A SIMPLE AND EFFICIENT CONVEX OPTIMIZATION BASED BOUND-PRESERVING HIGH ORDER ACCURATE LIMITER FOR CAHN–HILLIARD–NAVIER–STOKES SYSTEM [*]

CHEN LIU[†], BEATRICE RIVIERE[‡], JIE SHEN [§], AND XIANGXIONG ZHANG [†]

Abstract. For time-dependent PDEs, the numerical schemes can be rendered bound-preserving without losing conservation and accuracy, by a post processing procedure of solving a constrained minimization in each time step. Such a constrained optimization can be formulated as a nonsmooth convex minimization, which can be efficiently solved by first order optimization methods, if using the optimal algorithm parameters. By analyzing the asymptotic linear convergence rate of the generalized Douglas–Rachford splitting method, optimal algorithm parameters can be approximately expressed as a simple function of the number of out-of-bounds cells. We demonstrate the efficiency of this simple choice of algorithm parameters by applying such a limiter to cell averages of a discontinuous Galerkin scheme solving phase field equations for 3D demanding problems. Numerical tests on a sophisticated 3D Cahn–Hilliard–Navier–Stokes system indicate that the limiter is high order accurate, very efficient, and well-suited for large-scale simulations. For each time step, it takes at most 20 iterations for the Douglas–Rachford splitting to enforce bounds and conservation up to the round-off error, for which the computational cost is at most $80N$ with N being the total number of cells.

Key words. Douglas–Rachford splitting, nearly optimal parameters, bound-preserving limiter, discontinuous Galerkin method, Cahn–Hilliard–Navier–Stokes, high order accuracy

MSC codes. 65K10, 65M60, 65M12, 90C25

1. Introduction.

1.1. Objective and motivation. We are interested in a simple approach to enforce bound-preserving property of a high order accurate scheme for phase field models, without destroying conservation and accuracy. Many numerical methods, especially high order accurate schemes, do not preserve bounds. For the sake of both physical meaningfulness and robustness of numerical computation, it is critical to enforce both conservation and bounds.

Bound-preserving schemes have been well studied in the literature for equations like hyperbolic and parabolic PDEs. One popular approach of constructing a bound-preserving high order scheme was introduced in [44, 45] for conservation laws, which can be extended to parabolic equations [40, 39] and Navier–Stokes equations [12, 43], as well as implicit or semi-implicit time discretizations [35, 31]. However, this method, and most of other popular bound-preserving schemes for conservation laws and parabolic equations such as exponential time differencing [10], are based on the fact that the simplest low order scheme is bound-preserving, which is no longer true for a fourth order PDE like the Cahn–Hilliard equation, unless a very special implementation is used such as implicit treatment of a logarithmic potential [6].

A simple cut-off without enforcing conservation does not destroy accuracy but it is of little interest, because convergence might be lost due to loss of conservation. A meaningful objective is to enforce bounds without destroying conservation. For the Cahn–Hilliard equation, an exponential function transform approach was used in

[†]Department of Mathematics, Purdue University, 150 North University Street, West Lafayette, Indiana 47907 (liu3373@purdue.edu, zhan1966@purdue.edu).

[‡]Department of Computational Applied Mathematics and Operations Research, Rice University, 6100 Main Street, Houston, Texas 77005 (riviere@rice.edu).

[§] Eastern Institute of Technology, Ningbo, Zhejiang 315200, P.R. China (jshen@eitech.edu.cn).

[23], with conservation achieved up to some small time error. If the logarithmic energy potential is used and treated implicitly, bounds can also be ensured [6]. A Lagrange multiplier approach in [7, 8] provides a new interpretation for the cut-off method, and can preserve mass by solving a nonlinear algebraic equation for the additional space independent Lagrange multiplier. Even though the flux limiting [25, 42, 22, 11] can be formally extended to Cahn–Hilliard equation [17, 30], it is not clear whether flux limiters can preserve high order accuracy for a fourth order PDE. Recently a bound-preserving finite volume scheme, which is first order accurate in time and second order accurate in space, has been constructed for the Cahn–Hilliard equation [1].

In practice, the logarithmic potential causes additional difficulty in nonlinear system solvers in many schemes, thus the double well polynomial potential with a degenerate mobility is often used as an easier surrogate. With the double well potential, numerical schemes might violate the bounds much more since it does not enforce bounds $\phi \in [-1, 1]$ like the log potential. In this paper, we will explore a simple and efficient high order accurate post processing procedure for preserving bounds and conservation up to round-off errors, such that it can be easily applied to any numerical method solving the Cahn–Hillard equation, especially for the polynomial potential.

1.2. A bound-preserving limiter via convex minimization. Consider a scalar PDE as an example. Assume its solution u satisfies $m \leq u \leq M$ for all time and location, where m and M are constant bounds. For simplicity, we only consider enforcing cell averages in a high order accurate discontinuous Galerkin (DG) scheme by the convex minimization, then using the simple Zhang–Shu limiter in [44, 45] to enforce bounds of point values of the DG solution. But this convex minimization approach can be easily extended to enforcing bounds of point values for any other numerical scheme such as finite difference and continuous finite element methods.

Let $\bar{u}_i$ $(i = 1, \cdots, N)$ be all the DG solution cell averages at time step n on a uniform mesh. Given $\boldsymbol{u} = \begin{bmatrix} \bar{u}_1 & \bar{u}_2 & \cdots & \bar{u}_N \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^N$, we would like to post process it to $\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^N$ such that it is bound preserving $x_i \in [m, M]$, conservative $\sum_i x_i = \sum_i \bar{u}_i$, and accurate in the sense that $\|\boldsymbol{x} - \boldsymbol{u}\|$ should be small. Namely, we consider minimize $\|\boldsymbol{x} - \boldsymbol{u}\|$ under constraints $x_i \in [m, M]$ and $\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} \bar{u}_i$. To change as few cell averages as possible, the convex ℓ^1-norm is often used to approximate the NP-hard ℓ^0-norm. The ℓ^1-norm is nonsmooth without any strong convexity, thus the minimization might still be too expensive to solve. For the sake of efficiency, we propose the ℓ^2-norm instead:

$$(1.1) \qquad \min_{\boldsymbol{x}} \|\boldsymbol{x} - \boldsymbol{u}\|_2^2 \quad \text{s.t.} \quad x_i \in [m, M] \ \text{ and } \ \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} \bar{u}_i.$$

Obviously, the minimizer to (1.1) is conservative and bound-preserving. The justification of accuracy is also straightforward, as long as $\boldsymbol{u}$ is an accurate numerical solution, which is a reasonable assumption and has been proved to hold for many DG schemes of a variety of PDEs, e.g., see [29] for Cahn–Hilliard–Navier–Stokes (CHNS) equations. Let $\bar{u}_i^*$ and $\bar{u}_i^0$ be the cell averages of the exact solution at time t^n and initial condition, respectively. Then $\sum_i \bar{u}_i^* = \sum_i \bar{u}_i^0 = \sum_i \bar{u}_i$ and $\bar{u}_i^* \in [m, M]$ imply that $\boldsymbol{u}^*$ is a feasible point satisfying the constraints of (1.1). The minimizer $\boldsymbol{x}^*$ to (1.1) then satisfies $\|\boldsymbol{x}^* - \boldsymbol{u}\|_2 \leq \|\boldsymbol{u}^* - \boldsymbol{u}\|_2$, thus $\|\boldsymbol{x}^* - \boldsymbol{u}^*\|_2 \leq \|\boldsymbol{x}^* - \boldsymbol{u}\|_2 + \|\boldsymbol{u} - \boldsymbol{u}^*\|_2 \leq 2\|\boldsymbol{u}^* - \boldsymbol{u}\|_2$. Therefore, the limiter (1.1) does not lose the order of accuracy.

1.3. Efficient convex optimization algorithms. The main catch of using (1.1) in a large scale computation, is the possible huge cost of solving (1.1) to machine accuracy, unless proven or shown otherwise, which is our main focus. It is a convention use the indicator function $\iota_\Omega(x) = \begin{cases} 0, & x \in \Omega \\ +\infty, & x \notin \Omega \end{cases}$ for any set Ω, to rewrite (1.1) as:

$$(1.2) \qquad \min_x \frac{\alpha}{2}\|x - u\|_2^2 + \iota_{\Lambda_1}(x) + \iota_{\Lambda_2}(x),$$

where $\alpha > 0$ is a parameter and the sets Λ_1 and Λ_2 are $\Lambda_1 = \{x : \sum_i x_i = \sum_i \bar{u}_i\}$, $\Lambda_2 = \{x : x_i \in [m, M]\}$. The two indicator functions in (1.2) are convex but nonsmooth, and the ℓ^2 term is strongly convex, thus (1.2) has a unique minimizer x^*. Many optimization algorithms, e.g., fast proximal gradient (FISTA) [34, 3] applied to (1.2), can be proven to converge linearly. But a provable global linear rate is usually quite pessimistic, much slower than the actual convergence rate. It is possible to obtain sharp asymptotic rate for methods like the generalized Douglas–Rachford splitting solving ℓ^1 minimization [9], which can be used for designing best parameters. So we consider the generalized Douglas–Rachford splitting [26], which is equivalent to some other popular methods such as PDHG [5], ADMM [13], dual split Bregman [20], see also [9] and references therein for the equivalence.

1.4. The generalized Douglas–Rachford splitting method. Splitting algorithms naturally arise for composite optimization of the form

$$(1.3a) \qquad \min_x f(x) + g(x),$$

where functions f and g are convex and have simple subdifferentials and resolvents. Let ∂f and ∂g denote the subdifferentials of f and g. Their resolvents are defined as

$$J_{\gamma \partial f} = (I + \gamma \partial f)^{-1} = \text{argmin}_z \gamma f(z) + \frac{1}{2}\|z - x\|_2^2, \quad \gamma > 0,$$

$$J_{\gamma \partial g} = (I + \gamma \partial g)^{-1} = \text{argmin}_z \gamma g(z) + \frac{1}{2}\|z - x\|_2^2, \quad \gamma > 0.$$

We rewrite (1.2) into $\min_x f(x) + g(x)$ by defining

$$(1.3b) \qquad f(x) = \frac{\alpha}{2}\|x - u\|_2^2 + \iota_{\Lambda_1}(x) \quad \text{and} \quad g(x) = \iota_{\Lambda_2}(x),$$

where two sets are $\Lambda_1 = \{x : \mathbf{A}x = b\}$ and $\Lambda_2 = \{x : m \le x \le M\}$, with $\mathbf{A} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$, $b = \sum_i \bar{u}_i$, and $m \le x \le M$ denoting entrywise inequality. The subdifferentials and resolvents can be explicitly given as

$$(1.4) \quad \partial f(x) = \alpha(x - u) + \mathcal{R}(\mathbf{A}^{\mathrm{T}}), \quad J_{\gamma \partial f}(x) = \frac{1}{\gamma \alpha + 1}\left(\mathbf{A}^+(b - \mathbf{A}x) + x\right) + \frac{\gamma \alpha}{\gamma \alpha + 1}u,$$

$$(1.5) \quad [\partial g(x)]_i = \begin{cases} [0, +\infty], & \text{if } x_i = M, \\ 0, & \text{if } x_i \in (m, M), \\ [-\infty, 0], & \text{if } x_i = m. \end{cases} \quad [J_{\gamma \partial g}(x)]_i = \min\left(\max\left(x_i, m\right), M\right),$$

where $\mathcal{R}(\mathbf{A}^{\mathrm{T}})$ denotes the range of the matrix $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^+ = \mathbf{A}^{\mathrm{T}}(\mathbf{A}\mathbf{A}^{\mathrm{T}})^{-1}$.

Define reflection operators as $R_{\gamma \partial f} = 2J_{\gamma \partial f} - I$ and $R_{\gamma \partial g} = 2J_{\gamma \partial g} - I$, where I denotes the identity operator. The generalized Douglas–Rachford splitting for (1.3a) can be written as:

$$(1.6) \quad \begin{cases} \boldsymbol{y}^{k+1} = \lambda \dfrac{R_{\gamma \partial f} R_{\gamma \partial g} + I}{2} \boldsymbol{y}^k + (1 - \lambda) \boldsymbol{y}^k = \lambda J_{\gamma \partial f} \circ (2J_{\gamma \partial g} - I) \boldsymbol{y}^k + (I - \lambda J_{\gamma \partial g}) \boldsymbol{y}^k \\ \boldsymbol{x}^{k+1} = J_{\gamma \partial g}(\boldsymbol{y}^{k+1}) \end{cases}.$$

where $\boldsymbol{y}$ is an auxiliary variable, $\gamma > 0$ is step size, and $\lambda \in (0, 2)$ is a parameter. For two convex functions $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$, (1.6) converges for any $\gamma > 0$ and any fixed $\lambda \in (0, 2)$, see [26]. If one function is strongly convex, then $\lambda = 2$ also converges.

1.5. The bound-preserving post processing procedure for DG schemes.
At time step n, let $u_i(x, y, z)$ be the DG polynomial on a uniform mesh in the i-th cell with cell average $\bar{u}_i$. We define the following bound-preserving limiter:

Step I: Solve (1.2) to post process the cell averages. Let $c = \frac{1}{\alpha \gamma + 1}$, then the iteration (1.6) on (1.3) can be explicitly written as:

$$(1.7a) \quad \begin{cases} \boldsymbol{x}^k & = \min(\max(\boldsymbol{y}^k, m), M) \\ \boldsymbol{z}^k & = 2\boldsymbol{x}^k - \boldsymbol{y}^k \\ \boldsymbol{y}^{k+1} & = \lambda c(\boldsymbol{z}^k - \mathbf{1}[\frac{1}{N}(\sum_i z_i^k - b)]) + \lambda(1 - c)\boldsymbol{u} + \boldsymbol{y}^k - \lambda \boldsymbol{x}^k \end{cases},$$

where $\mathbf{1}$ is the constant one vector of size N and $b = \sum_i \bar{u}_i$ is a constant, $\lambda \in (0, 2]$ is the fixed relaxation parameter. Each iterate $\boldsymbol{x}^k$ is bound-preserving but is not conservative until converging to the minimizer $\boldsymbol{x}^*$. We iterate (1.7a) until relative change is small enough $\|\boldsymbol{y}^{k+1} - \boldsymbol{y}^k\|_2 \le \epsilon$, to get an approximated minimizer $\boldsymbol{x}^*$ to (1.2), for which the conservation would be satisfied up to round-off errors. We then modify DG polynomials by modifying the cell averages, i.e., shift them by a constant:

$$(1.7b) \quad \widetilde{u}_i(x, y, z) = u_i(x, y, z) - \bar{u}_i + x_i^*, \quad i = 1, \cdots, N.$$

Step II: Cell averages of modified DG polynomials $\widetilde{u}_i(x, y, z)$ are in the range $[m, M]$, so we can apply the simple scaling limiter by Zhang and Shu in [44, 45] to further enforce bounds at quadrature points, without losing conservation and accuracy. Let S_i be the set of interested points in each cell, then the Zhang–Shu limiter for the polynomial $\widetilde{u}_i(x, y, z)$ with cell average $x_i^* \in [m, M]$ is given as

$$(1.8) \quad \widehat{u}_i(x, y, z) = \theta(\widetilde{u}_i(x, y, z) - x_i^*) + x_i^*, \quad \theta = \min\left\{1, \frac{|m - x_i^*|}{|m_i - x_i^*|}, \frac{|M - x_i^*|}{|M_i - x_i^*|}\right\},$$

where $m_i = \min_{(x,y,z) \in S_i} \widetilde{u}_i(x, y, z)$ and $M_i = \max_{(x,y,z) \in S_i} \widetilde{u}_i(x, y, z)$. See the appendix in [43] for a rigorous proof of the high order accuracy of (1.8).

We emphasize that the Zhang-Shu limiter (1.8) can preserve bounds or positivity provided that the cell averages are within bounds or are positive, which can be proven for DG methods coupled with the limiter (1.8) for hyperbolic problems including scalar conservation laws, compressible Euler and compressible Navier-Stokes equations [44, 45, 43], because DG methods with suitable numerical fluxes satisfy a weak monotonicity property for these problems [43]. However, such a weak monotonicity property is simply not true for high order DG schemes solving fourth order PDEs. Thus, if using only the limiter (1.8), the high order DG methods will not be bound-preserving for Cahn-Hilliard equations. For all the numerical tests shown in this paper, DG methods with only the Zhang-Shu limiter will produce cell averages outside of the range $[-1, 1]$.

