SIAM J. SCI. COMPUT.
Vol. 40, No. 3, pp. B982–B1006


# NEWTON SOLVERS FOR DRIFT-DIFFUSION AND ELECTROKINETIC EQUATIONS*

ARTHUR BOUSQUET†, XIAOZHE HU‡, MAXIMILIAN S. METTI§, AND JINCHAO XU†

**Abstract.** A Newton solver for equations modeling drift-diffusion and electrokinetic phenomena is investigated. For drift-diffusion problems, modeled by the nonlinear Poisson–Nernst–Planck (PNP) equations, the linearization of the model equations is shown to be well-posed. Furthermore, a fast solver for the linearized PNP and electrokinetic equations is proposed and numerically demonstrated to be effective on some physically motivated benchmarks. This work builds on a formulation of the PNP and electrokinetic equations that is investigated in [M. S. Metti, J. Xu, and C. Liu, *J. Comput. Phys.*, 306 (2016), pp. 1–18] and shown to have some favorable stability properties.

**Key words.** finite elements, Poisson–Nernst–Planck, stability analysis, numerical solvers

**AMS subject classifications.** 35K61, 65L20, 65M15, 65M50, 65M60

**DOI.** 10.1137/17M1146956

## 1. Introduction.

Drift-diffusion equations are often used to model physical phenomena where the flux of a charge carrier, such an as ion or electron, is described by a generalized diffusive flux that is strongly coupled with an electric field. In such systems, the diffusion is inherently nonlinear and nonlocal, due to the nature of the electric field.

Computational studies for these types of systems date back to the 1960s [18, 22], and have been regularly investigated due to their application to semiconductor physics [2, 6, 11, 17, 27, 28, 38, 39, 44, 52, 55]. Since then, other applications have also been explored using equations of these types that study applications to batteries [41, 58], physical chemistry [33], and biology [19, 23, 24, 35, 56, 62].

Electrokinetic systems are modeled by augmenting the drift-diffusion system by coupling an additional kinetic force from a fluidic velocity field in order to model the charge carriers being suspended in a fluid. The result is a multiphysics system that serves as the foundation for studying electrokinetic devices, as in [21, 29, 32, 46].

Many formulations and discretizations have been investigated for solving these drift-diffusion and electrokinetic equations, including changes of variables to Slotboom variables [55, 56] and quasi-Fermi variables [6, 11, 28]. Common discretizations include finite difference or finite element schemes for drift-diffusion and, for electrokinetics, smoothed particle hydrodynamics methods have also been investigated [33, 46]. These discretizations often are modified to improve the conditioning of these typically convection-dominated equations using upwinding techniques or edge-based methods, or modifying the charge carrier flux [5, 6, 9, 22, 40, 44].

*Submitted to the journal's Computational Methods in Science and Engineering section September 12, 2017; accepted for publication (in revised form) March 27, 2018; published electronically June 26, 2018.

http://www.siam.org/journals/sisc/40-3/M114695.html

**Funding:** The work of the authors was supported by DOE grant DE-SC0009249 as part of the Collaboratory on Mathematics for Mesoscopic Modeling of Materials.

†Department of Mathematics, The Pennsylvania State University, University Park, PA 16802 (akb5670@psu.edu, http://personnal.edu/~akb5670/, xu@math.psu.edu, http://www.math.psu.edu/xu/).

‡Department of Mathematics, Tufts University, Medford, MA 02155 (xiaozhe.hu@tufts.edu, http://math.tufts.edu/faculty/xhu/).

§Adtile Technologies, San Diego, CA 92121 (maxx@adtile.me).

This work uses a finite element discretization that employs a special change of variables that has been recently shown to have some favorable stability properties for the electrokinetic equations as well as guaranteeing positivity of computed densities of charge carriers [42]. A Newton approach is used to linearize the model equations and a proof of the well-posedness of the linearized drift-diffusion equations is given. An edge-averaged finite element approximation [60] is used for the linearized system to provide additional stability.

Section 2 provides the relevant notation, preliminaries, and discretization details needed to properly describe the proposed solver. In section 3, a Newton solver is outlined and the well-posedness of the resulting linearized equations for the drift-diffusion subsystem is established; section 3 continues, presenting a stabilization scheme based on edge-averaged finite elements [60] and a fast solver for the linear system that arises from the Newton scheme. Numerical experiments that demonstrate the accuracy and applicability of the solver to physical benchmarks are presented and discussed in section 4. Section 5 closes with some final remarks and proposed future work.

**2. The PNP models and their discretizations.** Consider a system of $N \geq 2$ charge carrying species, where the density of the $i$th species is denoted by $\rho_i > 0$ and the sign and strength of its charge, known as its valency, are given by $q_i$. Denote the electrostatic potential by $\phi$. The charge carriers have nonlocal spatial interaction with one another and the strength of these nonlocal interactions is described by the electric permittivity coefficient, $\epsilon > 0$. In addition to these mobile charge carriers, there may also be some fixed charges present in the system, which are given by some signed and bounded quantity, $\rho_0$, that represents the total charge of these immobile charge carriers.

Let $\Omega \subset \mathbb{R}^d$ for $d = 1, 2$, or 3, and $T$ be a positive and finite real number. Then, the PNP system is described by the initial-boundary value problem:

$$-\nabla \cdot (\epsilon \nabla \phi) = \sum_{i=1}^{N} q_i \rho_i + \rho_0, \tag{1}$$

$$\frac{\partial \rho_i}{\partial t} = -\nabla \cdot \vec{J}_i, \qquad i = 1, \ldots, N, \tag{2}$$

in $\Omega \times (0, T]$, where

$$\rho_i(x, 0) = \rho_{i,0}(x) \qquad \text{for } x \in \Omega,\ i = 1, \ldots, N,$$

and

$$-\nabla \cdot \big(\epsilon \nabla \phi(x, 0)\big) = \sum_{i=1}^{N} q_i \rho_{i,0}(x) + \rho_0(x) \qquad \text{for } x \in \Omega.$$

To complete the description of this system, the charge carrier fluxes, $\vec{J}_i$, need to be defined.

