]$  separated into subdomains  $\Omega^- = [0, x_I)$  and  $\Omega^+ = (x_I, 1]$ , where  $x_I = 2 - \sqrt{2}$ . The solution to the one-dimensional equation  $\beta u_{xx} = f$  is  $u^- = \exp(-x) - 0.3646x + 0.4$  and  $u^+ = \exp(-x)/2 + x^2/2 + 0.5005x$  where  $\beta = 100$  in  $\Omega^-$  and  $\beta = 200$  in  $\Omega^+$ , with  $f = 100 \exp(-x)$  in  $\Omega^-$  and  $f = 100 \exp(-x) + 200$  in  $\Omega^+$ . The jump conditions connecting the two equations at  $x_I$  are  $a(x_I) = u^+ - u^- = 0$  and  $b(x_I) = 100(2u_x^+ - u_x^-) = 253.72$ .

The exact solution and error analysis of this example can be found in Figure 4.

**3.2. Two dimensions.** The following 2D and 3D examples are tested on some rectangular domain  $\Omega$  where  $\Omega$  will be divided into  $\Omega^+$  and  $\Omega^-$  by an interface  $\Gamma$ . It will sometimes be convenient to describe the interface  $\Gamma$  in terms of a level-set function  $\phi(x)$  as  $\Gamma = \{\mathbf{x} \in \Omega : \phi(\mathbf{x}) = 0\}$ , where the two sets  $\Omega^+$  and  $\Omega^-$  can be described as  $\Omega^+ = \{\mathbf{x} \in \Omega : \phi(\mathbf{x}) > 0\}$  and  $\Omega^- = \{\mathbf{x} \in \Omega : \phi(\mathbf{x}) < 0\}$ . The coefficients



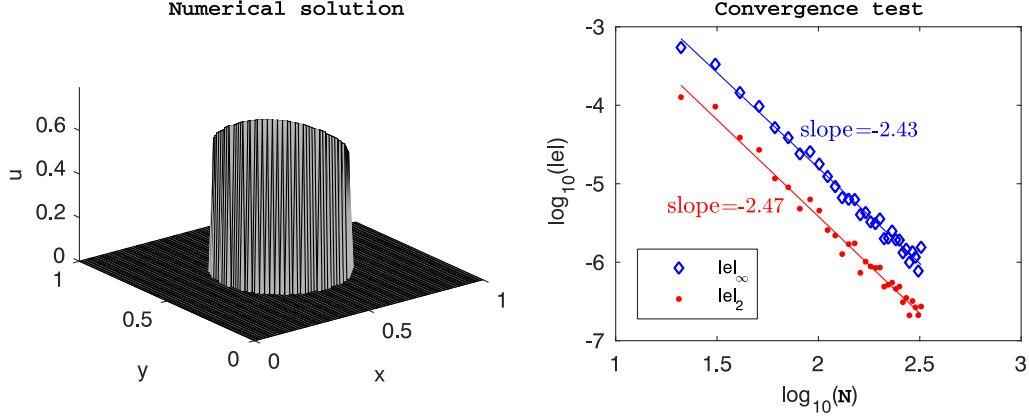
**Figure 4.** Example 1D-1. Left: numerical solution with number of grid points  $N = 61$ . Right: error  $\|e\|$  as a function of number of grid points  $N$ , as a log-log plot including slope of its linear fit.

$\beta$  are assumed to be smooth in both  $\Omega^+$  and  $\Omega^-$ , but may have a jump across the interface  $\phi$ . The piecewise smooth  $\beta$  in  $\Omega^+$  and  $\Omega^-$  will be denoted by  $\beta^+$  and  $\beta^-$ , respectively. As a consequence, the solution  $u$  may be discontinuous across  $\phi$ , but is  $\mathcal{C}^2$  in both  $\Omega^+$  and  $\Omega^-$ , and will similarly be denoted by  $u^+$  and  $u^-$ , respectively. The examples in this section will provide tests of the numerical method for several factors that could influence the numerical method's convergence, such as geometry of the interface, spatial variations in the coefficients  $\beta^\pm(x, y)$ , and contrast  $\beta^+/\beta^-$  due to jumps in the coefficients.

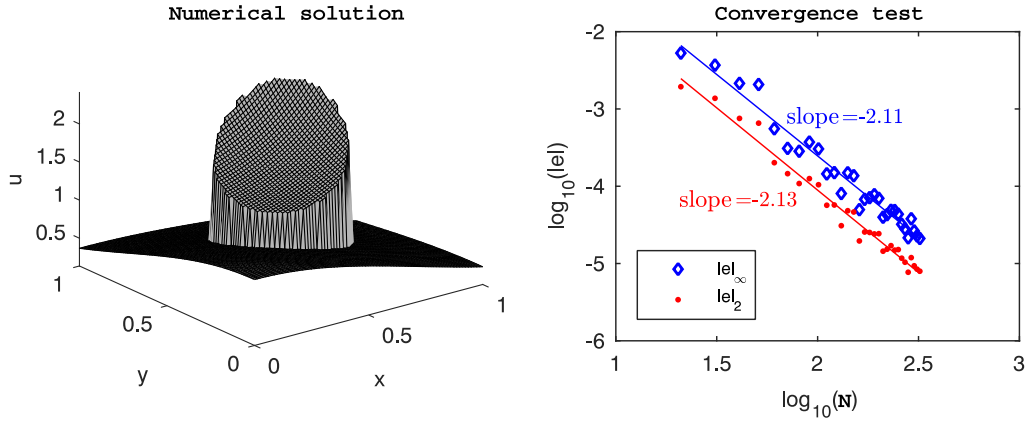
**Example 2D-1** (constant coefficient). In this example, we take  $\beta$  be a piecewise constant function with  $\beta^- = 2$  and  $\beta^+ = 1$ , and the interface is a circle described by the level set function  $\phi(x, y) = (x - 0.5)^2 + (y - 0.5)^2 - 0.25^2$ . The solution is  $u^- = \exp(-x^2 - y^2)$ ,  $u^+ = 0$ , with  $f^- = 8(x^2 + y^2 - 1) \exp(-x^2 - y^2)$ ,  $f^+ = 0$ , on the domain  $\Omega = [0, 1] \times [0, 1]$ . Second-order convergence can be seen in Figure 5, right.

**Example 2D-2** (variable coefficient). The next example we take  $\beta$  to be a piecewise smooth function with  $\beta^- = x^2 + y^2 + 1$ , and  $\beta^+ = 1$  with the same domain and level set function as the previous example. The solution is  $u^- = \exp(x^2 + y^2)$ , and  $u^+ = \exp(-x^2 - y^2)$  and source term is  $f^- = 4(\beta^-(x^2 + y^2 + 1) + (x^2 + y^2)) \exp(x^2 + y^2)$ ,  $f^+ = 4(x^2 + y^2 - 1) \exp(-x^2 - y^2)$ . Error analysis is presented in Figure 6, right.