1.6. The main results. We will analyze asymptotic convergence rate of iteration (1.7a) and give a sharp convergence rate formula, by which it is possible to pick up nearly optimal combination of parameters $c = \frac{1}{\alpha\gamma+1}$ and λ to achieve fast convergence for the iteration (1.7a). The asymptotic linear convergence rate we derive for (1.2) is similar to the one for ℓ^1 minimization in [9]. These rate formulae depend on the unknown $\boldsymbol{x}^*$, so usually it is impossible to use the formulae for tuning algorithm parameters, unless $\boldsymbol{x}^*$ can be easily estimated. For (1.2), it is possible to pick up a nearly optimal combination of optimization algorithm parameters by only calculating number of bad cells $\bar{u}_i \notin [m, M]$, which is the first main result of this paper.

Let $\hat{r}$ be the number of bad cells $\bar{u}_i \notin [m, M]$, and let $\hat{\theta} = \cos^{-1}\sqrt{\frac{\hat{r}}{N}}$, then our analysis suggests the following simple choice of nearly optimal parameters:

$$(1.9) \quad \begin{cases} c = \frac{1}{2}, \lambda = \frac{4}{2-\cos(2\hat{\theta})}, & \text{if } \hat{\theta} \in (\frac{3}{8}\pi, \frac{1}{2}\pi], \\ c = \frac{1}{(\cos\hat{\theta}+\sin\hat{\theta})^2}, \lambda = \frac{2}{1+\frac{1}{1+\cot\hat{\theta}}-\frac{1}{(\cos\hat{\theta}+\sin\hat{\theta})^2}}, & \text{if } \hat{\theta} \in (\frac{1}{4}\pi, \frac{3}{8}\pi], \\ c = \frac{1}{(\cos\hat{\theta}+\sin\hat{\theta})^2}, \lambda = 2, & \text{if } \hat{\theta} \in (0, \frac{1}{4}\pi]. \end{cases}$$

We emphasize that both c and λ should be the constants w.r.t. iteration index k in (1.7a), once they are chosen by (1.9). Notice that $\lambda(1-c)\boldsymbol{u}$ is a constant for the iteration (1.7a) and each entry of $\boldsymbol{z}^k - \mathbf{1}[\frac{1}{N}(\sum_i z_i^k - b)]$ can be computed by $z_i^k - [\frac{1}{N}(\sum_i z_i^k - b)]$, thus if only counting number of computing multiplications, min, and max, the computational complexity of each iteration in (1.7a) is $4N$. By using the parameters (1.9), it takes at most 20 iterations of (1.7a) to converge in all our numerical tests, thus the cost of iterating (1.7a) until convergence would be at most $80N$, which is highly efficient and well-suited for large-scale simulations.

The numerical observation of at most 20 iterations can also be explained by the asymptotic convergence rate analysis, which is another main result. Assuming the number of bad cells $\bar{u}_i \notin [m, M]$ is much smaller than the number of total cells N, we will show that the asymptotic convergence rate of (1.7a) using (1.9) is given by

$$(1.10) \quad -\frac{\cos(2\theta)}{2-\cos(2\theta)} \approx -\frac{\cos(2\hat{\theta})}{2-\cos(2\hat{\theta})} = \frac{1-2\cos\hat{\theta}^2}{3-2\cos\hat{\theta}^2} = \frac{1-2\frac{\hat{r}}{N}}{3-2\frac{\hat{r}}{N}} \approx \frac{1}{3}, \quad \text{if } \hat{r} \ll N,$$

with $\theta(\boldsymbol{x}^*)$ being an unknown angle, which can be approximated by $\hat{\theta}$. If the ratio of bad cells is very small, (1.7a) will have a local convergence rate almost like $\|\boldsymbol{y}^k - \boldsymbol{y}^*\| \le C\left(\frac{1}{3}\right)^k$, which would take around 30 iterations to reach around 1E-15 if $C = 1$.

1.7. Organization of the paper. The rest of the paper is organized as follows. In Section 2, we analyze the asymptotic linear convergence rate of the Douglas–Rachford splitting (1.6) and (1.7a), and derive the parameter guideline (1.9). In Section 3, we discuss an application of our bound-preserving limiting strategy to an important phase-field model, the CHNS system. The numerical tests are given in Section 4. Section 5 are concluding remarks.

2. Asymptotic linear convergence rate analysis. In this section, we derive the asymptotic linear convergence rate of generalized Douglas–Rachford splitting (1.6) for solving the minimization problem (1.3). The discussion in this section follows closely the analysis for ℓ^1 minimization in [9]. Even though ℓ^1 minimization is harder than ℓ^2 minimization, the analysis for (1.3) is not necessarily a straightforward extension of those in [9] because (1.4) and (1.5) are different from operators in [9].

For convenience, let $F = \partial f$ and $G = \partial g$ denote the subdifferential operators. Let $S(\boldsymbol{x})$ be the cut-off operator, i.e., $[J_{\gamma G}(\boldsymbol{x})]_i = [S(\boldsymbol{x})]_i = \min(\max(x_i, m), M)$.

We keep the discussion a bit more general by considering a general linear constraint $\mathbf{A}\boldsymbol{x} = b = \mathbf{A}\boldsymbol{u}$ in the function $f(\boldsymbol{x})$ in (1.3b), and assume $\mathbf{A}$ has less number of rows than the number of columns, with full row rank such that $\mathbf{A}^+ = \mathbf{A}^{\mathrm{T}}(\mathbf{A}\mathbf{A}^{\mathrm{T}})^{-1}$ is well defined. When needed, we will plug in the special case $\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$.

2.1. The fixed point set. Let $P(\boldsymbol{x}) = \mathbf{A}^+(b - \mathbf{A}\boldsymbol{x}) + \boldsymbol{x}$ denote the projection operator. Then, the resolvents can be written as $J_{\gamma F}(\boldsymbol{x}) = \frac{1}{\gamma\alpha+1}P(\boldsymbol{x}) + \frac{\gamma\alpha}{\gamma\alpha+1}\boldsymbol{u}$ and $J_{\gamma G}(\boldsymbol{x}) = S(\boldsymbol{x})$. Let T_γ denote the iteration operator for $\boldsymbol{y}$ in (1.6), then it becomes:

$$(2.1) \qquad T_\gamma = \frac{\lambda}{\gamma\alpha+1}P \circ (2S - I) + (I - \lambda S) + \frac{\lambda\gamma\alpha}{\gamma\alpha+1}\boldsymbol{u}.$$

The fixed point $\boldsymbol{y}^*$ of T_γ is not the minimizer of (1.3), while $\boldsymbol{x}^* = J_{\gamma G}(\boldsymbol{y}^*) = S(\boldsymbol{y}^*)$ is the minimizer. The fixed point set of the operator T_γ has the following structure.

THEOREM 2.1. *The set of fixed point of operator* T_γ *is*

$$\Pi = \{\boldsymbol{y}^* : \boldsymbol{y}^* = \boldsymbol{x}^* + \gamma\boldsymbol{\eta}, \ \boldsymbol{\eta} \in -\partial f(\boldsymbol{x}^*) \cap \partial g(\boldsymbol{x}^*)\}.$$

Proof. We first show any $\boldsymbol{y}^* \in \Pi$ is a fixed point of the operator T_γ. $\forall \boldsymbol{\eta} \in \partial g(\boldsymbol{x}^*)$ in (1.5), we have $S(\boldsymbol{y}^*) = \boldsymbol{x}^*$, since the i-th entry of the vector $\boldsymbol{y}^* = \boldsymbol{x}^* + \gamma\boldsymbol{\eta}$ satisfies

$$[\boldsymbol{y}^*]_i \begin{cases} \in [M, +\infty], & \text{if } x_i^* = M, \\ = x_i^*, & \text{if } x_i^* \in (m, M), \\ \in [-\infty, m], & \text{if } x_i^* = m. \end{cases}$$

Thus, we have $P \circ (2S - I)\boldsymbol{y}^* = P(2\boldsymbol{x}^* - \boldsymbol{y}^*) = P(\boldsymbol{x}^* - \gamma\boldsymbol{\eta}) = \boldsymbol{x}^* - \gamma\boldsymbol{\eta} + \gamma\mathbf{A}^+\mathbf{A}\boldsymbol{\eta}$, where $\mathbf{A}\boldsymbol{x}^* = b$ is used. And $\boldsymbol{\eta} \in -\partial f(\boldsymbol{x}^*)$ in (1.4) implies that there exists $\boldsymbol{\xi}$ such that $\boldsymbol{\eta} = -\alpha(\boldsymbol{x}^* - \boldsymbol{u}) + \mathbf{A}^{\mathrm{T}}\boldsymbol{\xi}$. Multiplying both sides by $\mathbf{A}$, with $\mathbf{A}\boldsymbol{x}^* = b = \mathbf{A}\boldsymbol{u}$ we get $\mathbf{A}\boldsymbol{\eta} = \mathbf{A}\mathbf{A}^{\mathrm{T}}\boldsymbol{\xi}$, thus $\boldsymbol{\xi} = (\mathbf{A}\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{A}\boldsymbol{\eta}$ and $\gamma\boldsymbol{\eta} = -\gamma\alpha(\boldsymbol{x}^* - \boldsymbol{u}) + \gamma\mathbf{A}^+\mathbf{A}\boldsymbol{\eta}$. Then, we have $P \circ (2S - I)\boldsymbol{y}^* = (\gamma\alpha + 1)\boldsymbol{x}^* - \gamma\alpha\boldsymbol{u}$. Therefore

$$T_\gamma(\boldsymbol{y}^*) = \frac{\lambda}{\gamma\alpha+1}\Big((\gamma\alpha + 1)\boldsymbol{x}^* - \gamma\alpha\boldsymbol{u}\Big) + \boldsymbol{y}^* - \lambda\boldsymbol{x}^* + \frac{\lambda\gamma\alpha}{\gamma\alpha+1}\boldsymbol{u} = \boldsymbol{y}^*.$$

Next, we show any fixed point $\boldsymbol{y}^*$ belongs to set Π. Let $\boldsymbol{\eta} = (\boldsymbol{y}^* - \boldsymbol{x}^*)/\gamma$. Then, $\boldsymbol{y}^*$ being a fixed point implies $J_{\gamma G}(\boldsymbol{y}^*) = \boldsymbol{x}^*$. Recall that $J_{\gamma G} = S$, we have

    i. if $x_i^* + \gamma\eta_i \geq M$, then $x_i^* = S(x_i^* + \gamma\eta_i) = M$, thus $\eta_i \in [0, +\infty]$;
    ii. if $x_i^* + \gamma\eta_i \in (m, M)$, then $x_i^* = S(x_i^* + \gamma\eta_i) = x_i^* + \gamma\eta_i$, thus $\eta_i = 0$;
    iii. if $x_i^* + \gamma\eta_i \leq m$, then $x_i^* = S(x_i^* + \gamma\eta_i) = m$, thus $\eta_i \in [-\infty, 0]$.

So $\boldsymbol{\eta} \in \partial g(\boldsymbol{x}^*)$. And $\boldsymbol{y}^* = T_\gamma(\boldsymbol{y}^*)$ is equivalent to $\boldsymbol{y}^* = \frac{\lambda}{2}(R_{\gamma F}R_{\gamma G} + I)\boldsymbol{y}^* + (1 - \lambda)\boldsymbol{y}^*$, which implies $\boldsymbol{y}^* = R_{\gamma F}R_{\gamma G}(\boldsymbol{y}^*)$. Recall $J_{\gamma G}(\boldsymbol{y}^*) = \boldsymbol{x}^*$ and $\boldsymbol{y}^* = \boldsymbol{x}^* + \gamma\boldsymbol{\eta}$, we have

$$\boldsymbol{y}^* = R_{\gamma F}(2J_{\gamma G}(\boldsymbol{y}^*) - \boldsymbol{y}^*) = R_{\gamma F}(\boldsymbol{x}^* - \gamma\boldsymbol{\eta}) = 2J_{\gamma F}(\boldsymbol{x}^* - \gamma\boldsymbol{\eta}) - (\boldsymbol{x}^* - \gamma\boldsymbol{\eta}).$$

So $\boldsymbol{x}^* = J_{\gamma F}(\boldsymbol{x}^* - \gamma\boldsymbol{\eta})$, which implies $\boldsymbol{x}^* = \operatorname{argmin}_{\boldsymbol{z}} \gamma f(\boldsymbol{z}) + \frac{1}{2}\|\boldsymbol{z} - (\boldsymbol{x}^* - \gamma\boldsymbol{\eta})\|_2^2$. By the critical point equation, we have $\boldsymbol{0} \in \gamma\partial f(\boldsymbol{x}^*) + \gamma\boldsymbol{\eta}$ thus $\boldsymbol{\eta} \in -\partial f(\boldsymbol{x}^*)$.   □

Let $\mathcal{B}_r(z)$ denote a closed ball in ℓ^2-norm centered at z with radius r. Define set Q:

$$Q = Q_1 \otimes Q_2 \otimes \cdots \otimes Q_n, \quad \text{where } Q_i = \begin{cases} [M, +\infty], & \text{if } x_i^* = M, \\ (m, M), & \text{if } x_i^* \in (m, M), \\ [-\infty, m], & \text{if } x_i^* = m. \end{cases}$$

For any fixed point y^*, the Theorem 2.1 implies there exists an $\eta = \frac{1}{\gamma}(y^* - x^*) \in \partial g(x^*)$ and by (1.5) we have $x^* + \gamma\eta \in Q$ for any $\gamma > 0$, which gives $y^* \in Q$. Let $\epsilon \geq 0$ be the least upper bound such that $\mathcal{B}_\epsilon(y^*) \subset Q$. If $\epsilon > 0$, then y^* is an interior fixed point and we call this the standard case; otherwise, y^* is a boundary fixed point and we call this the non-standard case. In the standard case that the sequence y^k converges to an interior fixed point y^*. There exists a large enough integer $K > 0$ such that $\|y^K - y^*\|_2 < \epsilon$ holds. For any $k \geq K$, the operator T_γ is nonexpansive [26], so

$$\|y^k - y^*\|_2 = \|\mathrm{T}_\gamma(y^{k-1}) - \mathrm{T}_\gamma(y^*)\|_2 \leq \|y^{k-1} - y^*\|_2 \leq \cdots \leq \|y^K - y^*\|_2 < \epsilon.$$

Thus, after taking the generalized Douglas–Rachford iteration (1.6) sufficiently many times, the iterates will always belong to the ball $\mathcal{B}_\epsilon(y^*) \subset Q$, namely the iteration enters the asymptotic convergence regime and the cut-off location does not change.

In the rest of this paper, we only focus on the standard case. The non-standard case can be analyzed by utilizing the same technique as in [9]. The non-standard case has not been observed in our numerical experiments.

2.2. The characterization of the operator T_γ. Assume the unique solution x^* of the minimization problem (1.3) has r components equal to m or M. We further assume $r < N$, e.g., not all the cell averages will touch the boundary m or M, which is a quite reasonable assumption. We emphasize that r is unknown, unless x^* is given.

Let e_i $(i = 1, \cdots, N)$ be the standard basis of $\mathbb{R}^N$. Let e_j $(j = i_1, \cdots, i_r)$ denote the basis vectors corresponding to entries x^* of being m or M. Let $\mathbf{B}$ be the corresponding $r \times N$ selector matrix, i.e., $\mathbf{B} = [e_{i_1}, \cdots, e_{i_r}]^{\mathrm{T}}$.

Recall that we only discuss the standard case, i.e., y^* is in the interior of Q. Then, in the asymptotic convergence regime, i.e., after sufficiently many iterations, the iterate y_k will stay in the interior of Q, thus the operator S has an expression

$$(2.2) \qquad \mathrm{S}(y) = y - \mathbf{B}^+\mathbf{B}y + \sum_{j \in \{i_1, \cdots, i_r\}} x_j^* e_j.$$

Note, the j-th component of x^*, namely the x_j^* in (2.2), takes value m or M for any $j \in \{i_1, \cdots, i_r\}$. Let $\mathbf{I}_N$ denote an $N \times N$ identity matrix.

LEMMA 2.2. *For any y in the interior of Q, and a standard fixed point y^* in the interior of Q, we have $\mathrm{T}_\gamma(y) - \mathrm{T}_\gamma(y^*) = \mathbf{T}_{c,\lambda}(y - y^*)$, where the matrix $\mathbf{T}_{c,\lambda}$ is given by*

$$\mathbf{T}_{c,\lambda} = \lambda\Big(c(\mathbf{I}_N - \mathbf{A}^+\mathbf{A})(\mathbf{I}_N - \mathbf{B}^+\mathbf{B}) + c\mathbf{A}^+\mathbf{A}\mathbf{B}^+\mathbf{B} + (1-c)\mathbf{B}^+\mathbf{B}\Big) + (1-\lambda)\mathbf{I}_N.$$

Here, $c = \frac{1}{\gamma\alpha+1}$ is a constant in $(0, 1)$.