This paper describes solvers for two choices of charge carrier fluxes. The first and simplest is the drift-diffusion flux

$$\vec{J}_i = -D_i(\nabla \rho_i + q_i \rho_i \nabla \phi), \tag{3}$$

where $D_i$ is a positive coefficient that represents the diffusivity of the $i$th charge carrying species of valence $q_i$. Note that the electric field is described in terms of

the potential, $\vec{E} = -\nabla\phi$ (and proper scaling has taken place for $\phi$ to suppress the coefficients in the Einstein relation). When the charge carriers are suspended in some velocity field, denoted $\vec{u}$, the flux contains an additional kinetic term:

$$\vec{J_i} = -D_i(\nabla\rho_i + q_i\rho_i\nabla\phi) + \rho_i\vec{u}. \tag{4}$$

**2.1. The drift-diffusion model.** It is most natural to begin by understanding the drift-diffusion model, which describes a subsystem of the electrokinetic model. In [42], a finite element discretization using a change of variables for these equations is analyzed and shown to be stable. Denote the log-density of the charge carriers by

$$\eta_i = \log\rho_i \quad \text{so that} \quad \nabla\rho_i = e^{\eta_i}\nabla\eta_i.$$

The corresponding PNP equations are given by

$$-\nabla\cdot(\epsilon\nabla\phi) = \sum_{i=1}^{N} q_i e^{\eta_i} + \rho_0, \tag{5}$$

$$\frac{\partial e^{\eta_i}}{\partial t} = \nabla\cdot\big(D_i e^{\eta_i}\nabla(\eta_i + q_i\phi)\big), \qquad i = 1,\dots,N, \tag{6}$$

in $\Omega\times(0,T]$.

The boundary conditions for this system are crucial to determining the qualitative behavior of the solution to the PNP system. For the charge carriers and electrostatic potential, some mix of Dirichlet and no-flux boundary conditions is imposed. Let $\partial\Omega = \Gamma_N\cup\Gamma_D$ partition the boundary into two disjoint sets, where

$$\big[D_i\nabla(\eta_i + q_i\phi)\big]\cdot\vec{n} = 0 \qquad \text{on } \Gamma_N, \tag{7}$$

$$\epsilon\nabla\phi\cdot\vec{n} = S \qquad \text{on } \Gamma_N, \tag{8}$$

$$\eta_i = \eta_{i,D} \qquad \text{on } \Gamma_D, \tag{9}$$

$$\phi = \delta V \qquad \text{on } \Gamma_D, \tag{10}$$

with $\vec{n}$ denoting the outward unit normal vector to $\Gamma_N$. The terms $\eta_{i,D}$ model the bulk density of ions in an ionic solution or contact boundary conditions in a semiconductor. The voltage-drop is given by $\delta V$, which models an applied electric field. Surface charges on the boundary of the domain are modeled by the Neumann boundary condition $S$. In general, one may wish to consider pure no-flux boundary conditions for the charge carriers or a Robin-type boundary condition on the electrostatic potential, which models a capacitor at the boundary; such boundary conditions lead to no significant modifications to this investigation.

A method of lines approach is taken to discretize the system (5)–(10), which permits the numerical scheme to solve for a computed solution by solving a sequence of elliptic equations. Let

$$\mathcal{L}_2(\Omega) \equiv \left\{ w : \Omega\to\mathbb{R} \,|\, \int_\Omega w^2\,dx < \infty \right\},$$
$$\mathcal{H}^1(\Omega) \equiv \left\{ w\in\mathcal{L}_2(\Omega) \,|\, \int_\Omega |\nabla w|^2\,dx < \infty \right\},$$
$$\mathcal{H}^1_{\Gamma_D} \equiv \big\{ w\in\mathcal{H}^1(\Omega) \,|\, w|_{\Gamma_D} = 0 \big\}$$

be the spaces with associated norms:

$$\|w\|_0 \equiv \left(\int_\Omega w^2\,dx\right)^{1/2}, \quad |w|_1 \equiv \left(\int_\Omega |\nabla w|^2\,dx\right)^{1/2}, \quad \|w\|_1^2 \equiv |w|_1^2 + \|w\|_0^2,$$

and $\|w\|_\infty \equiv \sup_{x\in\Omega} |w(x)|$.

Partition the time domain

$$0 = t_0 < t_1 < \cdots < t_m = T,$$

with $\Delta t_j \equiv t_j - t_{j-1}$ for $j = 1, \ldots, m$. Then, the solution to the weak formulation of the semidiscrete solution is given by $\eta_i(t_j) = \eta_i^{(j)}, \phi(t_j) = \phi^{(j)} \in \mathcal{H}^1(\Omega)$ such that

$$(\epsilon\nabla\phi^{(j)}, \nabla\psi) - \sum_{i=1}^{N} q_i\big(e^{\eta_i^{(j)}}, \psi\big) = (\rho_0, \psi) \qquad \text{for all } \psi \in \mathcal{H}^1_{\Gamma_D}, \tag{11}$$

$$\big(D_i e^{\eta_i^{(j)}} \nabla(\eta_i^{(j)} + q_i\phi^{(j)}), \nabla w\big) + \frac{1}{\Delta t_i}\big(e^{\eta_i^{(j)}}, w\big) = \frac{1}{\Delta t_j}\big(e^{\eta_i^{(j-1)}}, w\big) \quad \text{for all } w \in \mathcal{H}^1_{\Gamma_D}, \tag{12}$$

for $j = 1, \ldots, m$, with $\eta_i^{(j)}|_{\Gamma_D} = \eta_{i,D}$ and $\phi^{(j)}|_{\Gamma_D} = \delta V$. For time dependent problems, the initial conditions must be prescribed:

$$\big(e^{\eta_i^{(0)}}, w\big) = \big(e^{\eta_{i,0}}, w\big) \quad \text{and} \quad (\epsilon\nabla\phi, \nabla\psi) = (\rho_0, \psi) + \sum_{i=1}^{N} q_i\big(e^{\eta_{i,0}}, \psi\big)$$

for all $w, \psi \in \mathcal{H}^1_{\Gamma_D}$. For solutions at equilibrium, the charge carriers satisfy the condition $\eta_i^{(j)} = \eta_i^{(j-1)}$ so that the terms corresponding to the time derivative cancel out in (12). If no Dirichlet boundary conditions are imposed for the steady state problem, the charge carrier densities are defined up to an additive constant so that the total mass of each species is needed, $\int_\Omega e^{\eta_i}\, dx = m_i$, for the solution to be well-defined. In order to be consistent with (11) for $\psi = 1$, the total charge must be zero, $\sum_{i=1}^{N} q_i m_i + (\rho_0, 1) = 0$. This constraint must also be satisfied by the initial condition in the time dependent problem with no Dirichlet boundary condition, for consistency.