**Example 2D-3** (variable coefficient). With the same solution  $u$  in Example 2D-2, this example is computed on a domain  $\Omega = [-1, 1] \times [-1, 1]$ , with  $\beta^- = x^2 + y^2 + 1$ ,  $\beta^+ = \sqrt{x^2 + y^2 + 2}$ , and  $f^- = 4(\beta^-(x^2 + y^2 + 1) + (x^2 + y^2)) \exp(x^2 + y^2)$ ,  $f^+ = (4\beta^+(x^2 + y^2 - 1) - 2(x^2 + y^2)/\sqrt{x^2 + y^2 + 2}) \exp(-x^2 + y^2)$ . The interface



**Figure 5.** Example 2D-1: constant coefficient. Left: numerical solution,  $N = 81$ . Right: error  $\|e\|$  as a function of number of grid points in each coordinate direction,  $N$ , as a log-log plot including slope of its linear fit.



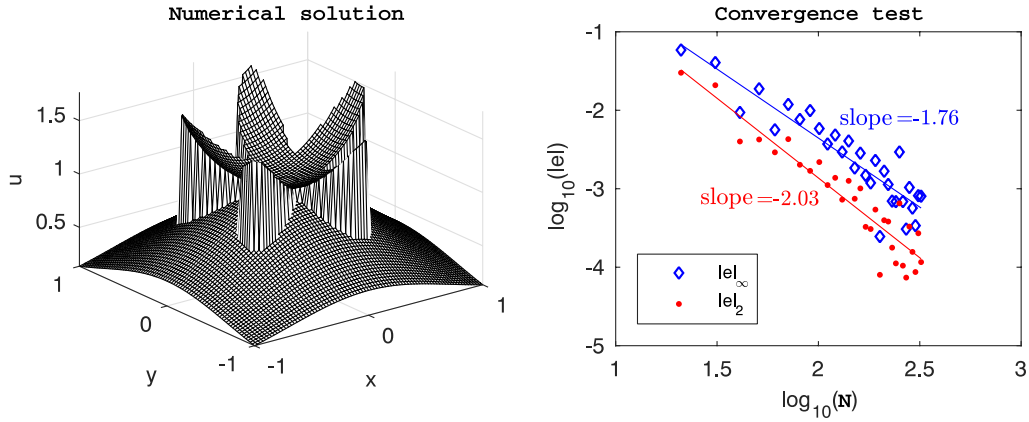
**Figure 6.** Example 2D-2: variable coefficient. Left: numerical solution,  $N = 81$ . Right: error  $\|e\|$  as a function of number of grid points in each coordinate direction,  $N$ , as a log-log plot including slope of its linear fit.

is parametrized by

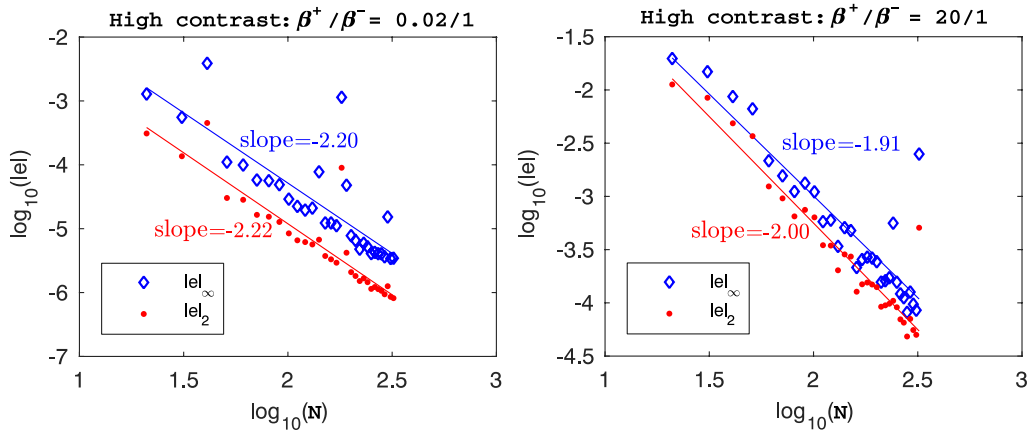
$$\begin{cases} x(t) = 0.02\sqrt{5} + (0.5 + 0.2 \sin(5t)) \cos t, \\ y(t) = 0.02\sqrt{5} + (0.5 + 0.2 \sin(5t)) \sin t, \end{cases} \quad (26)$$

with  $t \in [0, 2\pi]$ . Second-order convergence is demonstrated in Figure 7, right.

For the spatial variations of the error, we describe the two cases of Figures 6 and 7. In these cases the error appears to typically take its maximum value near the interface. In the special case of the outlier of  $N = 251$  from Figure 7, right, the error furthermore takes its maximum value at a single localized spike near one point close to the interface. These features are possibly related to the curvature of the interface or other factors.



**Figure 7.** Example 2D-3: variable coefficient. Left: numerical solution,  $N = 81$ . Right: error  $\|e\|$  as a function of number of grid points in each coordinate direction,  $N$ , as a log-log plot including slope of its linear fit.

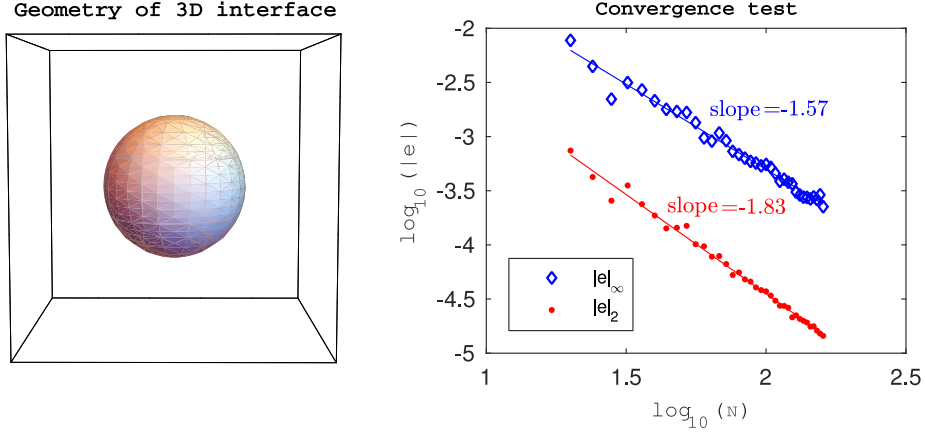


**Figure 8.** Error plots for high-contrast case, Example 2D-4. Left:  $\beta^+/\beta^- = 0.02/1$ . Right:  $\beta^+/\beta^- = 20/1$ .