Proof. By (2.2), $S(\boldsymbol{y}) - S(\boldsymbol{y}^*) = (\mathbf{I}_N - \mathbf{B}^+\mathbf{B})(\boldsymbol{y} - \boldsymbol{y}^*)$. So by (2.1),

$$
\begin{aligned}
T_\gamma(\boldsymbol{y}) - T_\gamma(\boldsymbol{y}^*) &= \frac{\lambda}{\gamma\alpha + 1}\Big(P(2S(\boldsymbol{y}) - \boldsymbol{y}) - P(2S(\boldsymbol{y}^*) - \boldsymbol{y}^*)\Big) + (\boldsymbol{y} - \boldsymbol{y}^*) - \lambda(S(\boldsymbol{y}) - S(\boldsymbol{y}^*)) \\
&= \frac{\lambda}{\gamma\alpha + 1}(\mathbf{I}_N - \mathbf{A}^+\mathbf{A})(\mathbf{I}_N - 2\mathbf{B}^+\mathbf{B})(\boldsymbol{y} - \boldsymbol{y}^*) + (\boldsymbol{y} - \boldsymbol{y}^*) - \lambda(\mathbf{I}_N - \mathbf{B}^+\mathbf{B})(\boldsymbol{y} - \boldsymbol{y}^*) \\
&= \frac{\lambda}{\gamma\alpha + 1}(\mathbf{I}_N - \mathbf{A}^+\mathbf{A})(\mathbf{I}_N - \mathbf{B}^+\mathbf{B})(\boldsymbol{y} - \boldsymbol{y}^*) + \frac{\lambda}{\gamma\alpha + 1}\mathbf{A}^+\mathbf{A}\mathbf{B}^+\mathbf{B}(\boldsymbol{y} - \boldsymbol{y}^*) \\
&\quad + \frac{\lambda\gamma\alpha}{\gamma\alpha + 1}\mathbf{B}^+\mathbf{B}(\boldsymbol{y} - \boldsymbol{y}^*) + (1 - \lambda)(\boldsymbol{y} - \boldsymbol{y}^*).
\end{aligned}
$$

Therefore, the matrix $\mathbf{T}_{c,\lambda}$ can be expressed as follows:

$$
\mathbf{T}_{c,\lambda} = \frac{\lambda}{\gamma\alpha + 1}\Big((\mathbf{I}_N - \mathbf{A}^+\mathbf{A})(\mathbf{I}_N - \mathbf{B}^+\mathbf{B}) + \mathbf{A}^+\mathbf{A}\mathbf{B}^+\mathbf{B}\Big) + \frac{\lambda\gamma\alpha}{\gamma\alpha + 1}\mathbf{B}^+\mathbf{B} + (1 - \lambda)\mathbf{I}_N.
$$

$\square$

DEFINITION 2.3. *Let $\mathcal{U}$ and $\mathcal{V}$ be two subspaces of $\mathbb{R}^N$ with $\dim(\mathcal{U}) = p \leq \dim(\mathcal{V})$. The principal angles $\theta_k \in [0, \frac{\pi}{2}]$ $(k = 1, \cdots, p)$ between $\mathcal{U}$ and $\mathcal{V}$ are recursively defined by*

$$
\cos\theta_k = \boldsymbol{u}_k^{\mathrm{T}}\boldsymbol{v}_k = \max_{\boldsymbol{u}\in\mathcal{U}}\max_{\boldsymbol{v}\in\mathcal{V}}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v},
$$

$$
such\ that\ \ \|\boldsymbol{u}\|_2 = \|\boldsymbol{v}\|_2 = 1,\ \boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{u} = 0,\ \boldsymbol{v}_j^{\mathrm{T}}\boldsymbol{v} = 0,\ j = 1, 2, \cdots, k-1.
$$

The vectors $(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p)$ and $(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p)$ are principal vectors.

Our next goal is to decompose the matrix $\mathbf{T}_{c,\lambda}$ with principal angles between subspaces $\mathcal{N}(\mathbf{A})$ and $\mathcal{N}(\mathbf{B})$. To simplify the writeup, we define matrix $\mathbf{T} = (\mathbf{I}_N - \mathbf{A}^+\mathbf{A})(\mathbf{I}_N - \mathbf{B}^+\mathbf{B}) + \mathbf{A}^+\mathbf{A}\mathbf{B}^+\mathbf{B}$. Thus, we rewrite $\mathbf{T}_{c,\lambda} = \lambda(c\mathbf{T} + (1 - c)\mathbf{B}^+\mathbf{B}) + (1 - \lambda)\mathbf{I}_N$. Let $\mathbf{A}_0$ be an $N \times (N - 1)$ matrix whose columns are orthogonal basis of $\mathcal{N}(\mathbf{A})$ and $\mathbf{A}_1$ be an $N \times 1$ matrix whose columns are orthogonal basis of $\mathcal{R}(\mathbf{A}^{\mathrm{T}})$. Similarly, let $\mathbf{B}_0$ be an $N \times (N - r)$ matrix whose columns are orthogonal basis of $\mathcal{N}(\mathbf{B})$ and $\mathbf{B}_1$ be an $N \times r$ matrix whose columns are orthogonal basis of $\mathcal{R}(\mathbf{B}^{\mathrm{T}})$.

Since both $\mathbf{A}^+\mathbf{A}$ and $\mathbf{A}_1\mathbf{A}_1^{\mathrm{T}}$ represent the projection to $\mathcal{R}(\mathbf{A}^{\mathrm{T}})$, we have $\mathbf{A}^+\mathbf{A} = \mathbf{A}_1\mathbf{A}_1^{\mathrm{T}}$. Similarly, $\mathbf{I}_N - \mathbf{A}^+\mathbf{A} = \mathbf{A}_0\mathbf{A}_0^{\mathrm{T}}$. Thus we have $\mathbf{T} = \mathbf{A}_0\mathbf{A}_0^{\mathrm{T}}\mathbf{B}_0\mathbf{B}_0^{\mathrm{T}} + \mathbf{A}_1\mathbf{A}_1^{\mathrm{T}}\mathbf{B}_1\mathbf{B}_1^{\mathrm{T}}$.

Define matrix $\mathbf{E}_0 = \mathbf{A}_0^{\mathrm{T}}\mathbf{B}_0$ and matrix $\mathbf{E}_1 = \mathbf{A}_1^{\mathrm{T}}\mathbf{B}_0$. Since $\mathbf{A}_0\mathbf{A}_0^{\mathrm{T}} + \mathbf{A}_1\mathbf{A}_1^{\mathrm{T}} = \mathbf{I}_N$, we have $\mathbf{B}_0 = (\mathbf{A}_0\mathbf{A}_0^{\mathrm{T}} + \mathbf{A}_1\mathbf{A}_1^{\mathrm{T}})\mathbf{B}_0 = \mathbf{A}_0\mathbf{E}_0 + \mathbf{A}_1\mathbf{E}_1$. Therefore, we rewrite

$$
(2.3) \quad \mathbf{B}_0\mathbf{B}_0^{\mathrm{T}} = (\mathbf{A}_0\mathbf{E}_0 + \mathbf{A}_1\mathbf{E}_1)(\mathbf{E}_0^{\mathrm{T}}\mathbf{A}_0^{\mathrm{T}} + \mathbf{E}_1^{\mathrm{T}}\mathbf{A}_1^{\mathrm{T}}) = \begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_1 \end{bmatrix} \begin{bmatrix} \mathbf{E}_0\mathbf{E}_0^{\mathrm{T}} & \mathbf{E}_0\mathbf{E}_1^{\mathrm{T}} \\ \mathbf{E}_1\mathbf{E}_0^{\mathrm{T}} & \mathbf{E}_1\mathbf{E}_1^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{A}_0^{\mathrm{T}} \\ \mathbf{A}_1^{\mathrm{T}} \end{bmatrix}.
$$

The singular value decomposition (SVD) of the $(N - 1) \times (N - r)$ matrix $\mathbf{E}_0$ is $\mathbf{E}_0 = \mathbf{U}_0\cos\boldsymbol{\Theta}\mathbf{V}^{\mathrm{T}}$ with singular values $\cos\theta_1, \cdots, \cos\theta_{N-r}$ in nonincreasing order. We know that θ_i $(i = 1, \cdots, N - r)$ are the principal angles between $\mathcal{N}(\mathbf{A})$ and $\mathcal{N}(\mathbf{B})$.

Notice $\mathbf{E}_1^{\mathrm{T}}\mathbf{E}_1 = \mathbf{B}_0^{\mathrm{T}}\mathbf{A}_1\mathbf{A}_1^{\mathrm{T}}\mathbf{B}_0$ and $\mathbf{A}_1\mathbf{A}_1^{\mathrm{T}} = \mathbf{I}_N - \mathbf{A}_0\mathbf{A}_0^{\mathrm{T}}$, we have $\mathbf{E}_1^{\mathrm{T}}\mathbf{E}_1 = \mathbf{B}_0^{\mathrm{T}}\mathbf{B}_0 - \mathbf{B}_0^{\mathrm{T}}\mathbf{A}_0\mathbf{A}_0^{\mathrm{T}}\mathbf{B}_0 = \mathbf{I}_{N-r} - \mathbf{E}_0^{\mathrm{T}}\mathbf{E}_0$. Recall the SVD of $\mathbf{E}_0$, we have $\mathbf{E}_1^{\mathrm{T}}\mathbf{E}_1 = \mathbf{V}\sin^2\boldsymbol{\Theta}\mathbf{V}^{\mathrm{T}}$. Thus $\mathbf{E}_1$ can be expressed as $\mathbf{U}_1\sin\boldsymbol{\Theta}\mathbf{V}^{\mathrm{T}}$, which is however not the SVD of $\mathbf{E}_1$. To this end, let matrix $\widetilde{\mathbf{A}} = [\mathbf{A}_0\mathbf{U}_0\ \mathbf{A}_1\mathbf{U}_1]$, then (2.3) becomes

$$
(2.4) \qquad\qquad \mathbf{B}_0\mathbf{B}_0^{\mathrm{T}} = \widetilde{\mathbf{A}}\begin{bmatrix} \cos^2\boldsymbol{\Theta} & \sin\boldsymbol{\Theta}\cos\boldsymbol{\Theta} \\ \sin\boldsymbol{\Theta}\cos\boldsymbol{\Theta} & \sin^2\boldsymbol{\Theta} \end{bmatrix}\widetilde{\mathbf{A}}^{\mathrm{T}}.
$$

Because of $\mathbf{B}_1\mathbf{B}_1^{\mathrm{T}} = \mathbf{I}_N - \mathbf{B}_0\mathbf{B}_0^{\mathrm{T}}$ and $\widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^{\mathrm{T}} = \mathbf{I}_N$, we have the decomposition

$$(2.5) \qquad \mathbf{B}_1\mathbf{B}_1^{\mathrm{T}} = \widetilde{\mathbf{A}} \begin{bmatrix} \sin^2\mathbf{\Theta} & -\sin\mathbf{\Theta}\cos\mathbf{\Theta} \\ -\sin\mathbf{\Theta}\cos\mathbf{\Theta} & \cos^2\mathbf{\Theta} \end{bmatrix} \widetilde{\mathbf{A}}^{\mathrm{T}}.$$

Notice $\mathbf{A}_0\mathbf{A}_0^{\mathrm{T}}\widetilde{\mathbf{A}} = [\mathbf{A}_0\mathbf{U}_0 \ \mathbf{O}_{N\times(N-r)}]$ and $\mathbf{A}_1\mathbf{A}_1^{\mathrm{T}}\widetilde{\mathbf{A}} = [\mathbf{O}_{N\times(N-r)} \ \mathbf{A}_1\mathbf{U}_1]$, by (2.4) and (2.5), we obtain

$$(2.6) \qquad \mathbf{T} = \widetilde{\mathbf{A}} \begin{bmatrix} \cos^2\mathbf{\Theta} & \sin\mathbf{\Theta}\cos\mathbf{\Theta} \\ -\sin\mathbf{\Theta}\cos\mathbf{\Theta} & \cos^2\mathbf{\Theta} \end{bmatrix} \widetilde{\mathbf{A}}^{\mathrm{T}}.$$

Therefore, use (2.6) and consider $\mathbf{B}^+\mathbf{B} = \mathbf{B}_1\mathbf{B}_1^{\mathrm{T}}$, the matrix $\mathbf{T}_{c,\lambda}$ becomes

$$(2.7) \ \mathbf{T}_{c,\lambda} = \widetilde{\mathbf{A}} \begin{bmatrix} \lambda c\cos^2\mathbf{\Theta} + \lambda(1-c)\sin^2\mathbf{\Theta} + (1-\lambda)\mathbf{I}_{N-r} & \lambda(2c-1)\sin\mathbf{\Theta}\cos\mathbf{\Theta} \\ -\lambda\sin\mathbf{\Theta}\cos\mathbf{\Theta} & \lambda\cos^2\mathbf{\Theta} + (1-\lambda)\mathbf{I}_{N-r} \end{bmatrix} \widetilde{\mathbf{A}}^{\mathrm{T}}.$$

2.3. Asymptotic convergence rate. With the assumption $r < N$, there exists a nonzero principal angle between subspaces $\mathcal{N}(\mathbf{A})$ and $\mathcal{N}(\mathbf{B})$. The following lemma gives values of all the principal angles.

LEMMA 2.4. *The principal angles θ_i, $i = 1, \cdots, N - r$, between subspaces $\mathcal{N}(\mathbf{A})$ and $\mathcal{N}(\mathbf{B})$ satisfy*

$$(2.8) \qquad \cos\theta_1 = \cdots = \cos\theta_{N-r-1} = 1 \quad and \quad \cos\theta_{N-r} = \sqrt{\frac{r}{N}}.$$

Proof. Let $\mathcal{N}(\mathbf{A})^{\perp}$ denote the orthogonal complement of space $\mathcal{N}(\mathbf{A})$. Since $\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{1\times N}$, we have $\mathcal{N}(\mathbf{A})^{\perp} = \mathrm{span}\{\mathbf{1}\}$. Recall the columns of $\mathbf{B}_0$ are the orthogonal basis of $\mathcal{N}(\mathbf{B})$. The principal angles between $\mathcal{N}(\mathbf{A})^{\perp}$ and $\mathcal{N}(\mathbf{B})$ can be computed via the SVD of $\frac{1}{\sqrt{N}}\mathbf{1}^{\mathrm{T}}\mathbf{B}_0$. Each column of $\mathbf{B}_0$ is a standard basis $\mathbf{e}_j$, where $j \neq i_1, \cdots, i_r$. Thus

$$\left(\frac{1}{\sqrt{N}}\mathbf{1}^{\mathrm{T}}\mathbf{B}_0\right)^{\mathrm{T}}\left(\frac{1}{\sqrt{N}}\mathbf{1}^{\mathrm{T}}\mathbf{B}_0\right) = \frac{1}{N}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}_{(N-r)\times(N-r)} .$$

The eigenvalues of the $(N - r) \times (N - r)$ matrix consisting of all ones, are $N - r$ and $0, \cdots, 0$. So the singular values of $\frac{1}{\sqrt{N}}\mathbf{1}^{\mathrm{T}}\mathbf{B}_0$ are $\sqrt{\frac{N-r}{N}}$ and $0, \cdots, 0$. We conclude $\cos\theta_{N-r} = \sqrt{\frac{r}{N}}$, since the non-trivial principal angles between $\mathcal{N}(\mathbf{A})$ and $\mathcal{N}(\mathbf{B})$ and the corresponding non-trivial principal angles between $\mathcal{N}(\mathbf{A})^{\perp}$ and $\mathcal{N}(\mathbf{B})$ sum up to $\frac{\pi}{2}$, see the Theorem 2.7 in [24]. In addition, since the dimension of $\mathcal{N}(\mathbf{A})$ is $N - 1$ and the dimension of $\mathcal{N}(\mathbf{B})$ is $N - r$, then as long as $N - r > 1$, from the definition of principal angles, it is straightforward to see $\cos\theta_1 = \cdots = \cos\theta_{N-r-1} = 1$. $\square$

By Lemma 2.4, there exists only one nonzero principal angle θ_{N-r}. By eliminating zero columns in (2.7), (2.7) can be simplified as

$$\mathbf{T}_{c,\lambda} = [\mathbf{A}_0\mathbf{U}_0 \ \mathbf{A}_1]\begin{bmatrix} \mathbf{0}_{r-1} & & \\ & (1-\lambda+\lambda c)\mathbf{I}_{N-r-1} & \\ & & \begin{matrix} \lambda c\cos^2\theta_{N-r} + \lambda(1-c)\sin^2\theta_{N-r} + (1-\lambda) & \lambda(2c-1)\sin\theta_{N-r}\cos\theta_{N-r} \\ -\lambda\sin\theta_{N-r}\cos\theta_{N-r} & \lambda\cos^2\theta_{N-r} + (1-\lambda) \end{matrix} \end{bmatrix}\begin{bmatrix} \mathbf{U}_0^{\mathrm{T}}\mathbf{A}_0^{\mathrm{T}} \\ \mathbf{A}_1^{\mathrm{T}} \end{bmatrix}.$$