In general, the equations that need to be solved are for $\eta_i, \phi \in \mathcal{H}^1(\Omega)$:

$$(\epsilon\nabla\phi, \nabla\psi) - \sum_{i=1}^{N} q_i\big(e^{\eta_i}, \psi\big) = (\rho_0, \psi) \qquad \text{for all } \psi \in \mathcal{H}^1_{\Gamma_D}, \tag{13}$$

$$\big(D_i e^{\eta_i} \nabla(\eta_i + q_i\phi), \nabla w\big) + \big(R(\eta_i), w\big) = \big(G_i, w\big) \qquad \text{for all } w \in \mathcal{H}^1_{\Gamma_D}, \tag{14}$$

where $R(\eta_i) = \theta\eta_i$, with $\theta \geq 0$, and $G_i \geq 0$ is a given and bounded expression. For steady problems, $\theta = G_i = 0$; otherwise, $\theta = \Delta t_j^{-1}$ and $G_i$ is the exponential of the solution at the previous time step. (The semiconductor equations also follow a similar form, where $R$ and $G$ instead represent the generation and recombination of electrons and holes; however, these terms in the semiconductor equations take a more general form, may be negative, and may introduce additional coupling of charge carriers. This work does not aim to treat the additional complexities required in the analysis of the semiconductor equations.)

To define the finite element solution, let $W_h \subset \mathcal{H}^1(\Omega)$ and $W_{h,0} \subset \mathcal{H}^1_{\Gamma_D}$ denote continuous piecewise linear finite element spaces defined on some triangulation ($d = 2$) or tetrahedralization ($d = 3$) of $\Omega$. Then, the finite element solution $\eta_{i,h}, \psi_h \in W_h$ must satisfy

$$(\epsilon\nabla\phi_h, \nabla\psi_h) - \sum_{i=1}^{N} q_i\big(e^{\eta_{i,h}}, \psi_h\big) = (\rho_0, \psi_h) \quad \text{for all } \psi_h \in W_{h,0}, \tag{15}$$

$$\big(D_i e^{\eta_{i,h}} \nabla(\eta_{i,h} + q_i\phi_h), \nabla w_h\big) + \big(R(\eta_{i,h}), w_h\big) = \big(G_i, w_h\big) \quad \text{for all } w \in W_{h,0}, \tag{16}$$

where the boundary conditions are imposed via some interpolation operator, $\mathcal{I}_h$, with $\eta_{i,h} = \mathcal{I}_h(\eta_{i,D})$, $\phi_h = \mathcal{I}_h(\delta V)$ on $\Gamma_D$. For time dependent problems, such equations must be solved at each time step and the initial conditions are given by projection,

$$\left(e^{\eta_{i,h}^{(0)}}, w_h\right) = \left(e^{\eta_{i,0}}, w_h\right) \quad \text{and} \quad (\epsilon\nabla\phi_h, \nabla\psi_h) = (\rho_0, \psi_h) + \sum_{i=1}^{N} q_i\left(e^{\eta_{i,0}}, \psi_h\right)$$

for all $w_h, \psi_h \in W_{h,0}$. The stability of this finite element discretization has been established for the steady state solution in [28] with mixed Dirichlet and homogeneous Neumann boundary conditions on the electrostatic potential and charge carriers. The stability of this discretization for the time dependent problem has been established in [42] with mixed inhomogeneous Dirichlet and Neumann boundary conditions on the electrostatic potential and no-flux boundary conditions on the charge carriers. Both of these stability results require the finite element mesh to be Delaunay in the case of a constant permittivity coefficient, $\epsilon$, so that the electrostatic potential satisfies a discrete maximum principle. For a variable permittivity coefficient, alternate mesh conditions must hold and are found in [30, 60]. Other geometric constraints for a tensor-valued permittivity coefficient can be derived following the procedures outlined in those papers.

**2.2. The electrokinetic model.** In addition to drift and diffusion, the electrokinetic model includes a kinetic force that is coupled to an external velocity field. While this formulation is quite general, the presented formulation exclusively considers a velocity field that is defined to be the solution of the incompressible Navier–Stokes (NS) equations with an electric body force. As the minimal equations are already long enough, the fixed charge, $\rho_0$, is taken to be zero to avoid too many terms in this presentation.

Denoting the fluidic pressure by $p$, the fluidic density by $\rho_f > 0$, and the fluidic viscosity by $\mu > 0$, the electrokinetic system is described by

$$-\nabla \cdot (\epsilon\nabla\phi) = \sum_{i=1}^{N} q_i e^{\eta_i}, \tag{17}$$

$$\frac{\partial e^{\eta_i}}{\partial t} = \nabla \cdot \left(D_i e^{\eta_i}\nabla(\eta_i + q_i\phi) - e^{\eta_i}\vec{u}\right), \qquad i = 1, \ldots, N, \tag{18}$$

$$\rho_f\left(\frac{\partial \vec{u}}{\partial t} + \vec{u}\cdot\nabla\vec{u}\right) + \nabla p = \nabla \cdot \left(2\mu\varepsilon(\vec{u})\right) + \sum_{i=1}^{N} q_i e^{\eta_i}\nabla\phi, \tag{19}$$

$$\nabla \cdot \vec{u} = 0, \tag{20}$$

in $\Omega \times (0, T]$. For the viscosity term in (19), the symmetrized gradient is used: $\varepsilon(\vec{u}) = \frac{1}{2}(\nabla\vec{u} + (\nabla\vec{u})^T)$.