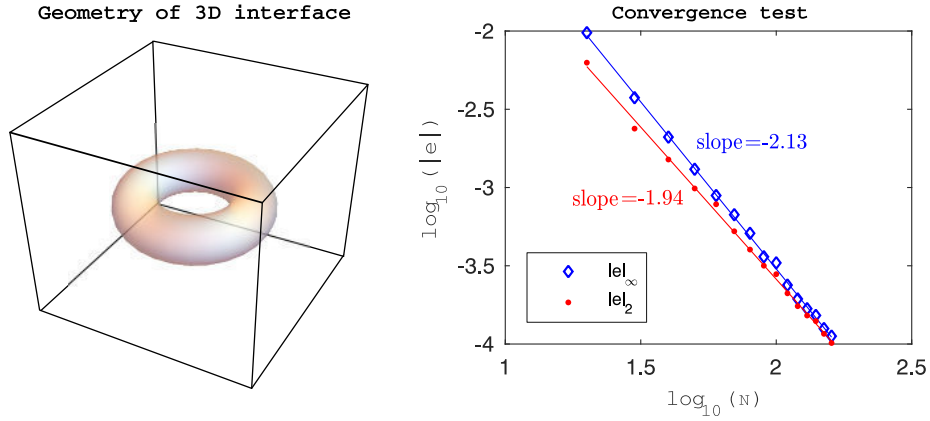
**Example 2D-4** (high-contrast coefficient cases). A series of tests were conducted on the large coefficient ratios, either  $\beta^+/\beta^- \ll 1$  or  $1 \ll \beta^+/\beta^-$ . Here we test with  $u^- = \exp(x^2 + y^2)$ , and  $u^+ = \exp(-x^2 - y^2)$  with a circular interface as in Example 2D-1, and  $(\beta^+, \beta^-) = (0.02, 1)$  and  $(20, 1)$ . Second-order convergence can still be obtained (see Figure 8).

### 3.3. Three dimensions.

**Example 3D-1** (variable coefficient with spherical interface). Domain  $\Omega = [0, 1] \times [0, 1] \times [0, 1]$  is divided into  $\Omega^+$  and  $\Omega^-$  by a sphere centered at  $(0.5, 0.5, 0.5)$  with radius 0.25. The variable coefficients  $\beta$  in (24) are  $\beta^- = 10 + \sin(xy + z)$  and  $\beta^+ = 10 + \cos(x + yz)$ , with solution  $u^- = \exp(x^2 + y^2 + z^2)$  and  $u^+ = 0$ .



**Figure 9.** Example 3D-1: variable coefficient with spherical interface. Left: geometry of the interface. Right: error  $\|e\|$  as a function of number of grid points in each coordinate direction,  $N$ , as a log-log plot including slope of its linear fit.



**Figure 10.** Example 3D-2: variable coefficient with torus interface. Left: numerical solution. Right: error  $\|e\|$  as a function of number of grid points in each coordinate direction,  $N$ , as a log-log plot including slope of its linear fit.

and  $f^- = (4\beta^-(x^2 + y^2 + z^2 + 3/2) + (4xy + 2z) \cos(xy + z)) \exp(x^2 + y^2 + z^2)$ ,  $f^+ = 0$ . See Figure 9 for the geometry of the spherical interface and second-order convergence in  $L^2$ .

**Example 3D-2** (variable coefficient with torus interface). For the same  $\beta$ ,  $u$ , and  $f$  in the previous example, we test this iterative method on  $\Omega = [-1, 1] \times [-1, 1] \times [-1, 1]$  with a toroid interface described by the level set function  $\phi(x, y, z) = (x^2 + y^2 + z^2 + R^2 - r^2)^2 - 4R^2(x^2 + y^2)$ , where  $R = 0.501 + \sqrt{2}/10$ ,  $r = 0.251$ . The geometry of the interface and second-order convergence in  $L^\infty$  and  $L^2$  are in Figure 10.



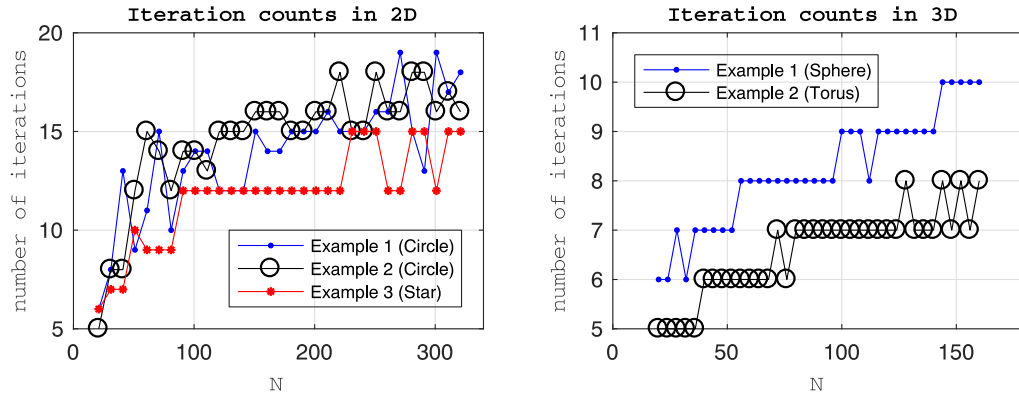
#### 4. Greater efficiency via alternative stopping criteria

**4.1. Iteration counts for standard stopping criterion.** In most of the cases shown above, the number of iterations required to reach the stopping criterion is small, which makes this iterative method efficient, as demonstrated in Figure 11. More specifically, approximately 10–20 iterations are used in 2D cases, and approximately 5–10 iterations in the 3D cases. For high-contrast cases (Example 2D-4), the number of iterations becomes larger (approximately 50–150, as seen in Figure 12), although it is smaller (approximately 20–40) for some  $N$ , and the bound on the number of iterations is essentially independent of the number of grid points.