From (2.7) we know the matrix $\mathbf{T}_{c,\lambda}$ is a nonnormal matrix, thus $\|\mathbf{T}_{c,\lambda}^k\|_2$ is significantly smaller than $\|\mathbf{T}_{c,\lambda}\|_2^k$ for sufficiently large k. Therefore, the asymptotic convergence rate is governed by $\lim_{k\to\infty} \|\mathbf{T}_{c,\lambda}^k\|_2^{\frac{1}{k}}$, which is equal to the norm of the eigenvalue of $\mathbf{T}_{c,\lambda}$ with the largest magnitude. We have

$$\det(\mathbf{T}_{c,\lambda} - \rho\mathbf{I}) = (\rho - 1 + \lambda - \lambda c)^{N-r-1}(\rho - 1 + \lambda c)^{r-1}$$
$$\times \left[\rho^2 - (\lambda(c\cos 2\theta_{N-r} - 1) + 2)\rho + \lambda^2 c \sin^2 \theta_{N-r} + \lambda(c\cos 2\theta_{N-r} - 1) + 1\right].$$

By Lemma 2.4, the matrix $\mathbf{T}_{c,\lambda}$ has eigenvalues $\rho_0 = 1 - \lambda c$ and $\rho_1 = 1 - \lambda(1-c)$ corresponding to the principle angles $\theta_1, \cdots, \theta_{N-r-1}$, Corresponding to the principle angle θ_{N-r}, the matrix $\mathbf{T}_{c,\lambda}$ has another two eigenvalues, ρ_2 and ρ_3, satisfying the following quadratic equation:

$$(2.9) \qquad \rho^2 - (\lambda(c\cos 2\theta_{N-r} - 1) + 2)\rho + \lambda^2 c \sin^2 \theta_{N-r} + \lambda(c\cos 2\theta_{N-r} - 1) + 1 = 0.$$

The discriminant of above equation is $\Delta = \lambda^2(c^2 \cos^2 2\theta_{N-r} - 2c + 1)$. The two solutions of $\Delta = 0$ are $[1 \pm \sin(2\theta_{N-r})]/\cos^2(2\theta_{N-r})$. Notice that $[1 + \sin(2\theta)]/\cos^2(2\theta) \geq 1$ for any $\theta \in [0, \frac{\pi}{2}]$ and $c \in (0,1)$. Let $c^* = [1 - \sin(2\theta_{N-r})]/\cos^2(2\theta_{N-r})$, then the magnitudes of ρ_2 and ρ_3 are:

$$\text{if } c \leq c^*, \quad \text{then } |\rho_2| = \frac{1}{2}|\lambda c\cos(2\theta_{N-r}) - \lambda + 2 + \lambda\sqrt{\cos^2(2\theta_{N-r})c^2 - 2c + 1}|,$$

$$|\rho_3| = \frac{1}{2}|\lambda c\cos(2\theta_{N-r}) - \lambda + 2 - \lambda\sqrt{\cos^2(2\theta_{N-r})c^2 - 2c + 1}|,$$

$$\text{if } c > c^*, \quad \text{then } |\rho_2| = |\rho_3| = \sqrt{c\lambda^2 \sin^2 \theta_{N-r} - (1 - c\cos(2\theta_{N-r}))\lambda + 1}.$$

Recall the generalized Douglas–Rachford splitting (1.6) and (1.7a) converges due to convexity [26]. When the iterations enter the asymptotic regime (after the cut-off location of the operator S does not change), the convergence rate is governed by the largest magnitude of eigenvalues ρ_0, ρ_1, ρ_2, and ρ_3:

THEOREM 2.5. *For a standard fixed point of generalized Douglas–Rachford splitting iteration as defined in Section 2.1, the asymptotic convergence rate of* (1.6) *solving* (1.3) *is linear. There exists a sufficiently large $K > 0$, such that for any integer $k \geq K$, we have*

$$\|\mathbf{y}^k - \mathbf{y}^*\|_2 \leq \widetilde{C}\left(\min_{c,\lambda} \max\{|\rho_0|, |\rho_1|, |\rho_2|, |\rho_3|\}\right)^k,$$

where K and $\widetilde{C}$ may depend on $\mathbf{A}$, b, and $\mathbf{y}^0$.

2.4. A simple strategy of choosing nearly optimal parameters. For solving problem (1.3), after the iteration of algorithm (1.6) enters the asymptotic linear convergence regime, the rate of convergence is governed by the largest magnitude of ρ_0, ρ_1, ρ_2, and ρ_3. For seeking optimal parameters, we can safely ignore ρ_0 because it is straightforward to verify that $\rho_0 \leq \rho_1$ with the optimal parameters derived below. It is highly preferred to construct a guideline for selecting parameters c and λ such that for $\max\{|\rho_1|, |\rho_2|, |\rho_3|\}$ is reasonably small.

We first consider the case $\theta_{N-r} \in (\frac{\pi}{4}, \frac{\pi}{2}]$. It is easy to check $c^* = \frac{1}{(\cos\theta_{N-r} + \sin\theta_{N-r})^2} \in (\frac{1}{2}, 1]$. Define surfaces $\Gamma_i = \{(c, \lambda, z) : 0 < c < c^*, 0 < \lambda \leq 2, z = |\rho_i|\}$, where

$i \in \{1, 2, 3\}$. For any point $(c, \lambda, z) \in \Gamma_2 \cap \Gamma_3$, due to the fact that $|a + b| = |a - b|$ implies $ab = 0$ for any $a, b \in \mathbb{R}$, we have $(\lambda c \cos(2\theta_{N-r}) - \lambda + 2)\sqrt{\Delta} = 0$. When $c < c^*$ the discriminant $\Delta > 0$, we get $\lambda c \cos(2\theta_{N-r}) - \lambda + 2 = 0$. Thus, if there exists a point belongs to $\Gamma_1 \cap \Gamma_2 \cap \Gamma_3$, then it satisfies

$$\begin{cases} |1 - \lambda(1 - c)| = \frac{\lambda}{2}\sqrt{\cos^2(2\theta_{N-r})c^2 - 2c + 1}, \\ \lambda c \cos(2\theta_{N-r}) - \lambda + 2 = 0. \end{cases}$$

On surfaces Γ_i, $i \in \{1, 2, 3\}$, the parameters $c \in (0, c^*)$ and $\lambda \in (0, 2]$ implies above equations only have one solution $c = \frac{1}{2}$ and $\lambda = \frac{4}{2 - \cos(2\theta_{N-r})}$. Thus, we have

$$(2.10) \qquad \Gamma_1 \cap \Gamma_2 \cap \Gamma_3 = \left\{ \left(\frac{1}{2}, \frac{4}{2 - \cos(2\theta_{N-r})}, -\frac{\cos(2\theta_{N-r})}{2 - \cos(2\theta_{N-r})} \right) \right\}.$$

Therefore, we know when $\theta_{N-r} \in (\frac{\pi}{4}, \frac{\pi}{2}]$, the minimum of $\max\{|\rho_1|, |\rho_2|, |\rho_3|\}$ for $c \in (0, c^*)$ and $\lambda \in (0, 2]$ is not greater than $-\frac{\cos(2\theta_{N-r})}{2 - \cos(2\theta_{N-r})}$. To deal with $c \in [c^*, 1)$, we need the following lemma.

LEMMA 2.6. *Assume ρ_1 and ρ_2 are functions of c and λ, for which the minimum can be attained. Then, the following inequality holds.*

$$\min_{c, \lambda} \max\{|\rho_1|, |\rho_2|\} \geq \max\{\min_{c, \lambda} |\rho_1|, \min_{c, \lambda} |\rho_2|\}.$$

Proof. Assume the minimum of $\max\{|\rho_1|, |\rho_2|\}$ is achieved at (c_0, λ_0). We have

   i. If $|\rho_1(c_0, \lambda_0)| \geq |\rho_2(c_0, \lambda_0)|$, then $\min_{c, \lambda} \max\{|\rho_1|, |\rho_2|\} = |\rho_1(c_0, \lambda_0)| \geq \min_{c, \lambda} |\rho_1|$.

   ii. If $|\rho_1(c_0, \lambda_0)| < |\rho_2(c_0, \lambda_0)|$, then $\min_{c, \lambda} \max\{|\rho_1|, |\rho_2|\} = |\rho_2(c_0, \lambda_0)| > |\rho_1(c_0, \lambda_0)|$. Proof by contradiction: assume $\min_{c, \lambda} \max\{|\rho_1|, |\rho_2|\} < \min_{c, \lambda} |\rho_1|$, then it implies $|\rho_1(c_0, \lambda_0)| < \min_{c, \lambda} |\rho_1|$, which is impossible.

Thus, $\min_{c, \lambda} \max\{|\rho_1|, |\rho_2|\} \geq \min_{c, \lambda} |\rho_1|$. Similarly, $\min_{c, \lambda} \max\{|\rho_1|, |\rho_2|\} \geq \min_{c, \lambda} |\rho_2|$. $\qquad \square$

When $c \in [c^*, 1)$, the magnitude of ρ_2 and ρ_3 are equal, namely we only need to find suitable parameters c and λ such that the $\max\{|\rho_1|, |\rho_2|\}$ is reasonably small. It is easy to verify that, when $c \in [c^*, 1)$ and $\lambda \in (0, 2]$, the function ρ_1 is monotonically increasing with respect to c and monotonically decreasing with respect to λ. Thus, $\rho_1(c^*, 2) = 2c^* - 1 > 0$ gives $|\rho_1| = \rho_1$. Associated with λ greater or less than $-\frac{\cos(2\theta_{N-r})}{\sin^2 \theta_{N-r}}$, we have two cases.

   1. When $\lambda \in (0, -\frac{\cos(2\theta_{N-r})}{\sin^2 \theta_{N-r}}]$, recall the monotonicity of ρ_1, we have

$$\min_{c \in [c^*, 1), \; \lambda \in (0, -\frac{\cos(2\theta_{N-r})}{\sin^2 \theta_{N-r}}]} |\rho_1| = \rho_1\left(c^*, -\frac{\cos(2\theta_{N-r})}{\sin^2 \theta_{N-r}} \right)$$

$$= 1 + \frac{\cos(2\theta_{N-r})}{\sin^2 \theta_{N-r}}\left(1 - \frac{1}{(\cos\theta_{N-r} + \sin\theta_{N-r})^2} \right) > \frac{1}{2} > -\frac{\cos(2\theta_{N-r})}{2 - \cos(2\theta_{N-r})}.$$

By Lemma 2.6, when the principal angle $\theta_{N-r} \in (\frac{\pi}{4}, \frac{\pi}{2}]$, we know

$$\min_{c \in [c^*, 1), \; \lambda \in (0, -\frac{\cos(2\theta_{N-r})}{\sin^2 \theta_{N-r}}]} \max\{|\rho_1|, |\rho_2|\} > -\frac{\cos(2\theta_{N-r})}{2 - \cos(2\theta_{N-r})}.$$

Therefore, the common point of the three surfaces Γ_1, Γ_2, and Γ_3 in (2.10) is still a good choice.

2. When $\lambda \in (-\frac{\cos(2\theta_{N-r})}{\sin^2\theta_{N-r}}, 2]$, define $\kappa = c\lambda^2 \sin^2\theta_{N-r} - (1 - c\cos(2\theta_{N-r}))\lambda + 1$. We have $\frac{\partial \kappa}{\partial c} = \lambda(\lambda \sin^2\theta_{N-r} + \cos(2\theta_{N-r})) > 0$, which implies κ is monotonically increasing with respect to c in the interval $[c^*, 1)$. Thus, for any $c \geq c^*$, the $|\rho_2(c, \lambda)| \geq |\rho_2(c^*, \lambda)|$ holds. Again, recall the monotonicity of ρ_1, we obtain

$$\min_{c \in [c^*, 1),\ \lambda \in (-\frac{\cos(2\theta_{N-r})}{\sin^2\theta_{N-r}}, 2]} \max\{|\rho_1|, |\rho_2|\} = \min_{\lambda \in (-\frac{\cos(2\theta_{N-r})}{\sin^2\theta_{N-r}}, 2]} \max\{|\rho_1(c^*, \lambda)|, |\rho_2(c^*, \lambda)|\}.$$

Since $|\rho_1(c^*, \lambda)| = 1 - \lambda(1 - c^*)$ and $|\rho_2(c^*, \lambda)| = |1 - \frac{\lambda}{1+\cot\theta_{N-r}}|$, when $\theta_{N-r} \in (\frac{\pi}{4}, \frac{\pi}{2}]$, $\frac{1}{1+\cot\theta_{N-r}} > 1 - c^*$, then the equation $|\rho_1(c^*, \lambda)| = |\rho_2(c^*, \lambda)|$ has one and only one root

$$\lambda^* = \frac{2}{1 + \frac{1}{1+\cot\theta_{N-r}} - \frac{1}{(\cos\theta_{N-r}+\sin\theta_{N-r})^2}}.$$

Therefore, we know when $\theta_{N-r} \in (\frac{\pi}{4}, \frac{\pi}{2}]$, the minimum of $\max\{|\rho_1|, |\rho_2|, |\rho_3|\}$ for $c \in [c^*, 1)$ and $\lambda \in (-\frac{\cos(2\theta_{N-r})}{\sin^2\theta_{N-r}}, 2]$ is not larger than $1 - \lambda^*(1 - c^*)$.

Next, let us consider the case $\theta_{N-r} \in (0, \frac{\pi}{4}]$. When $c \in (0, c^*)$ and $\lambda \in (0, 2]$, the discriminant $\Delta > 0$, namely the quadratic equation (2.9) has two real roots. Moreover, $|\rho_2| > |\rho_3|$ obviously. Thus, we only need to minimize the $\max\{|\rho_1|, |\rho_2|\}$. Define

$$\tilde{\kappa} = \lambda c \cos(2\theta_{N-r}) - \lambda + 2 + \lambda\sqrt{\cos^2(2\theta_{N-r})c^2 - 2c + 1}.$$

Since for any $\theta_{N-r} \in (0, \frac{\pi}{4}]$, $c \in (0, c^*)$, and $\lambda \in (0, 2]$ the $\lambda c\cos(2\theta_{N-r}) - \lambda + 2 > 0$, we have $|\rho_2| = \frac{1}{2}\tilde{\kappa}$. From

$$\frac{\partial \tilde{\kappa}}{\partial c} = \lambda\left(\cos(2\theta_{N-r}) + \frac{c\cos^2(2\theta_{N-r}) - 1}{\sqrt{\cos^2(2\theta_{N-r})c^2 - 2c + 1}}\right) \leq 0,$$

$$\frac{\partial \tilde{\kappa}}{\partial \lambda} = c\cos(2\theta_{N-r}) - 1 + \sqrt{\cos^2(2\theta_{N-r})c^2 - 2c + 1} \leq 0,$$

we know the $\tilde{\kappa}$ is monotonically decreasing with respect to both c and λ. Thus $\tilde{\kappa}$ take minimum at $c = c^*$ and $\lambda = 2$. By Lemma 2.6, when the principal angle $\theta_{N-r} \in (0, \frac{\pi}{4}]$, we know

(2.11) $$\min_{c \in (0,c^*),\ \lambda \in (0,2]} \max\{|\rho_1|, |\rho_2|\} \geq \min_{c \in (0,c^*),\ \lambda \in (0,2]} |\rho_2| = \frac{1}{2}\tilde{\kappa}(c^*, 2) = c^*\cos 2\theta_{N-r}.$$

Notice, when $c = c^*$ and $\lambda = 2$, the magnitude of ρ_1 and ρ_2 can be simplified as $|\rho_1| = |2c^* - 1|$ and $|\rho_2| = c^*\cos 2\theta_{N-r}$, where $c^* = \frac{1}{(\cos\theta_{N-r}+\sin\theta_{N-r})^2}$. It is easy to check that $|\rho_2| > |\rho_1|$ holds for any $\theta_{N-r} \in (0, \frac{\pi}{4}]$. We have

(2.12)
$$\min_{c \in (0,c^*),\ \lambda \in (0,2]} \max\{|\rho_1|, |\rho_2|\} \leq \max\{|\rho_1(c^*, 2)|, |\rho_2(c^*, 2)|\} = |\rho_2(c^*, 2)| = c^*\cos 2\theta_{N-r}.$$

From above (2.11) and (2.12), we obtain the minimum of $\max\{|\rho_1|, |\rho_2|, |\rho_3|\}$ equals $c^*\cos 2\theta_{N-r}$, which is achieved at $c = c^*$ and $\lambda = 2$. When $c \in [c^*, 2)$, following the similar argument as above, we can show $|\rho_1| = 1 - \lambda(1 - c)$, which is monotonically

increasing with respect to c and monotonically decreasing with respect to λ. In addition, we also have $|\rho_2| = |\rho_3|$ which is monotonically increasing with respect to c. Thus, we have

$$\min_{c\in[c^*,1),\ \lambda\in(0,2]} \max\{|\rho_1|,|\rho_2|,|\rho_3|\} = \min_{\lambda\in(0,2]} \max\{|\rho_1(c^*,\lambda)|,|\rho_2(c^*,\lambda)|\}$$

$$= \min_{\lambda\in(0,2]} \frac{1}{2}\lambda c^* \cos(2\theta_{N-r}) - \frac{1}{2}\lambda + 1.$$

The last equality above is due to the fact the $|\rho_1(c^*,\lambda)| \le |\rho_2(c^*,\lambda)|$ holds for any $\theta_{N-r} \in (0,\frac{\pi}{4}]$. From $\lambda c^* \cos(2\theta_{N-r}) - \lambda$ is monotonically decreasing with respect to λ, we know, in this case, the minimum equals $c^* \cos(2\theta_{N-r})$, which is taken at $c = c^*$ and $\lambda = 2$.