The boundary conditions applied to the charge carriers and the electrostatic potential are presumed to be prescribed as in the drift-diffusion model, except that the no-flux condition on the charge carriers includes the fluidic velocity as well:

$$\left[D_i\nabla(\eta_i + q_i\phi) + \vec{u}\right] \cdot \vec{n} = 0 \quad \text{on } \Gamma_N.$$

For the NS variables, the boundary conditions model a prescribed in/out flow at the boundary, or a pressure driven flow. For the purposes of this paper, it is sufficient to consider an example where the boundary conditions give the in/out flow of the

fluid at part of the boundary, and Dirichlet boundary conditions are imposed on the remaining portion of $\partial\Omega$:

$$\begin{aligned} \vec{u}\cdot\vec{n} = u_D, \qquad & (\mu\varepsilon(\vec{u})\vec{n})\cdot\vec{t} = 0 && \text{on } \Gamma_N, \\ & \vec{u} = \vec{u}_D && \text{on } \Gamma_D, \end{aligned}$$

where $\vec{t}$ and $\vec{n}$ denote the unit tangent and normal vectors to $\partial\Omega$, respectively. (Alternative boundary conditions that do not fix the fluid velocity but, instead, prescribe a surface tension can also be used; however, this discussion needs no significant modification to handle that case and it is ignored for brevity.)

In order to establish a stable numerical scheme, it is critical to preserve the incompressibility of the fluid velocity exactly as this divergence-free property yields the perfect cancellation of the energies of the PNP and NS cross-terms (namely, the energy of the kinetic force in the charge carrier flux and the electric body force in the NS continuity equation). This requirement is elaborated upon in [42], where an energy estimate is established for the electrokinetic system with no-flux boundary conditions using a strongly divergence-free finite element discretization of the fluid velocity.

To characterize the semidiscrete solution, let

$$V = \{\vec{v} \in [\mathcal{H}^1(\Omega)]^d \,|\, \vec{v}\cdot\vec{n} = 0 \text{ on } \Gamma_N \text{ and } \vec{v} = \vec{0} \text{ on } \Gamma_D\}$$

and $Q = \mathcal{L}_2(\Omega) \setminus \mathbb{R}$. Then, the semidiscrete solution of the electrokinetic system is described by

$$(\epsilon\nabla\phi^{(j)}, \nabla\psi) - \sum_{i=1}^{N} q_i\big(e^{\eta_i^{(j)}}, \psi\big) = 0 \qquad \text{for all } \psi \in \mathcal{H}^1_{\Gamma_D}, \tag{21}$$

$$\begin{aligned} \big(D_i e^{\eta_i^{(j)}}\nabla(\eta_i^{(j)} + q_i\phi^{(j)}) - e^{\eta_i^{(j)}}\vec{u}^{(j)}, \nabla w\big) & \\ + \frac{1}{\Delta t_i}\big(e^{\eta_i^{(j)}}, w\big) = \frac{1}{\Delta t_j}\big(e^{\eta_i^{(j-1)}}, w\big) & \quad \text{for all } w \in \mathcal{H}^1_{\Gamma_D}, \end{aligned} \tag{22}$$

$$\begin{aligned} & \rho_f\big(\Delta t_j^{-1}\vec{u}^{(j)} + \vec{u}^{(j)}\cdot\nabla\vec{u}^{(j)}, \vec{v}\big) + \big(2\mu\varepsilon(\vec{u}^{(j)}), \varepsilon(\vec{v})\big) \\ & \qquad -(p^{(j)}, \nabla\cdot\vec{v}) - \sum_{i=1}^{N} q_i\big(e^{\eta_i^{(j)}}\nabla\phi^{(j)}, \vec{v}\big) = \frac{\rho_f}{\Delta t_j}\big(\vec{u}^{(j-1)}, \vec{v}\big) \qquad \text{for all } \vec{v} \in V, \end{aligned} \tag{23}$$

$$\big(\nabla\cdot\vec{u}^{(j)}, q\big) = 0 \qquad \text{for all } q \in Q, \tag{24}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, m$.

The charge carriers and electrostatic potential are again discretized by continuous piecewise linear finite elements, and the fluid velocity and pressure are discretized using *divergence-conforming* finite elements, such as Raviart–Thomas elements, Brezzi–Douglas–Marini elements, or Brezzi–Douglas–Fortin–Marini elements, all of degree $k \geq 1$. Let $V_h \times Q_h$ denote one of the aforementioned div-conforming elements so that

$$V_h \subset \mathcal{H}(\mathrm{div}) = \{\vec{v} \in [\mathcal{L}_2(\Omega)]^d \,|\, \nabla\cdot\vec{v} \in \mathcal{L}_2(\Omega)\},$$

and $\mathrm{div}V_h = Q_h$. The unisolvence for this discretization of incompressible NS and Stokes equations is discussed in [1, 14]. Due to the div-conforming property of these elements, it is straightforward to verify that the discrete fluidic velocity is strongly

divergence-free on all subdomains of $\Omega$ and, furthermore, that the normal component of the fluid velocity is continuous across interelement facets.

The discretization of the NS subsystem is nonconforming, however, meaning that $V_h \not\subset V$, and the discontinuity of the tangent components between elements requires additional penalty terms to be introduced into the continuity equation. Let $\mathcal{T}_h$ denote the triangulation or tetrahedralization of $\Omega$ and let $\mathcal{F}_h$ denote the set of all interelement facets in the mesh. The jump of $\vec{v}_h \in V_h$ between neighboring elements, $e_h^+$ and $e_h^-$, is denoted by

$$\begin{aligned}\llbracket \vec{v}_h \rrbracket &= \lim_{\delta \to 0^+} \big[\vec{v}_h(x - \delta \vec{n}_+) - \vec{v}_h(x + \delta \vec{n}_+)\big] \otimes \vec{n}_+ \equiv (\vec{v}_h^+ - \vec{v}_h^-) \otimes \vec{n}_+, \\ \{\varepsilon(\vec{v}_h)\} &\equiv \tfrac{1}{2}\big[\varepsilon(\vec{v}_h^+) + \varepsilon(\vec{v}_h^-)\big] \qquad \text{on } \mathcal{F}_h,\end{aligned}$$

where $\vec{n}_+$ denotes the outward unit normal to $e_h^+$. Then, the kinematic derivative and the viscosity terms are discretized by