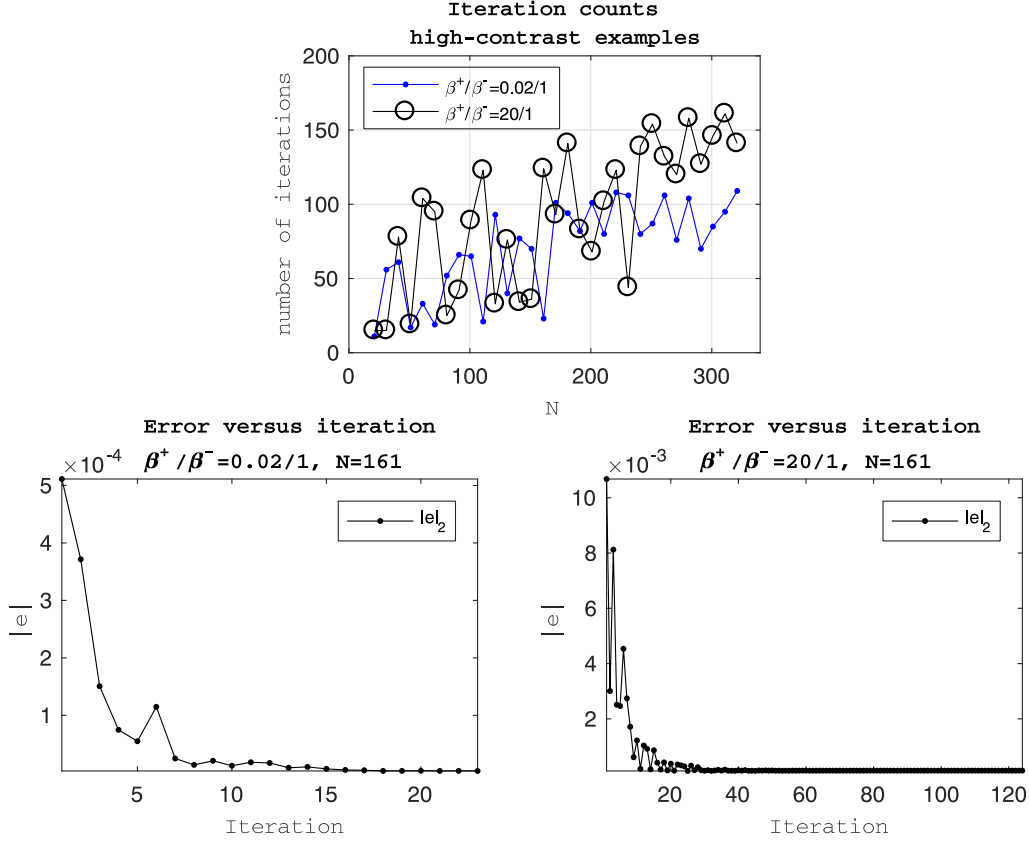
These examples demonstrate that the present method may be practical and efficient for time-dependent problems where the elliptic solver is needed at every time step. Also, further reduction of the iteration counts may be possible for time-dependent problems. For instance, in these first explorations, we are using a simple yet crude initial guess for the iterative methods; i.e., we are effectively using  $u^{[0]} = 0$  (see Section 2.2.2). For a better initial guess  $u^{[0]}$ , in a time-dependent problem, the solution from the previous time step could potentially be used, with the possibility of substantially reducing the number of iterations required.

Below we discuss other possibilities of further reducing the iterations counts through alternative stopping criteria—e.g., by using a small, fixed number of iterations in Section 4.2, and propose some other feasible stopping criteria in Section 4.3.

**4.2. Greater efficiency via a small, fixed number of iterations.** In most cases, the accuracy improves tremendously after only a few iterations; in other words, the latter iterations make only small modifications to the solution in order to satisfy the stopping criterion. Therefore, in practice, we may speed up this numerical method by using a fixed number of iterations without losing too much accuracy. Figure 13



**Figure 11.** Number of iterations for (left) 2D examples and (right) 3D examples.

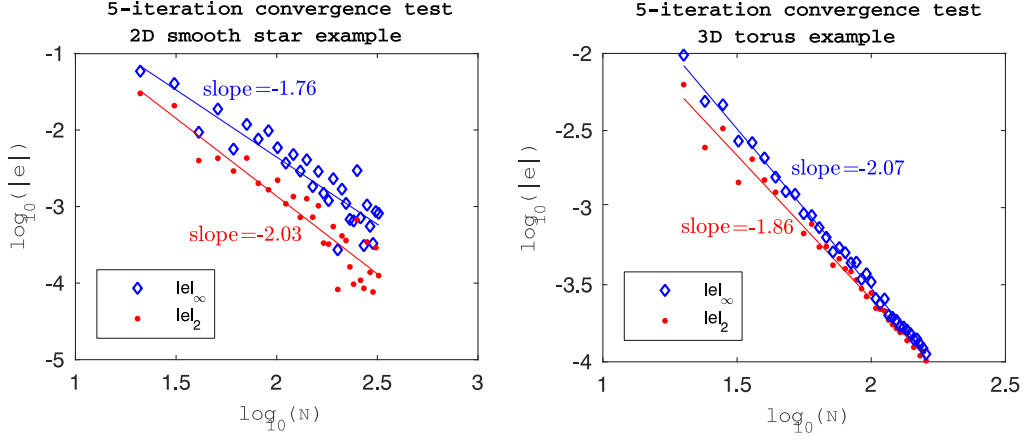


**Figure 12.** Errors and iterations for the high-contrast cases from Example 2D-4. Top: number of iterations as a function of the number of grid points in each coordinate direction,  $N$ . Bottom:  $L^2$  error as a function of iterations, for  $N = 161$ .

shows results of both 2D and 3D examples with only a small number of iterations (five), which still show second-order accuracy.

**4.3. Other stopping criteria.** Several other stopping criteria were also tested, beyond the standard criterion from Section 2.2.2, by using different combinations of criteria for the smallness of the differences  $u_d^{[k]} = \|\mathbf{u}^{[k+1]} - \mathbf{u}^{[k]}\|_\infty$  and/or  $F_d^{[k]} = \|\mathbf{F}^{[k+1]} - \mathbf{F}^{[k]}\|_\infty$ . A promising criterion may be to stop when  $u_d^{[k]} < h^2$ , without enforcing any smallness criterion on  $F_d^{[k]}$ ; in some tests, this led to second-order accuracy with fewer iterations, although we have not yet tested this criterion on a wide array of cases.

Also, it would be interesting in the future to tailor the stopping criterion to the particular problem at hand. For instance, the criterion  $u_d^{[k]} < h^2$  could be slightly generalized to  $u_d^{[k]} < C_u h^2$ , where  $C_u$  is a parameter that could be chosen based on, e.g., the expected accuracy or smoothness of the solution, given the parameters of the problem such as  $\beta^\pm(x, y, z)$ ,  $f(x, y, z)$ , and the geometry of the interface.



**Figure 13.** Error as a function of number of grid points in each coordinate direction,  $N$ , using a fixed number of iterations (five) for more efficient computations. Left: smooth star example in Section 3.2. Right: torus example in Section 3.3.

## 5. Comparisons with formulations of other methods

In this section we compare the present formulation with the formulations of other methods [36; 3; 32; 17; 26; 27], to add to the comparisons with the GFM [23; 24] and IIM [19; 20; 8; 21; 5] that were described above in Section 2.1.1. An IIM viewpoint of the present method is also described in the 1D case at the end of the section.

In [36], another approach had been taken to obtain a symmetric operator; the derivation used Taylor series expansions and derivatives of jump conditions, which can be somewhat complex compared to the simple derivations of the present paper that mainly involve centered finite difference formulas. Note that the present method and the method of [36] are, in fact, distinct. As one difference, in the 1D versions of the two methods, the method of [36] requires not only the symmetric operator but also some augmentation to account for jumps, whereas the method of the present paper has an operator in 1D that is symmetric on its own. Also, the method [36] utilizes a discretization of the standard Laplacian operator, whereas the present method maintains the symmetry of the elliptic operator that includes  $\beta$ .

In [3], an interesting approach was proposed which, like the present method, involves a symmetric operator and an iterative method to determine an adjusted forcing. The derivation is somewhat complex in that it is a version of the IIM and therefore uses Taylor series and derivatives of jump conditions. The derivation is presented in 2D, but no 3D results are presented. Also, their iterative procedure does not produce a first-order accurate solution at the first iteration, and therefore, it is likely to require a very large number of iterations (as possibly indicated by their very small relaxation parameter). The number of iterations, however, are not

reported, and the iterative methods and stopping criterion are not described in detail. In contrast, in the present paper, the first iteration is essentially the GFM, and the simple finite-difference formulation allows for efficient setup and computation even in 3D.