To this end, let us make a summary of the parameter selection principle as follows.
1. When $\theta_{N-r} \in (\frac{3}{8}\pi, \frac{1}{2}\pi]$, a suitable choice of parameters are: $c = \frac{1}{2}$, $\lambda = \frac{4}{2-\cos(2\theta_{N-r})}$. The associated asymptotic linear convergence rate is governed by $-\frac{\cos(2\theta_{N-r})}{2-\cos(2\theta_{N-r})}$.
2. When $\theta_{N-r} \in (\frac{1}{4}\pi, \frac{3}{8}\pi]$, a suitable choice of parameters are: $c = c^*$, $\lambda = \lambda^*$. The associated asymptotic linear convergence rate is governed by $1 - \lambda^*(1-c^*)$.
3. When $\theta_{N-r} \in (0, \frac{1}{4}\pi]$, a suitable choice of parameters are: $c = c^*$, $\lambda = 2$. The associated asymptotic linear convergence rate is governed by $c^* \cos(2\theta_{N-r})$.

Remark 2.7. The exact value of the principal angle θ_{N-r} in (2.8) is unknown. But it is simple to estimate θ_{N-r} by counting the number of bad cells, e.g., let $\hat{r}$ be the number of $u_i \notin [m, M]$ and use $\hat{r}$ instead of r in (2.8) to compute θ_{N-r}. This gives a simple guideline (1.9) for choosing nearly optimal parameters, which is efficient in all our numerical tests as shown in Section 4.

Remark 2.8. In a large scale 3D problem, usually the ratio of bad cells with cell averages out of bound in the DG scheme is quite small. In such a case, we expect $r \ll N$, with which θ_{N-r} is very close to zero. In this case, by the discussions above, the convergence rate in Theorem 2.5 becomes $-\frac{\cos(2\theta_{N-r})}{2-\cos(2\theta_{N-r})}$. If $\hat{r}$ is also a good approximation to r, which is usually true in this context, then we get the rate (1.10).

With the guideline (1.9) for choosing nearly optimal parameters in (1.7a), we can use the two-step limiter as explained in Section 1.5 to enforce bounds of DG solutions.

3. Application to phase-field equations. One of the popular approaches for modeling multi-phase fluid flow in micro-to-millimeter pore structures is to use phase-field equations [15]. Efficient and accurate pore-scale fluid dynamics simulators have important applications in digital rock physics (DRP), which has been extensively used in the petroleum industry for optimizing enhanced oil recovery schemes.

3.1. Mathematical model. In an open bounded domain $\Omega \subset \mathbb{R}^d$ over a time interval $(0, T]$, the dimensionless CHNS equations are given by:

$$(3.1a) \qquad \partial_t \phi - \frac{1}{\text{Pe}}\boldsymbol{\nabla}\cdot(\mathcal{M}(\phi)\boldsymbol{\nabla}\mu) + \boldsymbol{\nabla}\cdot(\phi\boldsymbol{v}) = 0 \quad \text{in } (0,T]\times\Omega,$$

$$(3.1b) \qquad \mu + \text{Cn}^2\Delta\phi - \Phi'(\phi) = 0 \quad \text{in } (0,T]\times\Omega,$$

$$(3.1c) \quad \partial_t\boldsymbol{v} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{v} - \frac{2}{\text{Re}}\boldsymbol{\nabla}\cdot\boldsymbol{\varepsilon}(\boldsymbol{v}) + \frac{1}{\text{ReCa}}\boldsymbol{\nabla}p - \frac{3}{2\sqrt{2}\,\text{ReCaCn}}\mu\boldsymbol{\nabla}\phi = 0 \quad \text{in } (0,T]\times\Omega,$$

$$(3.1d) \qquad \boldsymbol{\nabla}\cdot\boldsymbol{v} = 0 \quad \text{in } (0,T]\times\Omega,$$

where ϕ, μ, $\boldsymbol{v}$, and p are order parameter, chemical potential, velocity, and pressure. The non-dimensional quantities Pe, Cn, Re, and Ca denote the Péclet number, Cahn number, Reynolds number, and capillary number, respectively. The strain tensor is given by $\boldsymbol{\varepsilon}(\boldsymbol{v}) = \frac{1}{2}(\boldsymbol{\nabla}\boldsymbol{v} + (\boldsymbol{\nabla}\boldsymbol{v})^{\mathrm{T}})$. The function $\mathcal{M}$ denotes mobility. Typical choices of $\mathcal{M}$ include the constant mobility $\mathcal{M}(\phi) = \mathcal{M}_0 > 0$, where $\mathcal{M}_0$ can be set to 1 after nondimensionalization, and the degenerate mobility $\mathcal{M}(\phi) = 1 - \phi^2$. The function Φ is a scalar potential, which is also called chemical energy density. Classical and widely used forms are the polynomial Ginzburg–Landau (GL) double well potential: $\Phi(\phi) = \frac{1}{4}(1 - \phi)^2(1 + \phi)^2$ and the Flory–Huggins (FH) logarithmic potential with parameters α and β: $\Phi(\phi) = \frac{\alpha}{2}\left((1 + \phi)\ln\left(\frac{1+\phi}{2}\right) + (1 - \phi)\ln\left(\frac{1-\phi}{2}\right)\right) + \frac{\beta}{2}(1 - \phi^2)$.

We supplement (3.1) with initials $\phi = \phi^0$ and $\boldsymbol{v} = \boldsymbol{v}^0$ on $\{0\} \times \Omega$. Let $\boldsymbol{n}$ denote the unit outward normal to domain Ω. We decompose the boundary $\partial\Omega$ into three disjoint subsets $\partial\Omega = \partial\Omega^{\mathrm{wall}} \cup \partial\Omega^{\mathrm{in}} \cup \partial\Omega^{\mathrm{out}}$, where $\partial\Omega^{\mathrm{wall}}$ denotes fluid–solid interface and $\partial\Omega^{\mathrm{in}}$ and $\partial\Omega^{\mathrm{out}}$ are inflow boundary and outflow boundary

$$\partial\Omega^{\mathrm{in}} = \{\boldsymbol{x} \in \partial\Omega : \ \boldsymbol{v} \cdot \boldsymbol{n} < 0\} \ \text{ and } \ \partial\Omega^{\mathrm{out}} = \partial\Omega \setminus (\partial\Omega^{\mathrm{wall}} \cup \partial\Omega^{\mathrm{in}}).$$

We prescribe Dirichlet boundary conditions $\phi = \phi_{\mathrm{D}}$ and $\boldsymbol{v} = \boldsymbol{v}_{\mathrm{D}}$ on $(0, T] \times \partial\Omega^{\mathrm{in}}$. For velocity, the no-slip boundary condition $\boldsymbol{v} = \boldsymbol{0}$ is used on $(0, T] \times \partial\Omega^{\mathrm{wall}}$ and "do nothing" boundary condition $(2\boldsymbol{\varepsilon}(\boldsymbol{v}) - \frac{1}{\mathrm{Ca}}p\mathbf{I})\boldsymbol{n} = \boldsymbol{0}$ is applied on $(0, T] \times \partial\Omega^{\mathrm{out}}$. Wettability is modeled by a contact angle ϑ that is enforced by: $\boldsymbol{\nabla}\phi \cdot \boldsymbol{n} = -\frac{2\sqrt{2}\cos(\vartheta)}{3\mathrm{Cn}}g'(\phi)$ on $(0, T] \times (\partial\Omega^{\mathrm{wall}} \cup \partial\Omega^{\mathrm{out}})$, where the function g is a blending function. The closed-form expression of g depends on the choice of chemical energy density [4]. For the Ginzburg–Landau potential, we have $g(\phi) = \frac{1}{4}(\phi^3 - 3\phi + 2)$. In addition, we employ the homogeneous Neumann boundary condition $\mathcal{M}(\phi)\boldsymbol{\nabla}\mu \cdot \boldsymbol{n} = 0$ on $(0, T] \times \partial\Omega$ to ensure the global mass conservation.

The order parameter ϕ is the difference between the mass fraction ϕ_{A} and ϕ_{B} of the phase A and phase B. With constraint $\phi_{\mathrm{A}} + \phi_{\mathrm{B}} = 1$ for a two-component mixture as well as mass fractions belonging to $[0, 1]$, a physically meaningful range of the order parameter field is $[-1, 1]$. The Cahn–Hilliard equation with the degenerate mobility or with the logarithmic potential enjoys bound-preserving property [41]. However, for constant mobility with GL polynomial potential, the analytical solution of Cahn–Hilliard equation is not bound-preserving [2]. For a given initial data $\phi^0 \in [-1, 1]$, it is an open question whether the solution of a fully coupled CHNS system should have a bounded order parameter in $[-1, 1]$. On the other hand, empirically we would expect a reasonable solution, e.g., the discrete order parameter field, should be bounded by -1 and 1 for any time $t > 0$.

3.2. Time discretization. The CHNS equations form a highly nonlinear coupled system. One of the popular approaches of constructing efficient numerical algorithms for large-scale simulations in complex computational domains is to use splitting methods, e.g., to decouple the mass and momentum equations and to further split the convection from the incompressibility constraint [37]. Also, see [21, 19] for an overview of the splitting methods for time-dependent incompressible flows.

We uniformly partition the interval $[0, T]$ into N_{st} subintervals. Let τ denote the time step size. For the chemical energy density, we adopt a convex–concave decomposition of the form $\Phi = \Phi_+ + \Phi_-$, where the convex part Φ_+ is treated time implicitly and the concave part Φ_- is treated time explicitly. For the nonlinear convection $\boldsymbol{v} \cdot \boldsymbol{\nabla}\boldsymbol{v}$, the form $\mathcal{C}(\cdot, \cdot)$ is a semi-discretization that satisfies a positivity property, see the equation (12) in [27]. For any $1 \le n \le N_{\mathrm{st}}$, our first-order time discretization

algorithm consists of the following steps:

Step 1. Given $(\phi^{n-1}, \boldsymbol{w}^{n-1})$, compute (ϕ^n, μ^n) such that

$$\phi^n - \frac{\tau}{\text{Pe}} \boldsymbol{\nabla} \cdot (\mathcal{M}(\phi^{n-1}) \boldsymbol{\nabla} \mu^n) + \tau \boldsymbol{\nabla} \cdot (\phi^n \boldsymbol{w}^{n-1}) = \phi^{n-1} \qquad \text{in } \Omega,$$

$$-\mu^n - \text{Cn}^2 \Delta \phi^n + \Phi_+'(\phi^n) = -\Phi_-'(\phi^{n-1}) \qquad \text{in } \Omega.$$

Step 2. Given $(\phi^n, \mu^n, \boldsymbol{v}^{n-1}, p^{n-1}, \psi^{n-1})$, compute $\boldsymbol{v}^n$ such that

$$\boldsymbol{v}^n + \tau C(\boldsymbol{v}^{n-1}, \boldsymbol{v}^n) - \frac{2\tau}{\text{Re}} \boldsymbol{\nabla} \cdot \boldsymbol{\varepsilon}(\boldsymbol{v}^n) = \boldsymbol{v}^{n-1}$$

$$-\frac{\tau}{\text{ReCa}} \boldsymbol{\nabla}(p^{n-1} + \psi^{n-1}) + \frac{3\tau}{2\sqrt{2}\,\text{ReCaCn}} \mu^n \boldsymbol{\nabla} \phi^n \qquad \text{in } \Omega.$$

Step 3. Given $\boldsymbol{v}^n$, compute ψ^n such that

$$-\Delta \psi^n = -\frac{\text{ReCa}}{\tau} \boldsymbol{\nabla} \cdot \boldsymbol{v}^n \qquad \text{in } \Omega.$$

Step 4. Given $(\boldsymbol{v}^n, p^{n-1}, \psi^n)$, compute $(\boldsymbol{w}^n, p^n)$ such that

$$\boldsymbol{w}^n = \boldsymbol{v}^n - \frac{\tau}{\text{ReCa}} \boldsymbol{\nabla} \psi^n,$$

$$p^n = p^{n-1} + \psi^n - \sigma_\chi \text{Ca} \boldsymbol{\nabla} \cdot \boldsymbol{v}^n.$$

The parameter σ_χ is equal to $\frac{2}{d}$, namely, we use $\sigma_\chi = \frac{2}{3}$ for our numerical simulations in three dimensions. To start time marching, we set $p^0 = 0$ and $\psi^0 = 0$. The functions ϕ^0 and $\boldsymbol{w}^0 = \boldsymbol{v}^0$ are given initial data.

Remark 3.1. The above scheme is a combination of the convex splitting approach for the Cahn–Hilliard equation with the classical rotational pressure-correction algorithm (see Section 3.4 in [21]) for the Navier–Stokes equations. More precisely, Step 2 to Step 4 can be rewritten as follows:

$$\frac{1}{\tau}(\boldsymbol{v}^n - \boldsymbol{w}^{n-1}) + C(\boldsymbol{v}^{n-1}, \boldsymbol{v}^n) - \frac{2}{\text{Re}} \boldsymbol{\nabla} \cdot \boldsymbol{\varepsilon}(\boldsymbol{v}^n) = -\frac{1}{\text{ReCa}} \boldsymbol{\nabla} p^{n-1} + \frac{3}{2\sqrt{2}\,\text{ReCaCn}} \mu^n \boldsymbol{\nabla} \phi^n,$$

$$\begin{cases} \dfrac{1}{\tau}(\boldsymbol{w}^n - \boldsymbol{v}^n) + \dfrac{1}{\text{ReCa}} \boldsymbol{\nabla} \psi^n = 0, \\ \boldsymbol{\nabla} \cdot \boldsymbol{w}^n = 0, \end{cases} \qquad\qquad \psi^n = p^n - p^{n-1} + \sigma_\chi \text{Ca} \boldsymbol{\nabla} \cdot \boldsymbol{v}^n.$$

We use $\boldsymbol{w}^{n-1}$, instead of $\boldsymbol{v}^{n-1}$, in the advection term in Step 1, since $\boldsymbol{\nabla} \cdot \boldsymbol{w}^{n-1} = 0$.

For the sake of simplicity, we only presented a first-order version of the scheme, although high-order version can be constructed accordingly. On the other hand, it is also possible to construct energy dissipating schemes as in [38]. Since our focus in this paper is in preserving bounds for a DG spacial discretization, we employ a simple time-marching strategy.

3.3. Space discretization. Decoupled splitting algorithms combined with interior penalty DG spatial formations have been constructed to solve various CHNS models in large-scale complex-domain DRP simulations [15, 28, 30]. Also, see [29, 32, 33] for solvability, stability, and optimal error estimates on using DG with decoupled splitting schemes for CHNS equations and viscous incompressible flow. Here, we briefly review the fully discrete scheme.