$$\begin{aligned} D_{h,t}\big(\vec{w}_h; \vec{u}_h, \vec{v}_h\big) &\equiv \frac{\rho_f}{\Delta t_j}\big(\vec{u}_h, \vec{v}_h\big) - \rho_f\big(\vec{u}_h, w_h \cdot \nabla \vec{v}_h\big) \\ &\quad + \sum_{\tau \in \mathcal{T}_h} \big\langle (\vec{w}_h \cdot \vec{n}_\tau)\vec{u}_h^{\,w_h}, \vec{v}_h \big\rangle_{\partial \tau}, \\ A_h\big(\vec{u}_h, \vec{v}_h\big) &\equiv \sum_{\tau \in \mathcal{T}_h} \big(2\mu\varepsilon(\vec{u}_h), \varepsilon(\vec{v}_h)\big)_\tau + \alpha \sum_{e_h \in \mathcal{F}_h} |e_h|^{-1} \big\langle \llbracket \vec{u}_h \rrbracket, \llbracket \vec{v}_h \rrbracket \big\rangle_{e_h} \\ &\quad - 2\mu \sum_{e_h \in \mathcal{F}_h} \Big[\big\langle \{\varepsilon(\vec{u}_h)\}, \llbracket \vec{v}_h \rrbracket \big\rangle + \big\langle \{\varepsilon(\vec{v}_h)\}, \llbracket \vec{u}_h \rrbracket \big\rangle\Big], \end{aligned}$$

where the upwind flux of $\vec{u}_h$ is given by

$$\vec{u}_h^{\,w} \equiv \begin{cases} \lim_{\delta \to 0^+} \vec{u}_h(x - \delta \vec{w}), & x \in \partial\tau \setminus \Gamma_-, \\ 0, & x \in \partial\tau \cap \Gamma_-, \end{cases}$$

where $\Gamma_- \subset \partial\Omega$ is the inflow region of the boundary where $\vec{w} \cdot \vec{n} < 0$. The parameter, $\alpha \gg 1$, is the penalty parameter, which must be chosen to be sufficiently large to ensure a well-posed system. Then, the finite element solution to the NS subsystem is characterized by

$$\begin{aligned} &D_{h,t}\big(\vec{u}_h^{(j)}; \vec{u}_h^{(j)}, \vec{v}_h\big) + A_h\big(\vec{u}_h^{(j)}, \vec{v}_h\big) \\ &\quad -\big(p_h^{(j)}, \nabla \cdot \vec{v}_h\big) - \sum_{i=1}^{N} q_i\big(e^{\eta_{i,h}^{(j)}} \nabla \phi_h^{(j)}, \vec{v}_h\big) = \frac{\rho_f}{\Delta t_j}\big(\vec{u}_h^{(j-1)}, \vec{v}_h\big) \quad \text{for all } \vec{v}_h \in V_h, \\ &\qquad\qquad \big(\nabla \cdot \vec{u}_h^{(j)}, q_h\big) = 0 \qquad\qquad \text{for all } q_h \in Q_h. \end{aligned}$$

If only the steady state solution is sought, then the Stokes equation is often solved for highly viscous flows:

$$\begin{aligned} A_h\big(\vec{u}_h^{(j)}, \vec{v}_h\big) - \big(p_h^{(j)}, \nabla \cdot \vec{v}_h\big) &= \sum_{i=1}^{N} q_i\big(e^{\eta_{i,h}^{(j)}} \nabla \phi_h^{(j)}, \vec{v}_h\big) \qquad \text{for all } \vec{v}_h \in V_h, \\ \big(\nabla \cdot \vec{u}_h^{(j)}, q_h\big) &= 0 \qquad\qquad \text{for all } q_h \in Q_h. \end{aligned}$$

Since a method of lines approach is used to decouple the discretization of the temporal and spatial dimensions, the proposed solver needs only to solve a sequence

of elliptic PDEs. This motivates a formulation of the problem that encompasses both the time dependent and steady state problems, as in (15)–(16). The finite element solution to the electrokinetic system (at each time step) must satisfy

$$(\epsilon\nabla\phi_h, \nabla\psi_h) - \sum_{i=1}^{N} q_i\big(e^{\eta_{i,h}}, \psi_h\big) = 0 \qquad \text{for all } \psi_h \in W_{h,0}, \tag{25}$$

$$\begin{aligned}&\big(D_i e^{\eta_{i,h}}\nabla(\eta_{i,h} + q_i\phi_h) - e^{\eta_{i,h}}\vec{u}, \nabla w_h\big) + \big(R(\eta_{i,h}), w_h\big) = \big(G_i, w_h\big) \quad \text{for all } w \in W_{h,0},\\ &\quad D_{h,t}\big(\vec{u}_h; \vec{u}_h, \vec{v}_h\big) + A_h\big(\vec{u}_h, \vec{v}_h\big)\end{aligned} \tag{26}$$

$$-\big(p_h, \nabla\cdot\vec{v}_h\big) - \sum_{i=1}^{N} q_i\big(e^{\eta_{i,h}}\nabla\phi_h, \vec{v}_h\big) = \big(G_{\vec{u}}, \vec{v}_h\big) \qquad \text{for all } \vec{v}_h \in V_h, \tag{27}$$

$$\big(\nabla\cdot\vec{u}_h, q_h\big) = 0 \qquad \text{for all } q_h \in Q_h, \tag{28}$$

where $G_i, D_{h,t}, R$, and $G_{\vec{u}}$ are set to zero for the steady state solution.

**3. A numerical approach to solving the PNP equations.** Since most applications to engineering and multiphysics problems require a large number of degrees of freedom, it is important that the computational complexity of the numerical solver scales well with respect to this quantity. Additionally, the solver must be capable of solving nonlinear elliptic PDEs efficiently. This is particularly important for time dependent problems, as the numerical solver must run for each time step.

**3.1. The nonlinear solver.** There are two dominant strategies for computing solutions to these nonlinear equations, both of which require a prescribed initial guess as to what the true solution is. The first strategy uses fixed point iteration, wherein one iteration of the solver is defined by sequentially solving each model equation for an update to a specific variable in a fashion similar to Gauss–Seidel or Jacobi methods. Fixed point iteration is a useful tool for analysis and has been analytically shown to yield a convergent solver in some cases [28, 44, 49, 50]; it is noted, however, that convergence may be slower in high voltage scenarios [6].