In [17], following [32], another interesting approach is used to obtain a symmetric operator with corrections to the right-hand side. The method is implemented in 2D, but no 3D results are presented. Also, the method is presented for the standard Laplacian operator, not for the case of discontinuous and/or spatially varying coefficient  $\beta(\mathbf{x})$ .

Another interesting method called the correction-function method has been developed by building on the GFM and computing a corrected forcing function to achieve higher-order accuracy [26; 27]. In this method, the corrected forcing function is not derived explicitly; instead, the corrected forcing function is shown to satisfy a certain new PDE, and the new PDE is solved numerically to determine the corrected forcing function. The method has been demonstrated to achieve second-order and even fourth-order accuracy, although it has not yet been implemented for 3D problems and it has only been developed for cases with constant coefficients and piecewise-constant coefficients. It is similar to the method of the present paper in that both methods seek to compute corrections to the GFM; the present paper's method perhaps offers a simpler formulation (involving only one-dimensional finite differences) and simpler implementation for 3D problems.

Note that the present method appears to be well-behaved for any values of subcell location  $\theta$ , even values of  $\theta$  that are close to 0 or 1; in contrast, it has been noted in some applications of the GFM [9] that poor behavior could potentially result if  $\theta$  is close to 0 or 1. One might expect poor behavior in these cases, since the finite difference formulas in, e.g., (6)–(7) have factors of  $1/\theta$  and  $1/(1-\theta)$ ; however, these formulas are part of the derivation only, not part of the present numerical algorithm. Here, no special treatment of the case  $\theta \approx 0$  or  $\theta \approx 1$  was used in the numerical results, and in examining the possible influence of  $\theta$  on the numerical error, we found no systematic relationship. (For computing tangential derivatives  $\beta u_\tau$ , which are not part of the GFM but are used in the present method in higher dimensions, some values of  $\theta$  are given special treatment for computing  $u_\tau$ ; see Section 2.2.1.) A theoretical explanation can be seen from the finite difference formulas of the method; in particular, notice that (11) and (12) depend on  $\theta$  in a smooth, nonsingular way. In fact, a more precise error analysis shows that the error term  $O(\Delta x)$  in (11) is  $O(\theta \Delta x)$  and in (12) is  $O((1-\theta)\Delta x)$ , which theoretically suggests that  $\theta \approx 0$  or  $\theta \approx 1$  should lead to smaller truncation error, and also indicates that the proof of the convergence theorem in Section 2.1.2 is valid for any value with  $0 < \theta < 1$ . Hence, the method should work fine with  $\theta \approx 0$  or  $\theta \approx 1$  except for possible values of precisely  $\theta = 0$  or  $\theta = 1$ .

## 6. Conclusions

In this article, a simple numerical scheme is proposed to obtain second-order accuracy in solving the Poisson equation with sharp interfaces. One important contribution is a simple derivation that mainly involves centered finite difference formulas, with less reliance on the Taylor series expansions and derivatives of jump conditions used in typical immersed interface method derivations. The derivation here preserves the symmetry of the differential operator, and the method is formulated on a Cartesian grid. The accuracy of the method is proved rigorously in 1D and verified numerically in 2D and 3D. The three-dimensional problems are relatively easy to set up due to the method's simple derivation.

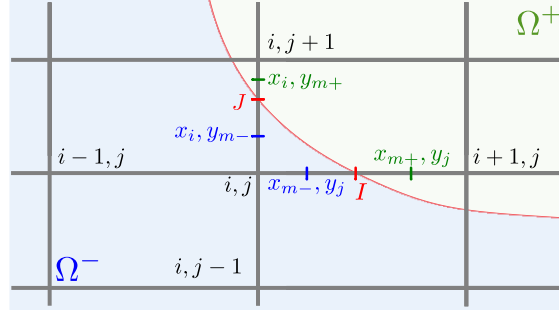
An iterative procedure was used for solving 2D or 3D problems, and the desired second-order accuracy can be obtained with only a small, fixed number of iterations (typically five), which makes this method efficient, even in 3D. In the future it would be interesting to investigate other algorithmic choices; for instance, perhaps an iterative method could be designed that requires an even smaller number (e.g., two or three) of iterations, or perhaps the method could be successful if the iterated correction terms were instead written as part of the left-hand-side linear operator, in which case the symmetry of the operator is lost but the nonsymmetric system could possibly be solved without the need for the outer iterations introduced in the present paper. Also, here we did not make a great effort to optimize the algorithms for cases with high-contrast coefficients, which require higher iteration counts, but such an effort would be interesting to pursue in the future.

The proposed method may be applied to solving time-dependent problems that require the solution of an elliptic PDE at each time step—for example, the heat equation with interfaces or multiphase flow problems [10; 34; 33]. In such applications, the present method could be used with any characterization of the interface (level set, Lagrangian markers, etc.), and the interface could have a location and shape that evolves in time. Also, for the iterative algorithms described here, some computational savings may be possible for time-dependent problems, since the solution from the previous time step could provide a good initial guess for the iterative method at the current time step.

## Appendix A. 2D discretization with two interface crossings

In this appendix, it is shown how to formulate the finite difference method in a case that is more general than in Section 2.2.1.

Suppose the interface crosses the stencil of point  $(x_i, y_j)$  in two places, as shown in Figure 14. The crossing between  $(x_i, y_j)$  and  $(x_{i+1}, y_j)$  is as in Section 2.2.1, and now a new, second crossing is present between  $(x_i, y_j)$  and  $(x_i, y_{j+1})$ . Accordingly,



**Figure 14.** Nonstandard point  $(x_i, y_j)$  with interface crossing the stencil in both  $x$  and  $y$  directions.

define  $\zeta = (y_{j+1} - y_J)/\Delta y$ , where  $(x_i, y_J) \in \Gamma$ , and assume  $(x_i, y_j) \in \Omega^-$  and  $(x_i, y_{j+1}) \in \Omega^+$ .

To obtain a finite difference method with a symmetric operator in this case, start by writing the 1D formula from (11) as

$$S^x u = (\beta u_x)_x \cdot (2 - \theta)/2 + F_{\text{cor}}^x + O(\Delta x), \quad (27)$$

where  $S^x$  is the symmetric finite difference operator and  $F_{\text{cor}}^x$  is the correction term. A similar formula can be derived for a symmetric finite difference operator in the  $y$  direction:

$$S^y u = (\beta u_y)_y \cdot (2 - \zeta)/2 + F_{\text{cor}}^y + O(\Delta y). \quad (28)$$

Summing up the two leads to

$$S^x u + S^y u = f - (\beta u_x)_x \cdot \theta/2 - (\beta u_y)_y \cdot \zeta/2 + F_{\text{cor}}^x + F_{\text{cor}}^y + O(\Delta x) + O(\Delta y), \quad (29)$$

which is the desired formula. Also note that the derivation in 3D follows the same simple principles by including the addition of a third component for  $S^z u$ .