Let $\mathcal{T}_h = \{E_i\}$ be a family of conforming nondegenerate (regular) meshes of the domain Ω with maximum element diameter h. Let Γ_h be the set of interior faces. For each interior face $e \in \Gamma_h$ shared by elements E_{i^-} and E_{i^+}, with $i^- < i^+$, we define a unit normal vector $\boldsymbol{n}_e$ that points from E_{i^-} into E_{i^+}. For a boundary face, $e \subset \partial\Omega$, the normal vector $\boldsymbol{n}_e$ is taken to be the unit outward vector to $\partial\Omega$. Let $\mathbb{P}_k(E_i)$ denote the set of all polynomials of degree at most k on an element E_i. Define the broken polynomial spaces X_h and $\mathbf{X}_h$, for any $k \geq 1$,

$$X_h = \{\chi_h \in L^2(\Omega) : \ \chi_h|_{E_i} \in \mathbb{P}_k(E_i), \ \forall E_i \in \mathcal{T}_h\},$$
$$\mathbf{X}_h = \{\boldsymbol{\theta}_h \in L^2(\Omega)^d : \ \boldsymbol{\theta}_h|_{E_i} \in \mathbb{P}_k(E_i)^d, \ \forall E_i \in \mathcal{T}_h\}.$$

The average and jump for any scalar quantity χ on a boundary face coincide with its trace; and on interior faces they are defined by

$$\{\!\{\chi\}\!\}|_e = \frac{1}{2}\,\chi|_{E_{i^-}} + \frac{1}{2}\,\chi|_{E_{i^+}}, \quad [\![\chi]\!]\,|_e = \chi|_{E_{i^-}} - \chi|_{E_{i^+}}, \quad \forall e = \partial E_{i^-} \cap \partial E_{i^+}.$$

The related definitions for any vector quantity are similar. For more details see [36].

Let $(\cdot,\cdot)_{\mathcal{O}}$ denote the L^2 inner product over $\mathcal{O}$. For instance, on any face e the L^2 inner product is denoted by $(\cdot,\cdot)_e$. We make use of the following compact notation for the L^2 inner product on the interior and boundary faces

$$(\cdot,\cdot)_{\mathcal{O}} = \sum_{e \in \mathcal{O}} (\cdot,\cdot)_e, \quad \text{where} \ \ \mathcal{O} = \Gamma_h, \ \partial\Omega, \ \partial\Omega^{\mathrm{in}}, \ \partial\Omega^{\mathrm{out}}, \ \cdots.$$

For convenience, we omit the subscript when $\mathcal{O} = \Omega$, namely denote $(\cdot,\cdot) = (\cdot,\cdot)_{\Omega}$. We still use $\boldsymbol{\nabla}$ and $\boldsymbol{\nabla}\cdot$ to denote the broken gradient and broken divergence.

For completeness, let us recall the DG forms below and we skip their derivation. Associated with the advection term $\boldsymbol{\nabla}\cdot(\phi\boldsymbol{w})$ and the convection term $\boldsymbol{v}\cdot\boldsymbol{\nabla}z$, we define

$$a_{\mathrm{adv}}(\phi, \boldsymbol{w}, \chi) = -(\phi, \boldsymbol{w}\cdot\boldsymbol{\nabla}\chi) + (\phi^{\uparrow}\{\!\{\boldsymbol{w}\cdot\boldsymbol{n}_e\}\!\}, [\![\chi]\!])_{\Gamma_h},$$

$$a_{\mathrm{conv}}(\boldsymbol{v}; z, \boldsymbol{\theta}) = (\boldsymbol{v}\cdot\boldsymbol{\nabla}z, \boldsymbol{\theta}) + \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{v}, z\cdot\boldsymbol{\theta})$$
$$- \frac{1}{2}([\![\boldsymbol{v}\cdot\boldsymbol{n}_e]\!], \{\!\{z\cdot\boldsymbol{\theta}\}\!\})_{\Gamma_h\cup\partial\Omega^{\mathrm{in}}} + \sum_{E\in\mathcal{T}_h}(|\{\!\{\boldsymbol{v}\}\!\}\cdot\boldsymbol{n}_E|, (z^{\mathrm{int}} - z^{\mathrm{ext}})\cdot\boldsymbol{\theta}^{\mathrm{int}})_{\partial E_-^v}.$$

The superscript int (resp. ext) refers to the trace of a function on a face of E coming from the interior (resp. exterior). The set ∂E_-^v is the upwind part of ∂E, defined by $\partial E_-^v = \{\boldsymbol{x} \in \partial E : \{\!\{\boldsymbol{v}\}\!\}\cdot\boldsymbol{n}_E < 0\}$, where $\boldsymbol{n}_E$ is the unit outward normal vector to E [18]. The upwind quantity $\phi^{\uparrow}$ on an interior face e is evaluated by

$$\phi^{\uparrow}\Big|_{e\in\Gamma_h} = \begin{cases} \phi\big|_{E_{i^-}} & \text{if } \{\!\{\boldsymbol{w}\}\!\}\cdot\boldsymbol{n}_e \geq 0, \\ \phi\big|_{E_{i^+}} & \text{if } \{\!\{\boldsymbol{w}\}\!\}\cdot\boldsymbol{n}_e < 0. \end{cases}$$

Associated with the operator $-\boldsymbol{\nabla}\cdot(z\boldsymbol{\nabla}\xi)$, we define

$$a_{\mathrm{diff}}(z; \xi, \chi) = (z\boldsymbol{\nabla}\xi, \boldsymbol{\nabla}\chi) - (\{\!\{z\boldsymbol{\nabla}\xi\cdot\boldsymbol{n}_e\}\!\}, [\![\chi]\!])_{\Gamma_h}$$
$$- (\{\!\{z\boldsymbol{\nabla}\chi\cdot\boldsymbol{n}_e\}\!\}, [\![\xi]\!])_{\Gamma_h} + \frac{\sigma}{h}([\![\xi]\!], [\![\chi]\!])_{\Gamma_h}.$$

Associated with the Laplace operator $-\Delta\xi$ (for terms $-\Delta\phi$ and $-\Delta\psi$), we define

$$-\Delta\xi + \text{Dirichlet on } \partial\Omega^{\text{in}} \;\rightsquigarrow\; a_{\text{diff,in}}(\xi,\chi) = a_{\text{diff}}(1;\xi,\chi) - (\nabla\xi\cdot\boldsymbol{n}_e,\chi)_{\partial\Omega^{\text{in}}}$$
$$- (\nabla\chi\cdot\boldsymbol{n}_e,\xi)_{\partial\Omega^{\text{in}}} + \frac{\sigma}{h}(\xi,\chi)_{\partial\Omega^{\text{in}}},$$

$$-\Delta\xi + \text{Dirichlet on } \partial\Omega^{\text{out}} \;\rightsquigarrow\; a_{\text{diff,out}}(\xi,\chi) = a_{\text{diff}}(1;\xi,\chi) - (\nabla\xi\cdot\boldsymbol{n}_e,\chi)_{\partial\Omega^{\text{out}}}$$
$$- (\nabla\chi\cdot\boldsymbol{n}_e,\xi)_{\partial\Omega^{\text{out}}} + \frac{\sigma}{h}(\xi,\chi)_{\partial\Omega^{\text{out}}}.$$

Associated with the diffusion term $-2\nabla\cdot\boldsymbol{\varepsilon}(\boldsymbol{v})$, we define

$$a_{\text{ellip}}(\boldsymbol{v},\boldsymbol{\theta}) = 2(\boldsymbol{\varepsilon}(\boldsymbol{v}),\boldsymbol{\varepsilon}(\boldsymbol{\theta})) - 2(\{\!|\boldsymbol{\varepsilon}(\boldsymbol{v})\boldsymbol{n}_e|\!\},[\![\boldsymbol{\theta}]\!])_{\Gamma_h} - 2(\{\!|\boldsymbol{\varepsilon}(\boldsymbol{\theta})\boldsymbol{n}_e|\!\},[\![\boldsymbol{v}]\!])_{\Gamma_h}$$
$$+ \frac{\sigma}{h}([\![\boldsymbol{v}]\!],[\![\boldsymbol{\theta}]\!])_{\Gamma_h} - 2(\boldsymbol{\varepsilon}(\boldsymbol{v})\boldsymbol{n}_e,\boldsymbol{\theta})_{\partial\Omega^{\text{in}}} - 2(\boldsymbol{\varepsilon}(\boldsymbol{\theta})\boldsymbol{n}_e,\boldsymbol{v})_{\partial\Omega^{\text{in}}} + \frac{\sigma}{h}(\boldsymbol{v},\boldsymbol{\theta})_{\partial\Omega^{\text{in}}}.$$

The remaining forms in the right-hand sides of the discrete equations account for the boundary conditions (see b_{diff} and b_{vel}) and the pressure and potential (see b_{pres}):

$$b_{\text{diff}}(\xi,\chi) = -(\phi_{\text{D}},\nabla\chi\cdot\boldsymbol{n}_e)_{\partial\Omega^{\text{in}}} + \frac{\sigma}{h}(\phi_{\text{D}},\chi)_{\partial\Omega^{\text{in}}} - \frac{2\sqrt{2}\delta\cos(\vartheta)}{3\,\text{Cn}}(g'(\xi),\chi)_{\partial\Omega^{\text{wall}}\cup\partial\Omega^{\text{out}}},$$

$$b_{\text{pres}}(p,\psi,\boldsymbol{\theta}) = -(p,\nabla\cdot\boldsymbol{\theta}) + (\{\!|p|\!\},[\![\boldsymbol{\theta}\cdot\boldsymbol{n}_e]\!])_{\Gamma_h\cup\partial\Omega} + (\nabla\psi,\boldsymbol{\theta}),$$

$$b_{\text{vel}}(\boldsymbol{\theta}) = -\frac{3}{2}(\boldsymbol{v}_{\text{D}}\cdot\boldsymbol{n},\boldsymbol{v}_{\text{D}}\cdot\boldsymbol{\theta})_{\partial\Omega^{\text{in}}} - \frac{2}{\text{Re}}(\boldsymbol{\varepsilon}(\boldsymbol{\theta})\boldsymbol{n}_e,\boldsymbol{v}_{\text{D}})_{\partial\Omega^{\text{in}}} + \frac{\sigma}{h\text{Re}}(\boldsymbol{v}_{\text{D}},\boldsymbol{\theta})_{\partial\Omega^{\text{in}}}.$$

In b_{diff}, the parameter δ is a scalar field that equals the constant one for smooth solid boundaries only and that otherwise corrects the numerical impact of the jaggedness of the solid boundaries obtained from micro-CT scanning. The derivation of this boundary condition and the wettability model can be found in [16].

For any $1 \le n \le N_{\text{st}}$, our fully discrete scheme for solving the CHNS equations (3.1) is as follows.

Algorithm CHNS. At time t^n, given scalar functions $\phi_h^{n-1}, p_h^{n-1}, \psi_h^{n-1}$ in X_h and vector functions $\boldsymbol{v}_h^{n-1}, \boldsymbol{w}_h^{n-1}$ in $\mathbf{X}_h$.

Step 1. Compute $\phi_h^n, \mu_h^n \in X_h$, such that for all $\chi_h \in X_h$,

$$(\phi_h^n,\chi_h) + \frac{\tau}{\text{Pe}}a_{\text{diff}}(\mathcal{M}(\phi_h^{n-1});\mu_h^n,\chi_h) + \tau a_{\text{adv}}(\phi_h^n,\boldsymbol{w}_h^{n-1},\chi_h)$$
$$= (\phi_h^{n-1},\chi_h) + \tau(\phi_{\text{D}}\boldsymbol{w}_h^{n-1}\cdot\boldsymbol{n}_e,\chi_h)_{\partial\Omega^{\text{in}}},$$
$$-(\mu_h^n,\chi_h) + \text{Cn}^2 a_{\text{diff,in}}(\phi_h^n,\chi_h) + (\Phi_+{}'(\phi_h^n),\chi_h)$$
$$= \text{Cn}^2 b_{\text{diff}}(\phi_h^{n-1},\chi_h) - (\Phi_-{}'(\phi_h^{n-1}),\chi_h).$$

Step 2. Compute $\boldsymbol{v}_h^n \in \mathbf{X}_h$, such that for all $\boldsymbol{\theta}_h \in \mathbf{X}_h$,

$$(\boldsymbol{v}_h^n,\boldsymbol{\theta}_h) + \tau a_{\text{conv}}(\boldsymbol{v}_h^{n-1},\boldsymbol{v}_h^n,\boldsymbol{\theta}_h) + \frac{\tau}{\text{Re}}a_{\text{ellip}}(\boldsymbol{v}_h^n,\boldsymbol{\theta}_h) = (\boldsymbol{v}_h^{n-1},\boldsymbol{\theta}_h)$$
$$- \frac{\tau}{\text{ReCa}}b_{\text{pres}}(p_h^{n-1},\psi_h^{n-1},\boldsymbol{\theta}_h) + \frac{3\tau}{2\sqrt{2}\,\text{ReCaCn}}(\mu_h^n\nabla\phi_h^n,\boldsymbol{\theta}_h) + \tau b_{\text{vel}}(\boldsymbol{\theta}_h).$$

Step 3. Compute $\psi_h^n \in X_h$, such that for all $\chi_h \in X_h$,

$$a_{\text{diff,out}}(\psi_h^n,\chi_h) = -\frac{\text{ReCa}}{\tau}(\nabla\cdot\boldsymbol{v}_h^n,\chi_h).$$

Step 4. Compute $w_h^n \in \mathbf{X}_h$ and $p_h^n \in X_h$, such that for all $\boldsymbol{\theta} \in \mathbf{X}_h$ and $\chi_h \in X_h$,

$$(w_h^n, \boldsymbol{\theta}_h) + \sigma_{\mathrm{div}}(\nabla \cdot w_h^n, \nabla \cdot \boldsymbol{\theta}_h) = (v_h^n, \boldsymbol{\theta}_h) - \frac{\tau}{\mathrm{ReCa}}(\nabla \psi_h^n, \boldsymbol{\theta}_h),$$

$$(p_h^n, \chi_h) = (p_h^{n-1}, \chi_h) + (\psi_h^n, \chi_h) - \sigma_\chi \mathrm{Ca}(\nabla \cdot v_h^n, \chi_h).$$

For the initial conditions, we set $p_h^0 = \psi_h^0 = 0$, $w_h^0 = v_h^0$; we compute ϕ_h^0 from the L^2 projection of ϕ^0 followed with Zhang–Shu limiter and we obtain v_h^0 from the L^2 projection of v^0.

To obtain a bound-preserving discrete order parameter field, at each time step after finishing computing Step 1 in the Algorithm CHNS, we apply the two-stage limiting strategy, see Section 1.5, to postprocess discrete order parameter ϕ_h^n. For the simulations in Section 4, we choose $m = -1$ and $M = 1$.

4. Numerical experiments. In this section, we first verify the high order accuracy of our cell average limiter (1.7) for a manufactured smooth solution. Then we verify the efficiency of the limiter (1.7) when using the parameters (1.9) on some representative physical simulations including spinodal decomposition, flows in micro structure, and merging droplets.

We use $\mathbb{P}_2$ scheme, e.g., discontinuous piecewise quadratic polynomials for space approximation, on cubic partitions of 3D domains. More details can be found in [14].

The penalty parameters for all tests are as follows. We use $\sigma = 8$ on Γ_h for a_{diff}; $\sigma = 16$ on $\partial\Omega$ for $a_{\mathrm{diff,in}}$ and $a_{\mathrm{diff,out}}$; $\sigma = 32$ on Γ_h and $\sigma = 64$ on $\partial\Omega^{\mathrm{in}}$ for a_{ellip}. In addition, we set tolerance $\epsilon = 10^{-13}$ to terminate Douglas–Rachford iterations.

4.1. Accuracy test. We use the manufactured solution method on domain $\Omega = (0, 1)^3$ with end time $T = 0.1$ to test the spatial order of convergence for our cell average limiter (1.7).

To trigger the cell average limiter (1.7), e.g., produce a fully discrete solution with cell average out of $[-1, 1]$ at each time step, we use constant mobility with GL polynomial potential and choose the prescribed order parameter field as an expression of a cosine function to power eight, as follows: $\phi = 1 - 2\cos^8\left(t + \frac{2\pi}{3}(x + y + z)\right)$. The chemical potential μ is an intermediate variable, which value is derived by the order parameter ϕ. The prescribed velocity and pressure fields are taken from the Beltrami flow [32], which enjoys the property that the nonlinear convection is balanced by the pressure gradient and the velocity is parallel to vorticity. We have

$$v = \begin{bmatrix} -e^{-t+x}\sin(y+z) - e^{-t+z}\cos(x+y) \\ -e^{-t+y}\sin(x+z) - e^{-t+x}\cos(y+z) \\ -e^{-t+z}\sin(x+y) - e^{-t+y}\cos(x+z) \end{bmatrix} \text{ and } p = -e^{-2t}(e^{x+z}\sin(y+z)\cos(x+y) +$$

$e^{x+y}\sin(x+z)\cos(y+z) + e^{y+z}\sin(x+y)\cos(x+z) + \frac{1}{2}e^{2x} + \frac{1}{2}e^{2y} + \frac{1}{2}e^{2z} - \overline{p^0})$, where $\overline{p^0} = 7.63958172715414$ guarantees zero average pressure over Ω for any $t > 0$ up to round-off error. The initial conditions and Dirichlet boundary condition for velocity are imposed by above manufactured solutions. For order parameter and chemical potential, we apply Neumann boundary condition. In addition, the right-hand side terms is evaluated by the prescribed solution.