The second strategy uses Newton's method, which computes the Frechét derivative of the nonlinear elliptic operator to simultaneously linearize all the equations of the PDE and the resulting linear system is solved to compute corrections to the solution of the nonlinear problem. In doing so, the solver aims to compute the solution of the residual equations of the drift-diffusion equations (15)–(16) or the electrokinetic equations (25)–(28). The efficacy of a Newton solver depends on its discretization of the PDE, modifications to improve the well-posedness, and stability of its linearized equations, on its update equations, and on its stopping criteria. Some of these solvers can be found in works like [4, 11, 35, 56], for instance.

While the authors do not know of any proof that guarantees the sequence of Newton iterates is convergent for the presented model equations, work in [42] shows that the solutions to (15)–(16) and (25)–(28) minimize their associated energy functionals, which are lower-semicontinuous, suggesting convergence if the initial guess is sufficiently close to the true solution. A theoretical contribution to the Newton solver approach is presented in section 3.2, where the linearized PNP system is shown to be well-posed, not only for the given choice of variables, but also for the primitive formulation of the problem, which is sometimes used for computing solutions [23, 26, 49, 50].

To avoid cluttered equations, the subscript, $h$, for the finite element functions is suppressed. The basis functions for the finite element space, $W_h$, are denoted by $\{w_j\}$, $j = 1, \ldots, n$. For the drift-diffusion equations, define the residuals

$$\mathcal{R}_{0,j}(\phi, \eta_1, \ldots, \eta_N) \equiv (\epsilon \nabla \phi, \nabla w_j) - \sum_{i=1}^{N} q_i \big(e^{\eta_i}, w_j\big) - (\rho_f, w_j),$$
$$\mathcal{R}_{i,j}(\phi, \eta_1, \ldots, \eta_N) \equiv \big(D_i e^{\eta_i} \nabla(\eta_i + q_i \phi), \nabla w_j\big) + \big(R(\eta_i), w_j\big) - \big(G_i, w_j\big).$$

The purpose of the solver is to compute $\phi, \eta_1, \ldots, \eta_N \in W_h$ such that $\mathcal{R}_{i,j}(\phi, \ldots, \eta_N)$ is zero for $i = 0, \ldots, N$, $j = 1, \ldots, n$. To solve the residual equation, the problem must be linearized; the Frechét derivative of the residuals is

$$\mathcal{J}_{0,j}(\phi, \eta_1, \ldots, \eta_N)[\delta\phi, \delta\eta_1, \ldots, \delta\eta_N] = \frac{\partial}{\partial \varepsilon}\bigg|_{\varepsilon=0} \mathcal{R}_{0,j}(\phi + \varepsilon\delta\phi, \eta_1 + \varepsilon\delta\eta_1, \ldots, \eta_N + \varepsilon\delta\eta_N)$$
$$= (\epsilon \nabla \delta\phi, \nabla w_j) - \sum_{i=1}^{N} q_i \big(e^{\eta_i} \delta\eta_i, w_j\big), \tag{29}$$

$$\mathcal{J}_{i,j}(\phi, \eta_1, \ldots, \eta_N)[\delta\phi, \delta\eta_1, \ldots, \delta\eta_N] = \frac{\partial}{\partial \varepsilon}\bigg|_{\varepsilon=0} \mathcal{R}_{i,j}(\phi + \varepsilon\delta\phi, \eta_1 + \varepsilon\delta\eta_1, \ldots, \eta_N + \varepsilon\delta\eta_N)$$
$$= \big(D_i e^{\eta_i} \big(\nabla \delta\eta_i + \nabla(\eta_i + q_i\phi)\delta\eta_i\big), \nabla w_j\big) + \big(\delta R(\eta_i)\delta\eta_i, w_j\big) + q_i \big(D_i e^{\eta_i} \nabla\delta\phi, \nabla w_j\big), \tag{30}$$

where $\delta R(\eta) = \frac{dR}{d\eta}$ is simply the derivative of the function $R$. It is clear that the Frechét derivative is a linear operator of $\delta\phi$ and $\delta\eta_i$. Given an initial guess, $\phi^{(0)}, \eta_1^{(0)}, \ldots, \eta_N^{(0)}$, that satisfies the Dirichlet boundary conditions, the solution to the residual equations is computed by the sequence of approximate solutions defined by solving for $\delta\phi, \delta\eta_1, \ldots, \delta\eta_N \in W_{h,0}$ such that

$$\mathcal{J}_{i,j}\big(\phi^{(k)}, \eta_1^{(k)}, \ldots, \eta_N^{(k)}\big)[\delta\phi, \delta\eta_1, \ldots, \delta\eta_N] = -\mathcal{R}_{i,j}\big(\phi^{(k)}, \eta_1^{(k)}, \ldots, \eta_N^{(k)}\big) \tag{31}$$

for $i = 0, 1, \ldots, N$, and updating with some damping factor $0 < \mu \leq 1$,

$$\phi^{(k+1)} \leftarrow \phi^{(k)} + \mu\delta\phi^{(k)}, \ldots, \qquad \eta_N^{(k+1)} \leftarrow \eta_N^{(k)} + \mu\delta\eta_N^{(k)}.$$

To simplify notation, the residual and Frechét derivative of the $k$th iterate in the Newton sequence are denoted

$$\mathcal{R}_{i,j}^{(k)} = \mathcal{R}_{i,j}\big(\phi^{(k)}, \eta_1^{(k)}, \ldots, \eta_N^{(k)}\big),$$
$$\mathcal{J}_{i,j}^{(k)}(\delta\phi, \delta\eta_1, \ldots, \delta\eta_N) = \mathcal{J}_{i,j}\big(\phi^{(k)}, \eta_1^{(k)}, \ldots, \eta_N^{(k)}\big)[\delta\phi, \delta\eta_1, \ldots, \delta\eta_N].$$

Proving that the sequence of Newton iterates is well-defined requires the residual, $\mathcal{R}_{i,j}^{(k)}$, to be in the range of the Frechét derivative, $\mathcal{J}_{i,j}^{(k)}$. In general, even if the sequence of Newton iterates is well-defined, it is still difficult to verify its convergence, as the sequence is closely tied to the initial guess.