Written out in detail, (29) takes the form

$$\begin{aligned} & \frac{1}{(\Delta x)^2} (\beta(x_{i-1/2}, y_j) \cdot u_{i-1,j} - (\beta(x_{i-1/2}, y_j) + \hat{\beta})u_{i,j} + \hat{\beta} \cdot u_{i+1,j}) \\ & + \frac{1}{(\Delta y)^2} (\beta(x_i, y_{j-1/2}) \cdot u_{i,j-1} - (\beta(x_i, y_{j-1/2}) + \tilde{\beta})u_{i,j} + \tilde{\beta} \cdot u_{i,j+1}) \\ & = f_{i,j} - (\beta u_x)_x(x_i, y_j) \cdot \frac{\theta}{2} - (\beta u_y)_y(x_i, y_j) \cdot \frac{\zeta}{2} + F_{\text{cor}}^x + F_{\text{cor}}^y + O(\Delta x), \end{aligned} \quad (30)$$

where  $\hat{\beta}$  is the same as (19) and

$$\tilde{\beta} = \frac{\beta(x_i, y_{m+}) \cdot \beta(x_i, y_{m-})}{(1 - \zeta) \cdot \beta(x_i, y_{m+}) + \zeta \cdot \beta(x_i, y_{m-})}, \quad (31)$$

with midpoints  $y_{m+} = (y_J + y_{J+1})/2$  and  $y_{m-} = (y_j + y_J)/2$ , and

$$F_{\text{cor}}^x = \frac{\hat{\beta}\theta}{\beta(x_{m+}, y_j)\Delta x} \left\{ \frac{\beta(x_{m+}, y_j)a(x_I, y_j)}{\theta\Delta x} + [\beta u_x] \right. \\ \left. + \left( \theta \cdot (f - (\beta u_y)_y)(x_{I+}, y_j) + (1 - \theta) \cdot (f - (\beta u_y)_y)(x_{I-}, y_j) \right) \frac{\Delta x}{2} \right\}, \quad (32)$$

$$F_{\text{cor}}^y = \frac{\tilde{\beta}\zeta}{\beta(x_i, y_{m+})\Delta y} \left\{ \frac{\beta(x_i, y_{m+})a(x_i, y_J)}{\zeta\Delta y} + [\beta u_y] \right. \\ \left. + \left( \zeta \cdot (f - (\beta u_x)_x)(x_i, y_{J+}) + (1 - \zeta) \cdot (f - (\beta u_x)_x)(x_i, y_{J-}) \right) \frac{\Delta y}{2} \right\}, \quad (33)$$

where  $[\beta u_x] = [\beta u_n]n^1 - [\beta u_\tau]n^2$  and  $[\beta u_y] = [\beta u_\tau]n^1 + [\beta u_n]n^2$ .

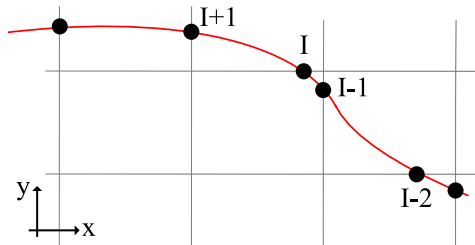
Several variations could also be used. For instance, on the right-hand side of (30), one may replace  $(\beta u_x)_x$  by  $f - (\beta u_y)_y$ , or one may replace  $(\beta u_y)_y$  by  $f - (\beta u_x)_x$ . Similar replacements could be made in (32) and (33). For our numerical tests, we used the  $(\beta u_y)_y$ -based version:  $S^x u + S^y u = f \cdot (2 - \theta)/2 + (\beta u_y)_y \cdot (\theta - \zeta)/2 + F_{\text{cor}}^x + F_{\text{cor}}^y$ .

## Appendix B. Computing tangential derivatives

Here we compute the tangential derivatives up to second order using the values of  $u$  at the interfacial points and two of the neighboring interfacial points (see Figure 15 for reference). By parametrizing our boundary of domain in either  $y = y(x)$  or  $x = x(y)$ , whichever is a properly defined function, we first compute  $dy/dx(x_I)$  or  $dx/dy(y_I)$  and  $dw/dx$  or  $dw/dy$  correspondingly, where  $w = u(x, y(x))$  or  $w = u(x(y), y)$ .

Suppose  $y = y(x)$  is a properly defined function as in Figure 15; the unit tangent vector of the interface  $\tau = (\tau_1, \tau_2)$  at  $(x_I, y_I)$  can be written as

$$(\tau_1, \tau_2) = \begin{cases} (1, dy/dx)/\sqrt{1 + (dy/dx)^2} & \text{if } \tau_1 > 0, \\ -(1, dy/dx)/\sqrt{1 + (dy/dx)^2} & \text{if } \tau_1 < 0. \end{cases} \quad (34)$$



**Figure 15.** Three neighboring interface points are used to compute  $u_\tau$  centered at point  $I$ .

Combining (34) with  $u_\tau = u_x \tau_1 + u_y \tau_2$ ,

$$u_\tau = \frac{(u_x + u_y y_x)}{\sqrt{1 + (dy/dx)^2}} \cdot \text{sign } \tau_1 = \frac{w_x \cdot \text{sign } \tau_1}{\sqrt{1 + (dy/dx)^2}} = w_x \cdot \tau_1. \quad (35)$$

Suppose  $w_y$  and  $x_y$  are both accurate to second order; we have

$$\begin{aligned} u_\tau &= \frac{(w_x + O(h^2))}{\sqrt{1 + (dy/dx + O(h^2))^2}} \sim \frac{(w_x + O(h^2))}{\sqrt{1 + (dy/dx)^2 + O(h^2)}} \\ &= \frac{(w_x + O(h^2))}{\sqrt{1 + (dy/dx)^2} \sqrt{1 + O(h^2)}} \sim \frac{(w_x + O(h^2))}{\sqrt{1 + (dy/dx)^2}} (1 - \frac{1}{2} O(h^2)) \\ &= \frac{w_x}{\sqrt{1 + (dy/dx)^2}} + O(h^2). \end{aligned} \quad (36)$$

Both the derivatives with parametrization should be accurate up to second order. Using the interfacial and neighboring two interfacial points, we can compute  $y_x$  to second order by

$$ay(x_{I+1}) + by(x_I) + cy(x_{I-1}) = y_x(x_I) + O(h^2), \quad (37)$$

where

$$D = \Delta x_l \Delta x_r (\Delta x_r - \Delta x_l), \quad a = -\Delta x_l^2 / D, \quad c = \Delta x_r^2 / D, \quad b = -(a + c), \quad (38)$$

and  $\Delta x_l = x_{I-1} - x_I$  and  $\Delta x_r = x_{I+1} - x_I$ . Note that the denominator  $D$  can be small when neighboring interface points are very close, which leads to numerical issues. This issue can be avoided by, for example, using  $(x_{I+1}, y_{I+1})$ ,  $(x_I, y_I)$ , and  $(x_{I-2}, y_{I-2})$  instead of  $(x_{I+1}, y_{I+1})$ ,  $(x_I, y_I)$ , and  $(x_{I-1}, y_{I-1})$  in Figure 15.