Let us estimate the spatial rates of convergence by computing solutions on a sequence of uniformly refined meshes with fixed time step size $\tau = 10^{-4}$. In our experiments, the time step size is small enough such that the spatial error dominates. We choose Re = 1, Ca = 1, Pe = 1, Cn = 1, and the contact angle $\vartheta = 90°$ on $\partial\Omega$. If err_h denotes the error on a mesh with resolution h, then the rate is given by $\ln(\mathrm{err}_h/\mathrm{err}_{h/2})/\ln 2$.

We compare the L_h^2 rate and the L_h^∞ rate of order parameter in three scenarios: not applying any limiter, only applying the cell average limiter (1.7), and applying both limiters (1.7) and (1.8). In those applied cell average limiter (1.7) cases, the limiter is triggered at each time step, see Figure 1 for the ratio of the number of trouble cells to the number of total elements. The convergence of our original DG scheme without applying any limiter is optimal, see the top rows in Table 1. The middle and bottom rows in Table 1 show optimal convergence of the cases that only apply cell average limiter (1.7) and apply both cell average limiter (1.7) and Zhang–Shu limiter (1.8). Our limiting strategy preserves high order accuracy. We emphasize that DG methods with only the Zhang-Shu limiter will produce cell averages outside of the range $[-1, 1]$ for this particular test.

FIG. 1. *The performance of limiting strategy in the accuracy test of applying both limiters* (1.7) *and* (1.8) *with mesh resolution* $h = 1/2^5$. *Left: the percentage of trouble cells at each time step for the cell average limiter* (1.7). *Right: the number of Douglas–Rachford iterations at each time step. For each time step, at most 15 iterations are needed for* (1.7a)

.

	h	$\|\phi_h^{N_{\mathrm{st}}} - \phi(T)\|_{L_h^2}$	rate	$\|\phi_h^{N_{\mathrm{st}}} - \phi(T)\|_{L_h^\infty}$	rate
no limiter	$1/2^2$	2.034 E−1	—	5.636 E−1	—
	$1/2^3$	4.903 E−2	2.053	1.400 E−1	2.009
	$1/2^4$	5.714 E−3	3.101	2.731 E−2	2.358
	$1/2^5$	4.833 E−4	3.564	4.699 E−3	2.548
DR	$1/2^2$	2.053 E−1	—	5.826 E−1	—
	$1/2^3$	4.954 E−2	2.051	1.485 E−1	1.972
	$1/2^4$	5.720 E−3	3.115	2.799 E−2	2.408
	$1/2^5$	4.834 E−4	3.565	4.734 E−3	2.564
DR+ZS	$1/2^2$	2.872 E−1	—	7.631 E−1	—
	$1/2^3$	5.970 E−2	2.266	2.561 E−1	1.575
	$1/2^4$	7.181 E−3	3.057	3.926 E−2	2.706
	$1/2^5$	4.833 E−4	3.893	4.734 E−3	3.052

TABLE 1
Errors and spatial convergence rates of order parameter. Top: the original DG scheme without applying any limiters. Middle: only apply the cell average limiter (1.7) *(DR). Bottom: apply both of the cell average limiter* (1.7) *and Zhang–Shu limiter* (1.8).

4.2. Spinodal decomposition.

Spinodal decomposition is a phase separation mechanism, by which an initially thermodynamically unstable homogeneous mixture spontaneously decomposes into two separated phases that are more thermodynamically favorable. The spinodal decomposition test is a widely used benchmark for

validating CHNS simulators. In this part, we employ the degenerate mobility with GL polynomial potential.

We define a trefoil-shaped pipe, which is a set of points whose distance away from the following parametric curve is less than 0.09. A trefoil knot: $x(t) = \frac{1}{8}(\cos t + 2\cos 2t) + \frac{1}{2}$, $y(t) = \frac{1}{8}(\sin t - 2\sin 2t) + \frac{1}{2}$, and $z(z) = \frac{1}{4}\sin 3t + \frac{1}{2}$, where $t \in [0, 2\pi]$. Let us uniformly partition the unit cube $(0, 1)^3$ into cubic cells with the mesh resolution $h = 1/100$. A cell is marked as fluid if its center is in the above pipe, otherwise is marked as solid. The computational domain Ω is defined as the union of all fluid cells. We consider a closed system, i.e., $\partial\Omega = \partial\Omega^{\mathrm{wall}}$. The initial order parameter field is generated by sampling numbers from a discrete uniform distribution, $c^0|_{E_i} \sim \mathrm{U}\{-1, 1\}$, and the initial velocity field is taken to be zero. We take the time step size $\tau = 1\times10^{-3}$. For physical parameters, we choose Re = 1, Ca = 0.1, Pe = 1, Cn = h, and the contact angle $\vartheta = 90°$ on $\partial\Omega$.

Figure 2 shows snapshots of the order parameter field. We employ a rainbow color scale that maps the values in $[-1, 1]$ from transparent blue to non-transparent red for plotting the order parameter field. The center of the diffusive interface is colored green. We observe that the homogeneous mixture decomposes into two separate phases. With a neutral wall, i.e., the contact angle $\vartheta = 90°$, in the final stage of the simulation, each of the two phases occupies several disjoint sections of the domain. The interfaces are perpendicular to the solid surface. Our limiters remove overshoots and undershoots. The global mass is conserved, see the left subfigure of Figure 3.

The middle subfigure of Figure 3 records the number of iterations of the Douglas–Rachford algorithm on each time step. To measure the convergence rate, we run the Douglas–Rachford algorithm for 10^3 iterations with a very small tolerance to approximate $\boldsymbol{y}^*$ and $\boldsymbol{x}^*$ numerically. Then we plot $\|\boldsymbol{y}^k - \boldsymbol{y}^*\|_2$ and $\|\boldsymbol{x}^k - \boldsymbol{x}^*\|_2$. The right subfigure of Figure 3 shows asymptotic linear convergence rates at the selected time step 128. We see the convergence rates match our analysis in Theorem 2.5. In addition, we check the convergence rates on all of the rest steps that match with our analysis.
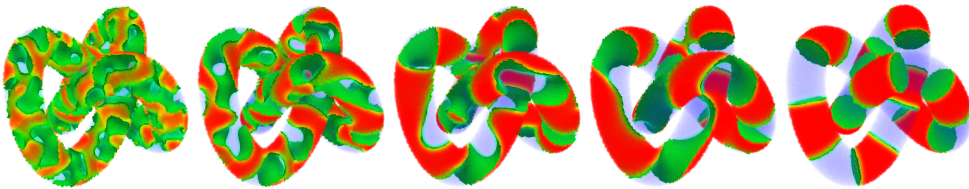

FIG. 2. *Selected snapshots at time steps* 2^n, *where* $n = 3, 5, \cdots, 11$. *3D views of the evolution of order parameter field.*

4.3. Micro structure simulations. This example involves large Péclet flows in a microfluidic device, making it an interesting test for validating our bound-preserving scheme in simulating advection-dominated CHNS problems. In this part, we use the constant mobility with GL polynomial potential.

The microstructure image is a set of $334 \times 210 \times 10$ cubic cells of resolution $h = 1/350$. Analogous to the lab experiment setup, we add a buffer of $16\times210\times70$ cells to the left side. The pore space together with the buffer region form our computational domain Ω, see Figure 4. We refer to phase A the bulk phase with order parameter equals to $+1$ and phase B the bulk phase with order parameter equals to -1. The buffer zone is initially filled with phase A and the microstructure is initially filled with
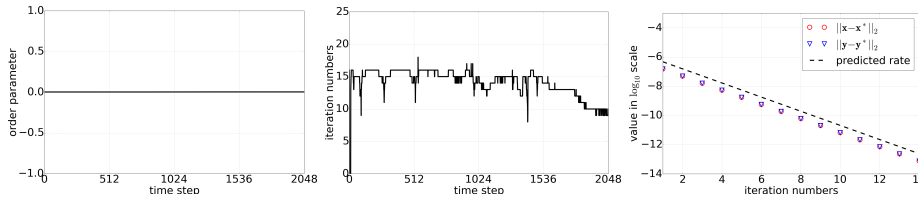
FIG. 3. *Left: the average of order parameter at each time step, which shows the conservation is preserved. Middle: the number of Douglas–Rachford iterations at each time step. Right: the asymptotic linear convergence at time step* 128. *The predicted rate is the rate given in Theorem 2.5.*

phase B, respectively. The initial velocity field is taken to be zero. The left boundary of Ω is inflow, the right boundary of Ω is outflow, and the rest boundaries of Ω are fluid–solid interfaces. On the inflow boundary, we prescribe $\phi_D = 1$, e.g., the phase A is injected, and $v_D = \frac{10000}{9}(y - 0.2)(y - 0.8)(z - 0.4)(z - 0.6)$. We the take time step size $\tau = 5 \times 10^{-4}$. For physical parameters, we choose Re = 1, Ca = 1, Pe = 100, and Cn = h. The microstructure surface is hydrophobic with respect to phase A with a contact angle $\vartheta = 135°$. The buffer surface and outflow boundary are neutral, namely $\vartheta = 90°$.

Figure 5 shows snapshots of the order parameter field as well as its values along the plane $\{(x, y, z) \in \Omega : z = 0.5\}$ in mountain views. Similar to the previous example, we employ a rainbow color scale that maps the values in $[-1, 1]$ from blue to red for plotting the order parameter field. The center of the diffusive interface is colored green. The values outside $[-1, 1]$ are marked in black. We observe that phase A invades the microstructure while staying away from the solid surfaces due to the wettability constraint. The top two rows correspond to the simulation without applying any limiter whereas the bottom two rows correspond to the simulation applying our two-stage limiting strategy. Our limiters remove overshoot and undershoot. The fluid dynamics are similar for both cases.

Figure 6 shows the number of iterations of the Douglas–Rachford algorithm on each time step as well as the asymptotic linear convergence rates of selected time steps. Here, the errors $\|y^k - y^*\|_2$ and $\|x^k - x^*\|_2$ are measured in a similar way as explained in the previous example. A numerical way of getting an exact value of r is to run the Douglas–Rachford iterations sufficiently many times with small enough tolerance and count the number of entries that stay out of the bounds in y^*. Using the exact r to compute the principal angle θ_{N-r}, the numerical results match our analysis, see Figure 6.
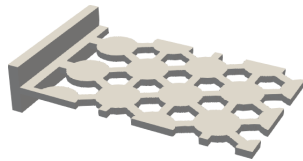

FIG. 4. *The computational domain of the microstructure simulation.*

4.4. Merging droplets. This example deals with droplets of fluid surrounded by another fluid. In a capillary-forces-dominated merging process, the large droplet wobbles several times and eventually evolves into the most thermodynamically favorable configuration, e.g., a single spherical droplet.
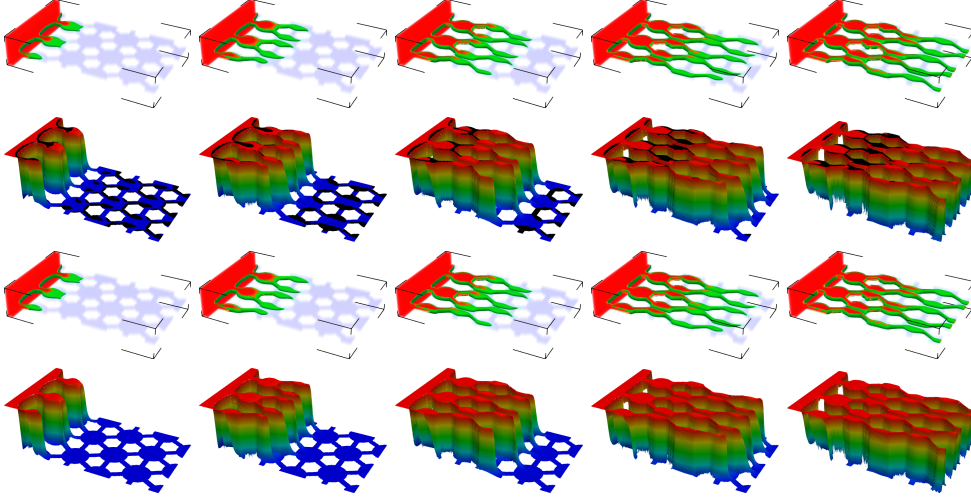
Fig. 5. *Selected snapshots at time steps* 50, 100, 150, 200, *and* 250. *The first and third rows: 3D views of the evolution of the order parameter field. The second and fourth rows: plots of order parameter warped along the plane* {z = 0.5}. *The top two rows are without limiters and the bottom two rows are with our limiters.*
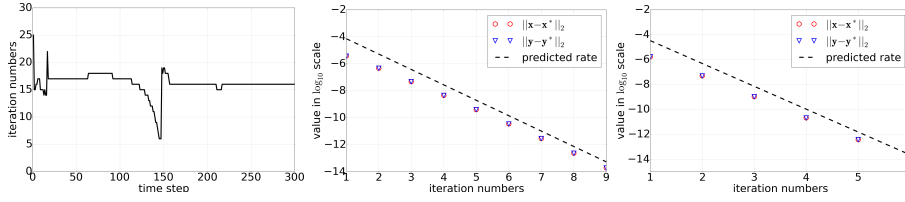

Fig. 6. *The left top figure shows the number of Douglas–Rachford iterations at each time step. The middle and right figures show the asymptotic linear convergence at time steps* 150 *and* 250, *where the principal angle* θ_{N-r} *is computed by using exact values of* r.

Let us consider four different scenarios. In the first scenario, we use constant mobility with GL polynomial potential and we do not apply any limiter. In the rest scenarios, we apply our two-stage limiting strategy. In the second scenario, we use constant mobility with GL polynomial potential. In the third scenario, we use constant mobility with FH logarithmic potential (parameters $\alpha = 0.3$ and $\beta = 1$). And in the fourth scenario, we use degenerate mobility with GL polynomial potential.

Let $\Omega = (0,1)^3$ to be a closed system, $\partial\Omega = \partial\Omega^{\text{wall}}$ and set the initial velocity field $\boldsymbol{v}^0 = \boldsymbol{0}$. Four droplets of phase A are initially in a non-equilibrium configuration, surrounded by phase B, i.e., the initial order parameter field is prescribed by

$$\phi^0 = \max\left\{-1, \tanh\left(\frac{r_1 - \|\boldsymbol{x} - \boldsymbol{a}_0\|}{\sqrt{2}\,\text{Cn}}\right), \tanh\left(\frac{r_1 - \|\boldsymbol{x} - \boldsymbol{a}_1\|}{\sqrt{2}\,\text{Cn}}\right), \tanh\left(\frac{r_2 - \|\boldsymbol{x} - \boldsymbol{a}_2\|}{\sqrt{2}\,\text{Cn}}\right), \tanh\left(\frac{r_2 - \|\boldsymbol{x} - \boldsymbol{a}_3\|}{\sqrt{2}\,\text{Cn}}\right)\right\},$$

where $\boldsymbol{a}_0 = [0.35, 0.35, 0.35]^{\text{T}}$ and $\boldsymbol{a}_1 = [0.65, 0.65, 0.65]^{\text{T}}$ are the centers of the two initial larger droplets with radius $r_1 = 0.25$; and $\boldsymbol{a}_2 = [0.75, 0.25, 0.25]^{\text{T}}$ and $\boldsymbol{a}_3 = [0.25, 0.75, 0.75]^{\text{T}}$ are the centers of the two initial smaller droplets with radius $r_2 = 0.16$. For the FH logarithmic potential, we use $0.997\phi^0$ as the initial order parameter field to make its value away from the singularity. We uniformly partition domain Ω by cubic elements with the mesh resolution $h = 1/50$ and take the time

step size $\tau = 10^{-4}$. For physical parameters, we choose Re = 1, Ca = 10^{-4}, Pe = 1, Cn = h, and the contact angle $\vartheta = 90°$ on $\partial\Omega$.

Figure 7 shows snapshots of the order parameter field. The center of the diffusive interface is colored green and the bulk phases are colored transparent. We see the merging of the four droplets, the intermediate wobbling stages, and the final equilibrium configuration of a spherical droplet. We observe from Figure 7 that the fluid dynamics are visually similar in these scenarios. However, there are visible differences in certain one dimensional profiles, see Figure 8 for the order parameters at the line $\{(x, y, z) \in \Omega : x = y = z\}$.

Figure 8 shows values of order parameter along the diagonal $\{(x, y, z) \in \Omega : x = y = z\}$ of the computational domain. In scenario 1, we observe bulk shift at near steady state, which is as expected since no limiters are applied. In secnarios 2 and 4, our limiters remove overshoots and undershoots. In scenario 3, the FH logarithmic potential ensures bounds without bulk shift. The cell average limiter (1.7) is not triggered but the Zhang–Shu limiter is triggered. The global mass is conserved, see the left subfigure in Figure 9.