**3.2. Well-posedness of the linearized PNP system.** In [42], the stability of the nonlinear drift-diffusion and electrokinetic problems is verified by proving energy estimates for the time dependent case. The stability for the steady state solutions is

established in [27, 28]. It should be emphasized, however, that linearization of these systems may destabilize the problem. Namely, the authors have not found a proof that the sequence of Newton iterates satisfies a uniform pointwise bound. As a matter of fact, the stability of the solution cannot be verified for general boundary or initial conditions, as the analytic solution might not generally be in $\mathcal{L}_\infty$, except for finite time.

To verify the well-posedness of the linearized equations, let $\alpha_i > 0$ and $\gamma_i, \theta_i \geq 0$ be bounded functions in $\mathcal{L}_2(\Omega)$, and $\vec{\beta}_i \in [\mathcal{L}_\infty(\Omega)]^d$ for $i = 1, \ldots, N$. The bilinear forms for the linearized PNP system generalize to

$$\begin{aligned}
\mathcal{J}_0(\phi, \ldots, \eta_N; \psi) &\equiv \big(\epsilon\nabla\phi, \nabla\psi\big) - \sum_{i=1}^{N} q_i(\gamma_i\eta_i, \psi), \\
\mathcal{J}_i(\phi, \ldots, \eta_N; \mu_i) &= \mathcal{A}_i\big(\eta_i, \mu_i\big) + q_i\big(\alpha_i\nabla\phi, \nabla\mu_i\big) \\
&\equiv \big(\alpha_i(\nabla\eta_i + \vec{\beta}_i\eta_i), \nabla\mu_i\big) + \big(\theta_i\eta_i, \mu_i\big) + q_i\big(\alpha_i\nabla\phi, \nabla\mu_i\big)
\end{aligned}$$

for $\psi, \mu_i \in W_{h,0}$ and $i = 1, \ldots, N$. The formulation of the PNP system discussed in this article corresponds to

$$\alpha_i = D_i e^{\tilde{\eta}_i}, \quad \vec{\beta}_i = \nabla(\tilde{\eta}_i + q_i\tilde{\phi}), \quad \gamma_i = e^{\tilde{\eta}_i}, \quad \theta_i = e^{\tilde{\eta}_i}/\Delta t_j,$$

or $\theta_i = 0$ for the steady state case, where $\tilde{\eta}_i$ and $\tilde{\phi}$ come from the previous Newton iterate or initial guess. When solving the linearized equations for the primitive variables (with $\rho_i = e^{\eta_i}$), the coefficients are the same except that $\tilde{\eta}_i = 0$ and the $\phi$-term is $(D_i\tilde{\rho}_i\nabla\phi, \nabla\mu_i)$ for a preceding iterate, $\tilde{\rho}_i$. This difference in formulations is inconsequential to the following argument.

The general linearized PNP equation is then to find $\phi, \eta_1, \ldots, \eta_N \in W_{h,0}$ such that

$$\big(\epsilon\nabla\phi, \nabla\psi\big) - \sum_{i=1}^{N} q_i(\gamma_i\eta_i, \psi) = F_0(\psi) \qquad \text{for all } \psi \in W_{h,0}, \tag{32}$$

$$\mathcal{A}_i\big(\eta_i, \mu_i\big) + q_i\big(\alpha_i\nabla\phi, \nabla\mu_i\big) = F_i(\mu_i) \qquad \text{for all } \mu_i \in W_{h,0}, \quad i = 1, \ldots, N, \tag{33}$$

where $F_0$ and $F_i$ are linear functionals defined on $W_{h,0}$. To discuss the criteria for the well-posedness of the system, the Schur complement of the linearized equations is constructed to determine an equivalent linear elliptic operator that only depends on $\phi$. Once this equivalence is established, the well-posedness of the linearized PNP system follows from the well-posedness of the reduced elliptic equation.

Recall that the inf-sup condition is satisfied for the advection-diffusion operator when the mesh is sufficiently refined [53, 59],

$$\sup_{\mu \in W_{h,0}} \frac{\mathcal{A}_i(\eta, \mu)}{\|\mu\|_1} \geq C_i\|\eta\|_1, \tag{34}$$

where $C_i$ is a positive constant that depends on $\alpha_i, \vec{\beta}_i$, and $\theta_i$. This result follows from the well-posedness of the problem on the continuous level, a Gårding-type inequality, and an a priori error estimate for elliptic projection. (Note that the inf-sup condition is always valid for the continuous problem if $\theta_i > 0$, as in the time dependent case.) The inf-sup bound implies that the discrete (reaction-)advection-diffusion operator is nonsingular when the mesh is sufficiently refined.

To write the Schur complement of the linearized PNP system, it is helpful to describe the problem using operator notation. Denote the dual space of $\mathcal{H}^1_{\Gamma_D}$ by $\mathcal{H}^{-1}$ and define $\mathcal{K}_i, \mathcal{L}_i : \mathcal{H}^1_{\Gamma_D} \to \mathcal{H}^{-1}$ such that

$$\begin{aligned}(\mathcal{K}_i\eta_i, \mu) &\equiv \mathcal{A}_i(\eta_i, \mu_i) = \big(\alpha_i(\nabla\eta_i + \vec{\beta}_i\eta_i), \nabla\mu\big) + \big(\theta_i\eta_i, \mu\big) && \text{for all } \mu \in \mathcal{H}^1_{\Gamma_D},\\ (\mathcal{L}_i\phi, \mu) &\equiv \big(\alpha_i\nabla\phi, \nabla\mu\big) && \text{for all } \mu \in \mathcal{H}^1_{\Gamma_D}.\end{aligned}$$

The linearized PNP equations are singular if and only if there exists some nonzero $\phi, \eta_1, \dots, \eta_N \in W_{h,0}$ such that

$$\begin{aligned}\mathcal{J}_0(\phi, \dots, \eta_N; \psi) &= \big(\epsilon\nabla\phi, \nabla\psi\big) - \sum_{i=1}^{N} q_i(\gamma_i\eta_i, \psi) = 0 && \text{for all } \psi \in W_{h,0},\\ \mathcal{J}_i(\phi, \dots, \eta_N; \mu_i) &= (\mathcal{K}_i\eta_i, \mu_i) + q_i(\mathcal{L}_i\phi, \mu_i) = 0 && \text{for all } \mu_i \in W_{h,0}.\end{aligned}$$

From (34), it follows that the convection-diffusion operator, $\mathcal{K}_i$, is invertible so that

$$\eta_i = -q_i\mathcal{K}_i^{-1}\mathcal{L}_i\phi.$$

This implies that the linearized PNP equations (32)–(33) are well-posed if and only if

$$(\epsilon\nabla\phi, \nabla\psi) + \sum_{i=1}^{N} q_i^2(\gamma_i\mathcal{K}_i^{-1}\mathcal{L}_i\phi, \psi) = 0 \qquad \text{for all } \psi \in W_{h,0} \tag{35}$$

has only the trivial solution, $\phi = 0$.