Similar set up for either  $w_x$  or  $w_y$  follows the above:

$$aw(x_{I+1}) + bw(x_I) + cw(x_{I-1}) = w_x(x_I) + O(h^2), \quad (39)$$

and the coefficients  $a$ ,  $b$ , and  $c$  are exactly the same as above. Second-order  $bu_\tau$  hence follows from the equation  $bu_\tau = \pm w_x/y_x$  or  $bu_\tau = \pm w_y/x_y$ , depending on  $|y_x| < 1$  or  $|x_y| < 1$ , and the sign adjustment comes from the sign of  $\tau_1$  and  $\tau_2$ , respectively.

### Appendix C. Relaxation

As discussed in Section 2.2.2, Picard iteration works well in many cases, but we found that it sometimes diverges. For this reason, as our standard iterative scheme, we instead use a simple relaxation scheme to bypass this difficulty and guarantee that the iterative scheme stops. The idea behind the relaxation scheme is to update



the forcing term as

$$\mathbf{F}^{[k]} = \alpha_k \mathbf{F}^{[T_k]} + (1 - \alpha_k) \mathbf{F}^{[k-1]}, \quad (40)$$

which is a mixture between the previous forcing  $\mathbf{F}^{[k-1]}$  and the temporary forcing  $\mathbf{F}^{[T_k]}$  that would have been used if a Picard update would have been followed. The parameter  $\alpha_k$  is chosen to guarantee that  $\mathbf{u}^{[k+1]}$  is not too far away from  $\mathbf{u}^{[k]}$ .

One cycle of the relaxation scheme goes as follows. Suppose  $\mathbf{u}^{[k]}$  was computed by solving  $A\mathbf{u}^{[k]} = \mathbf{F}^{[k-1]}$ , and we now want to compute the next iteration. With  $\mathbf{u}^{[k]}$ , compute the temporary right-hand-side  $\mathbf{F}^{[T_k]}$  by following the Picard update procedure from Section 2.2.2. A temporary solution  $\mathbf{u}^{[T_{k+1}]}$  is then obtained by solving  $A\mathbf{u}^{[T_{k+1}]} = \mathbf{F}^{[T_k]}$ . Now the parameter  $\alpha_k$  is determined to guarantee that  $\mathbf{u}^{[k+1]}$  is not too far away from  $\mathbf{u}^{[k]}$ ; to this end, define the ratio  $r_k = \|\mathbf{u}^{[T_{k+1}]} - \mathbf{u}^{[k]}\| / \|\mathbf{u}^{[k]} - \mathbf{u}^{[k-1]}\|$ . If this ratio is small ( $r_k < 1$ ), then there is no need for relaxation and we set  $\alpha_k = 1$ . If this ratio is large ( $r_k \geq 1$ ), then we set  $\alpha_k = \rho / r_k$ , where  $\rho$  is a preselected factor between 0 and 1. In practice, we pick  $\|\cdot\| = \|\cdot\|_\infty$  and  $\rho$  to be between 0.9 and 0.99. With this relaxation scheme for the forcing  $\mathbf{F}^{[k]}$ , the solution is likewise updated as  $\mathbf{u}^{[k+1]} = \alpha_k \mathbf{u}^{[T_{k+1}]} + (1 - \alpha_k) \mathbf{u}^{[k]}$ , as a mixture of the previous solution estimate  $\mathbf{u}^{[k]}$  and the temporary solution estimate  $\mathbf{u}^{[T_{k+1}]}$  that would have been used if a Picard update would have been followed.

The differences  $u_d^{[k]} = \|\mathbf{u}^{[k+1]} - \mathbf{u}^{[k]}\|_\infty$  and  $F_d^{[k]} = \|\mathbf{F}^{[k+1]} - \mathbf{F}^{[k]}\|_\infty$  are guaranteed to be decreasing as  $k$  increases if this relaxation procedure is followed. Specifically, the relaxation procedure leads to either  $u_d^{[k]} = r_k u_d^{[k-1]}$  (if  $r_k < 1$ ) or  $u_d^{[k]} = \rho u_d^{[k-1]}$  (if  $r_k \geq 1$ ). Therefore,  $u_d^{[k]}$  is decreasing in  $k$  and hence the stopping criterion will be met in a finite number of iterations. Note that this stopping criterion, based on  $\|\mathbf{u}^{[k+1]} - \mathbf{u}^{[k]}\|$ , does not guarantee that the relaxation procedure's iterate  $\mathbf{u}^{[k+1]}$  is actually close to the exact solution; nevertheless, one would expect that it should be at least a better estimate than the first iterate  $\mathbf{u}^{[1]}$ , which is the first-order accurate GFM solution; and in practice we find from the examples in Section 3 that the iterations terminate at a second-order accurate solution.

### Acknowledgments

The research of Stechmann is partially supported by a Sloan Research Fellowship and National Science Foundation grant AGS-1443325. The authors thank J. Thomas Beale, Aleksandar Donev, Randall J. LeVeque, and anonymous reviewers for helpful comments.

### References

- [1] J. T. Beale and A. T. Layton, *On the accuracy of finite difference methods for elliptic problems with interfaces*, Commun. Appl. Math. Comput. Sci. **1** (2006), 91–119. MR Zbl