We plot the number of iterations of the Douglas–Rachford algorithm on each time step, see the right two subfigures in Figure 9. We check the asymptotic linear convergence rates and they match with our analysis. The errors $\|y^k - y^*\|_2$ and $\|x^k - x^*\|_2$ are measured in a similar way as in the previous example.

5. Conclusion. In this paper, we have analyzed the asymptotic linear convergence rate for using Douglas–Rachford splitting methods of a simple nonsmooth convex minimization, which forms a high order accurate cell average limiter. We obtain an explicit dependence of the convergence rate on the parameters, which gives a principle of parameter selection for accelerating the asymptotic convergence rate. Our optimization scheme is efficient and our two-stage limiting strategy is well-suited for high order accurate DG schemes for large-scale simulations.

REFERENCES

[1] R. BAILO, J. CARRILLO, S. KALLIADASIS, AND S. PEREZ, *Unconditional bound-preserving and energy-dissipating finite-volume schemes for the Cahn-Hilliard equation*, Communications in Computational Physics, (2023).

[2] J. W. BARRETT, J. F. BLOWEY, AND H. GARCKE, *Finite element approximation of the Cahn–Hilliard equation with degenerate mobility*, SIAM Journal on Numerical Analysis, 37 (1999), pp. 286–318.

[3] A. BECK, *First-Order Methods in Optimization*, SIAM, 2017.

[4] A. CARLSON, M. DO-QUANG, AND G. AMBERG, *Dissipation in rapid dynamic wetting*, Journal of Fluid Mechanics, 682 (2011), pp. 213–240.

[5] A. CHAMBOLLE AND T. POCK, *An introduction to continuous optimization for imaging*, Acta Numerica, 25 (2016), pp. 161–319.

[6] W. CHEN, C. WANG, X. WANG, AND S. M. WISE, *Positivity-preserving, energy stable numerical schemes for the Cahn–Hilliard equation with logarithmic potential*, Journal of Computational Physics: X, 3 (2019), p. 100031.

[7] Q. CHENG AND J. SHEN, *A new Lagrange multiplier approach for constructing structure preserving schemes, I. Positivity preserving*, Computer Methods in Applied Mechanics and Engineering, 391 (2022), p. 114585.

[8] Q. CHENG AND J. SHEN, *A new Lagrange multiplier approach for constructing structure preserving schemes, II. Bound preserving*, SIAM Journal on Numerical Analysis, 60 (2022), pp. 970–998.

[9] L. DEMANET AND X. ZHANG, *Eventual linear convergence of the Douglas–Rachford iteration for basis pursuit*, Mathematics of Computation, 85 (2016), pp. 209–238.

[10] Q. DU, L. JU, X. LI, AND Z. QIAO, *Maximum bound principles for a class of semilinear parabolic equations and exponential time-differencing schemes*, SIAM Review, 63 (2021),
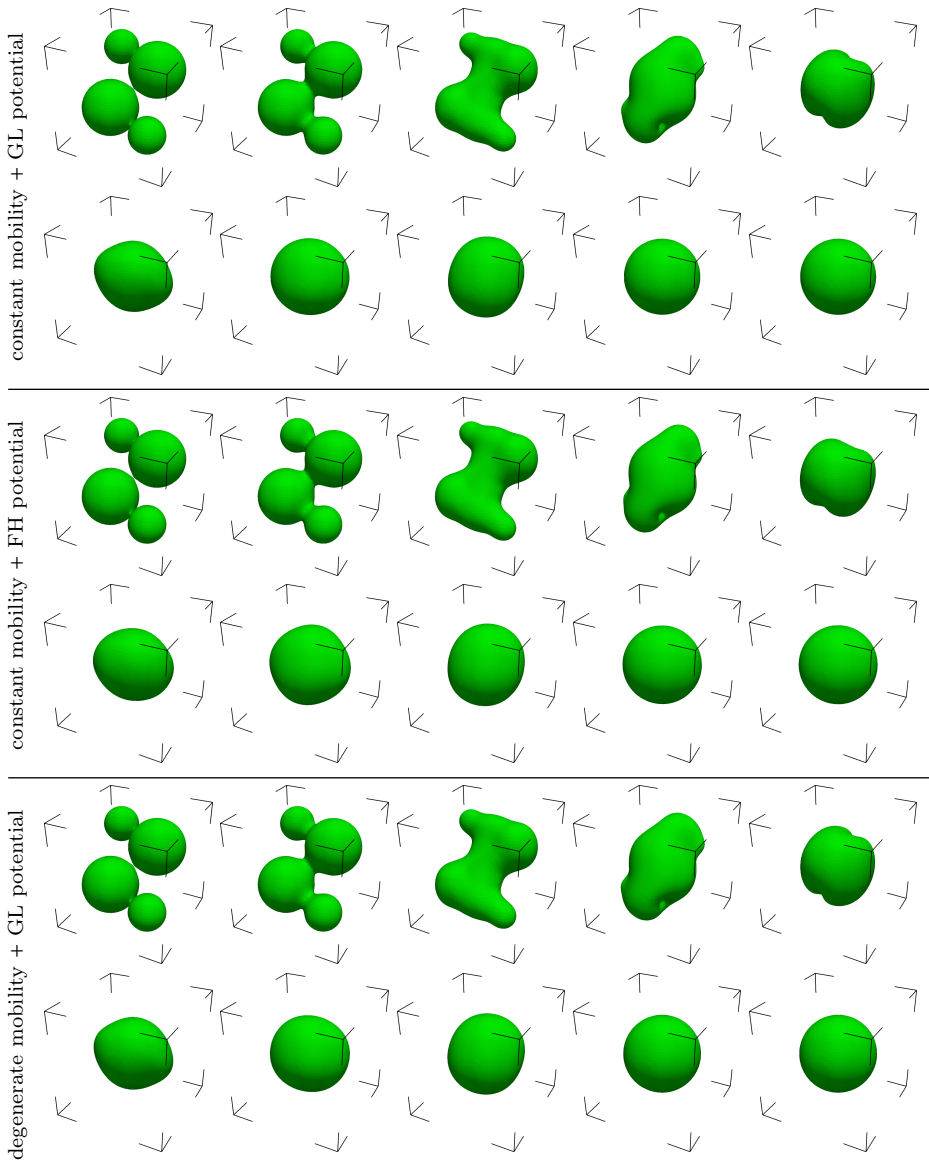
FIG. 7. *3D views of the evolution of the order parameter field. Selected snapshots at time steps 1, 3, 11, 23, 39, 56, 72, 90, 256, and 512. The dynamics are visually similar in these scenarios. However, there are visible differences in certain 2D profiles, see Figure 8.*

pp. 317–359.

[11] A. ERN AND J.-L. GUERMOND, *Invariant-Domain-Preserving High-Order Time Stepping: I. Explicit Runge–Kutta Schemes*, SIAM Journal on Scientific Computing, 44 (2022), pp. A3366–A3392.

[12] C. FAN, X. ZHANG, AND J. QIU, *Positivity-preserving high order finite difference WENO schemes for compressible Navier-Stokes equations*, Journal of Computational Physics, 467 (2022), p. 111446.

[13] M. FORTIN AND R. GLOWINSKI, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-value Problems*, Elsevier, 2000.

[14] F. FRANK, C. LIU, F. ALPAK, AND B. RIVIERE, *A finite volume/discontinuous Galerkin method for the advective Cahn–Hilliard equation with degenerate mobility on porous domains stem-*
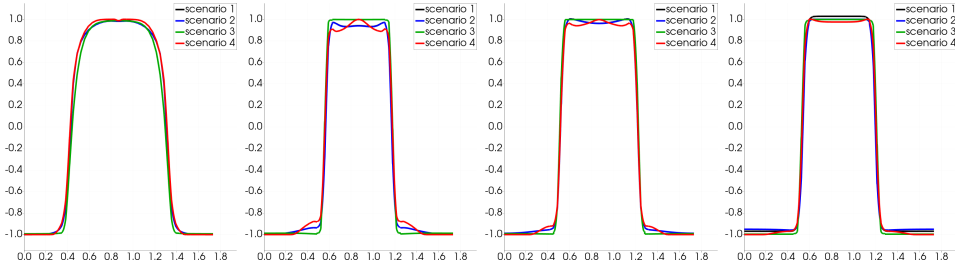
Fig. 8. *Plots of order parameter extracted along the line $\{(x, y, z) \in \Omega : x = y = z\}$. Selected snapshots at time steps 23, 56, 90, and 512. Scenario 1: constant mobility with GL polynomial potential and do not apply any limiter. The rest scenarios apply limiters. Scenario 2: constant mobility with GL polynomial potential. Scenario 3: constant mobility with FH logarithmic potential. Scenario 4: degenerate mobility with GL polynomial potential.*
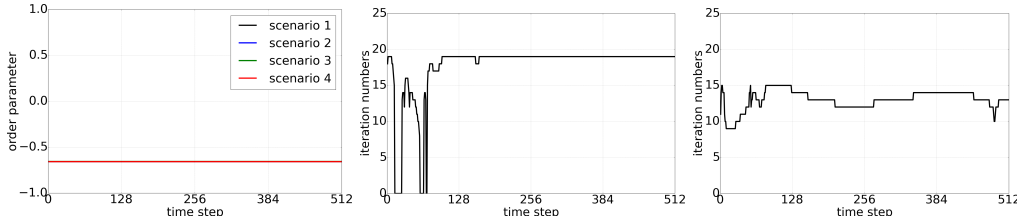

Fig. 9. *Left: the average of order parameter at each time step. Middle and right: the number of Douglas–Rachford iterations for scenario 2 and 4 at each time step. Scenario 1: constant mobility with GL polynomial potential and do not apply any limiter. The rest scenarios apply limiters. Scenario 2: constant mobility with GL polynomial potential. Scenario 3: constant mobility with FH logarithmic potential. Scenario 4: degenerate mobility with GL polynomial potential.*

ming from micro-CT imaging, Computational Geosciences, 22 (2018), pp. 543–563.

[15] F. Frank, C. Liu, F. O. Alpak, S. Berg, and B. Riviere, *Direct numerical simulation of flow on pore-scale images using the phase-field method*, SPE Journal, 23 (2018), pp. 1833–1850.

[16] F. Frank, C. Liu, A. Scanziani, F. O. Alpak, and B. Riviere, *An energy-based equilibrium contact angle boundary condition on jagged surfaces for phase-field methods*, Journal of Colloid and Interface Science, 523 (2018), pp. 282–291.

[17] F. Frank, A. Rupp, and D. Kuzmin, *Bound-preserving flux limiting schemes for DG discretizations of conservation laws with applications to the Cahn–Hilliard equation*, Computer Methods in Applied Mechanics and Engineering, 359 (2020), p. 112665.

[18] V. Girault, B. Riviere, and M. Wheeler, *A discontinuous Galerkin method with nonoverlapping domain decomposition for the Stokes and Navier-Stokes problems*, Mathematics of Computation, 74 (2005), pp. 53–84.

[19] R. Glowinski, *Finite element methods for incompressible viscous flow*, Handbook of Numerical Analysis, 9 (2003), pp. 3–1176.

[20] T. Goldstein and S. Osher, *The split Bregman method for L1-regularized problems*, SIAM Journal on Imaging Sciences, 2 (2009), pp. 323–343.

[21] J.-L. Guermond, P. Minev, and J. Shen, *An overview of projection methods for incompressible flows*, Computer Methods in Applied Mechanics and Engineering, 195 (2006), pp. 6011–6045.

[22] J.-L. Guermond, B. Popov, and I. Tomas, *Invariant domain preserving discretization-independent schemes and convex limiting for hyperbolic systems*, Computer Methods in Applied Mechanics and Engineering, 347 (2019), pp. 143–175.

[23] F. Huang, J. Shen, and K. Wu, *Bound/positivity preserving and unconditionally stable schemes for a class of fourth order nonlinear equations*, Journal of Computational Physics, 460 (2022), p. 111177.

[24] A. V. Knyazev and M. E. Argentati, *Majorization for changes in angles between subspaces, Ritz values, and graph Laplacian spectra*, SIAM Journal on Matrix Analysis and Applications, 29 (2007), pp. 15–32.

[25] D. KUZMIN, *Explicit and implicit FEM-FCT algorithms with flux linearization*, Journal of Computational Physics, 228 (2009), pp. 2517–2534.

[26] P.-L. LIONS AND B. MERCIER, *Splitting algorithms for the sum of two nonlinear operators*, SIAM Journal on Numerical Analysis, 16 (1979), pp. 964–979.

[27] C. LIU, F. FRANK, F. O. ALPAK, AND B. RIVIERE, *An interior penalty discontinuous Galerkin approach for 3D incompressible Navier–Stokes equation for permeability estimation of porous media*, Journal of Computational Physics, 396 (2019), pp. 669–686.

[28] C. LIU, F. FRANK, C. THIELE, F. O. ALPAK, S. BERG, W. CHAPMAN, AND B. RIVIERE, *An efficient numerical algorithm for solving viscosity contrast Cahn–Hilliard–Navier–Stokes system in porous media*, Journal of Computational Physics, 400 (2020), p. 108948.

[29] C. LIU, R. MASRI, AND B. RIVIERE, *Convergence of a decoupled splitting scheme for the Cahn–Hilliard–Navier–Stokes system*, SIAM Journal on Numerical Analysis (to appear), (2023). arXiv:2210.05625.

[30] C. LIU, D. RAY, C. THIELE, L. LIN, AND B. RIVIERE, *A pressure-correction and bound-preserving discretization of the phase-field method for variable density two-phase flows*, Journal of Computational Physics, 449 (2022), p. 110769.

[31] C. LIU AND X. ZHANG, *A positivity-preserving implicit-explicit scheme with high order polynomial basis for compressible Navier–Stokes equations*, arXiv:2305.05769, (2023).

[32] R. MASRI, C. LIU, AND B. RIVIERE, *A discontinuous Galerkin pressure correction scheme for the incompressible Navier–Stokes equations: Stability and convergence*, Mathematics of Computation, 91 (2022), pp. 1625–1654.

[33] R. MASRI, C. LIU, AND B. RIVIERE, *Improved a priori error estimates for a discontinuous Galerkin pressure correction scheme for the Navier–Stokes equations*, Numerical Methods for Partial Differential Equations, 39 (2023), pp. 3108–3144.

[34] Y. NESTEROV, *Gradient methods for minimizing composite functions*, Mathematical Programming, 140 (2013), pp. 125–161.

[35] T. QIN AND C.-W. SHU, *Implicit positivity-preserving high-order discontinuous Galerkin methods for conservation laws*, SIAM Journal on Scientific Computing, 40 (2018), pp. A81–A107.

[36] B. RIVIERE, *Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation*, SIAM, 2008.

[37] J. SHEN, *Modeling and numerical approximation of two-phase incompressible flows by a phase-field approach*, in Multiscale Modeling and Analysis for Materials Aimulation, World Scientific, 2012, pp. 147–195.

[38] J. SHEN AND X. YANG, *Decoupled, energy stable schemes for phase-field models of two-phase incompressible flows*, SIAM Journal on Numerical Analysis, 53 (2015), pp. 279–296.

[39] S. SRINIVASAN, J. POGGIE, AND X. ZHANG, *A positivity-preserving high order discontinuous Galerkin scheme for convection–diffusion equations*, Journal of Computational Physics, 366 (2018), pp. 120–143.

[40] Z. SUN, J. A. CARRILLO, AND C.-W. SHU, *A discontinuous Galerkin method for nonlinear parabolic equations and gradient flow problems with interaction potentials*, Journal of Computational Physics, 352 (2018), pp. 76–104.

[41] G. TIERRA AND F. GUILLÉN-GONZÁLEZ, *Numerical methods for solving the Cahn–Hilliard equation and its applicability to related energy-based models*, Archives of Computational Methods in Engineering, 22 (2015), pp. 269–289.

[42] Z. XU, *Parametrized maximum principle preserving flux limiters for high order schemes solving hyperbolic conservation laws: one-dimensional scalar problem*, Mathematics of Computation, 83 (2014), pp. 2213–2238.

[43] X. ZHANG, *On positivity-preserving high order discontinuous Galerkin schemes for compressible Navier–Stokes equations*, Journal of Computational Physics, 328 (2017), pp. 301–343.

[44] X. ZHANG AND C.-W. SHU, *On maximum-principle-satisfying high order schemes for scalar conservation laws*, Journal of Computational Physics, 229 (2010), pp. 3091–3120.

[45] X. ZHANG AND C.-W. SHU, *On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes*, Journal of Computational Physics, 229 (2010), pp. 8918–8934.