- [2] J. Bedrossian, J. H. von Brecht, S. Zhu, E. Sifakis, and J. M. Teran, *A second order virtual node method for elliptic problems with interfaces and irregular domains*, J. Comput. Phys. **229** (2010), no. 18, 6405–6426. MR Zbl
- [3] P. A. Berthelsen, *A decomposed immersed interface method for variable coefficient elliptic equations with non-smooth and discontinuous solutions*, J. Comput. Phys. **197** (2004), no. 1, 364–386. MR Zbl
- [4] R. K. Crockett, P. Colella, and D. T. Graves, *A Cartesian grid embedded boundary method for solving the Poisson and heat equations with discontinuous coefficients in three dimensions*, J. Comput. Phys. **230** (2011), no. 7, 2451–2469. MR Zbl
- [5] S. Deng, K. Ito, and Z. Li, *Three-dimensional elliptic solvers for interface problems and applications*, J. Comput. Phys. **184** (2003), no. 1, 215–243. MR Zbl
- [6] T. G. Fai, B. E. Griffith, Y. Mori, and C. S. Peskin, *Immersed boundary method for variable viscosity and variable density problems using fast constant-coefficient linear solvers, I: Numerical method and results*, SIAM J. Sci. Comput. **35** (2013), no. 5, B1132–B1161. MR Zbl
- [7] ———, *Immersed boundary method for variable viscosity and variable density problems using fast constant-coefficient linear solvers, II: Theory*, SIAM J. Sci. Comput. **36** (2014), no. 3, B589–B621. MR Zbl
- [8] A. L. Fogelson and J. P. Keener, *Immersed interface methods for Neumann and related problems in two and three dimensions*, SIAM J. Sci. Comput. **22** (2000), no. 5, 1630–1654. MR Zbl
- [9] F. Gibou, R. P. Fedkiw, L.-T. Cheng, and M. Kang, *A second-order-accurate symmetric discretization of the Poisson equation on irregular domains*, J. Comput. Phys. **176** (2002), no. 1, 205–227. MR Zbl
- [10] B. E. Griffith, X. Luo, D. M. McQueen, and C. S. Peskin, *Simulating the fluid dynamics of natural and prosthetic heart valves using the immersed boundary method*, Int. J. Appl. Mech. **1** (2009), no. 1, 137–177.
- [11] B. E. Griffith and C. S. Peskin, *On the order of accuracy of the immersed boundary method: higher order convergence rates for sufficiently smooth problems*, J. Comput. Phys. **208** (2005), no. 1, 75–105. MR Zbl
- [12] J. L. Hellrung, Jr., L. Wang, E. Sifakis, and J. M. Teran, *A second order virtual node method for elliptic problems with interfaces and irregular domains in three dimensions*, J. Comput. Phys. **231** (2012), no. 4, 2015–2048. MR Zbl
- [13] H. Ji, F. Lien, and E. Yee, *An efficient second-order accurate cut-cell method for solving the variable coefficient Poisson equation with jump conditions on irregular domains*, Int. J. Numer. Meth. Fluids **52** (2006), no. 7, 723–748. Zbl
- [14] B. Kallemov, A. P. S. Bhalla, B. E. Griffith, and A. Donev, *An immersed boundary method for rigid bodies*, Commun. Appl. Math. Comput. Sci. **11** (2016), no. 1, 79–141. MR Zbl
- [15] M. Kang, R. P. Fedkiw, and X.-D. Liu, *A boundary condition capturing method for multiphase incompressible flow*, J. Sci. Comput. **15** (2000), no. 3, 323–360. MR Zbl
- [16] M.-C. Lai and C. S. Peskin, *An immersed boundary method with formal second-order accuracy and reduced numerical viscosity*, J. Comput. Phys. **160** (2000), no. 2, 705–719. MR Zbl
- [17] M.-C. Lai and H.-C. Tseng, *A simple implementation of the immersed interface methods for Stokes flows with singular forces*, Comput. & Fluids **37** (2008), no. 2, 99–106. MR Zbl
- [18] L. Lee and R. J. LeVeque, *An immersed interface method for incompressible Navier–Stokes equations*, SIAM J. Sci. Comput. **25** (2003), no. 3, 832–856. MR Zbl
- [19] R. J. LeVeque and Z. Li, *The immersed interface method for elliptic equations with discontinuous coefficients and singular sources*, SIAM J. Numer. Anal. **31** (1994), no. 4, 1019–1044. MR Zbl

- [20] Z. Li, *A note on immersed interface method for three-dimensional elliptic equations*, Comput. Math. Appl. **31** (1996), no. 3, 9–17. MR Zbl
- [21] Z. Li and K. Ito, *Maximum principle preserving schemes for interface problems with discontinuous coefficients*, SIAM J. Sci. Comput. **23** (2001), no. 1, 339–361. MR Zbl
- [22] Z. Li and M.-C. Lai, *The immersed interface method for the Navier–Stokes equations with singular forces*, J. Comput. Phys. **171** (2001), no. 2, 822–842. MR Zbl
- [23] X.-D. Liu, R. P. Fedkiw, and M. Kang, *A boundary condition capturing method for Poisson’s equation on irregular domains*, J. Comput. Phys. **160** (2000), no. 1, 151–178. MR Zbl
- [24] X.-D. Liu and T. C. Sideris, *Convergence of the ghost fluid method for elliptic equations with interfaces*, Math. Comp. **72** (2003), no. 244, 1731–1746. MR Zbl
- [25] S. H. Lui, *Numerical analysis of partial differential equations*, Wiley, Hoboken, NJ, 2011. MR Zbl
- [26] A. N. Marques, J.-C. Nave, and R. R. Rosales, *A correction function method for Poisson problems with interface jump conditions*, J. Comput. Phys. **230** (2011), no. 20, 7567–7597. MR Zbl
- [27] ———, *High order solution of Poisson problems with piecewise constant coefficients and interface jumps*, J. Comput. Phys. **335** (2017), 497–515. MR Zbl
- [28] Y. Mori and C. S. Peskin, *Implicit second-order immersed boundary methods with boundary mass*, Comput. Methods Appl. Mech. Engrg. **197** (2008), no. 25–28, 2049–2067. MR Zbl
- [29] C. S. Peskin, *Flow patterns around heart valves: a numerical method*, J. Comput. Phys. **10** (1972), no. 2, 252–271. Zbl
- [30] ———, *Numerical analysis of blood flow in the heart*, J. Comput. Phys. **25** (1977), no. 3, 220–252. MR Zbl
- [31] ———, *The immersed boundary method*, Acta Numer. **11** (2002), 479–517. MR Zbl
- [32] D. Russell and Z. J. Wang, *A Cartesian grid method for modeling multiple moving objects in 2D incompressible viscous flow*, J. Comput. Phys. **191** (2003), no. 1, 177–205. MR Zbl
- [33] L. M. Smith and S. N. Stechmann, *Precipitating quasigeostrophic equations and potential vorticity inversion with phase changes*, J. Atmos. Sci. **74** (2017), no. 10, 3285–3303.
- [34] M. Sussman, E. Fatemi, P. Smereka, and S. Osher, *An improved level set method for incompressible two-phase flows*, Comput. & Fluids **27** (1998), no. 5–6, 663–680. Zbl
- [35] M. Sussman, K. M. Smith, M. Y. Hussaini, M. Ohta, and R. Zhi-Wei, *A sharp interface method for incompressible two-phase flows*, J. Comput. Phys. **221** (2007), no. 2, 469–505. MR Zbl
- [36] A. Wiegmann and K. P. Bube, *The explicit-jump immersed interface method: finite difference methods for PDEs with piecewise smooth solutions*, SIAM J. Numer. Anal. **37** (2000), no. 3, 827–862. MR Zbl
- [37] S. Xu and Z. J. Wang, *An immersed interface method for simulating the interaction of a fluid with moving boundaries*, J. Comput. Phys. **216** (2006), no. 2, 454–493. MR Zbl
- [38] ———, *Systematic derivation of jump conditions for the immersed interface method in three-dimensional flow simulation*, SIAM J. Sci. Comput. **27** (2006), no. 6, 1948–1980. MR Zbl

Received August 1, 2018. Revised March 12, 2019.

CHUNG-NAN TZOU: ctzou@wisc.edu

Department of Mathematics, University of Wisconsin–Madison, Madison, WI, United States

SAMUEL N. STECHMANN: stechmann@wisc.edu

Department of Mathematics, University of Wisconsin–Madison, Madison, WI, United States