LEMMA 3.1. *Assume that the inf-sup condition* (34) *holds for* $i = 1, \dots, N$. *Then, there exists a positive constant,* $\kappa$, *such that*

$$\sup_{\substack{\psi \in W_{h,0}\\ \psi \neq 0}} \frac{\|\mathcal{K}_i^{-1}\mathcal{L}_i\psi\|_1}{\|\psi\|_1} \leq \kappa \quad \textit{for} \quad i = 1, \dots, N. \tag{36}$$

*Furthermore, there exist positive constants,* $M, C, b$, *such that the reduced elliptic problem is bounded for all* $\phi, \psi \in W_{h,0}$,

$$\begin{aligned}(\epsilon\nabla\phi, \nabla\psi) + \sum_{i=1}^{N} q_i^2(\gamma_i\mathcal{K}_i^{-1}\mathcal{L}_i\phi, \psi) &\leq M\|\phi\|_1\|\psi\|_1,\\ (\epsilon\nabla\phi, \nabla\phi) + \sum_{i=1}^{N} q_i^2(\gamma_i\mathcal{K}_i^{-1}\mathcal{L}_i\phi, \phi) &\geq C\|\phi\|_1^2 - b\|\phi\|_0^2,\end{aligned}$$

*where* $M, C, b$ *are independent of* $\phi, \psi$, *and the mesh.*

*Proof.* It is first proved that (36) holds. The norm on $\mathcal{H}^{-1}$ is defined as

$$\|\mu\|_{-1} \equiv \sup_{\substack{\psi \in \mathcal{H}^1_{\Gamma_D}\\ \psi \neq 0}} \frac{(\mu, \psi)}{\|\psi\|_1}.$$

Furthermore, the coercivity and continuity of $\mathcal{L}_i : \mathcal{H}^1_{\Gamma_D} \to \mathcal{H}^{-1}$ imply that this operator induces an isomorphism between $\mathcal{H}^1_{\Gamma_D}$ and $\mathcal{H}^{-1}$. Thus, it follows from an

application of a Hölder inequality that

$$\sup_{\substack{\psi\in\mathcal{H}^1_{\Gamma_D}\\ \psi\neq 0}} \frac{\|\mathcal{K}_i^{-1}\mathcal{L}_i\psi\|_1}{\|\psi\|_1} = \sup_{\substack{\mu\in\mathcal{H}^{-1}(\Omega)\\ \mu\neq 0}} \frac{\|\mathcal{K}_i^{-1}\mu\|_1}{\|\mathcal{L}_i^{-1}\mu\|_1} \leq \sup_{\substack{\mu\in\mathcal{H}^{-1}(\Omega)\\ \mu\neq 0}} \frac{\|\mathcal{K}_i^{-1}\mu\|_1}{\|\mu\|_{-1}} \sup_{\substack{\nu\in\mathcal{H}^{-1}(\Omega)\\ \nu\neq 0}} \frac{\|\nu\|_{-1}}{\|\mathcal{L}_i^{-1}\nu\|_1}$$
$$= \sup_{\substack{\mu\in\mathcal{H}^{-1}(\Omega)\\ \mu\neq 0}} \frac{\|\mathcal{K}_i^{-1}\mu\|_1}{\|\mu\|_{-1}} \sup_{\substack{\psi\in\mathcal{H}^1_{\Gamma_D}\\ \psi\neq 0}} \frac{\|\mathcal{L}_i\psi\|_{-1}}{\|\psi\|_1} \leq C_i^{-1}C_{\alpha,i},$$

where $C_i$ is the inf-sup constant from (34) and $C_{\alpha,i}$ depends on the coercivity and continuity bounds for $\mathcal{L}_i$. Note that these bounding constants are independent of the mesh size and the bounding constant in (36) is chosen $\kappa \equiv \max_{1\leq i\leq N} C_i^{-1}C_{\alpha,i}$.

From the bound (36), the boundedness of $\epsilon$ and $\gamma_i$, and a Poincaré inequality, the continuity of the reduced elliptic operator is established:

$$(\epsilon\nabla\phi,\nabla\psi) + \sum_{i=1}^N q_i^2(\gamma_i\mathcal{K}_i^{-1}\mathcal{L}_i\phi,\psi) \leq \|\epsilon\|_\infty|\phi|_1|\psi|_1 + \sum_{i=1}^N q_i^2\|\gamma_i\|_\infty\|\mathcal{K}_i^{-1}\mathcal{L}_i\phi\|_1\|\psi\|_{-1}$$
$$\leq \|\epsilon\|_\infty|\phi|_1|\psi|_1 + c_P\kappa\sum_{i=1}^N q_i^2\|\gamma_i\|_\infty\|\phi\|_1\|\psi\|_1 \leq M\|\phi\|_1\|\psi\|_1,$$

with $M = \|\epsilon\|_\infty + c_P\kappa\sum_{i=1}^N q_i^2\|\gamma_i\|_\infty$.

The lower bound similarly follows in conjunction with a lower bound, $\epsilon \geq \bar{\epsilon} > 0$, and Young's inequality,

$$(\epsilon\nabla\phi,\nabla\phi) + \sum_{i=1}^N q_i^2(\gamma_i\mathcal{K}_i^{-1}\mathcal{L}_i\phi,\phi) \geq \bar{\epsilon}|\phi|_1^2 - \kappa\sum_{i=1}^N q_i^2\|\gamma_i\|_\infty\|\phi\|_1\|\phi\|_{-1}$$
$$\geq \bar{\epsilon}|\phi|_1^2 - \tilde{C}\|\phi\|_1\|\phi\|_0 \geq \frac{\bar{\epsilon}}{2}\|\phi\|_1^2 - \left(\frac{\tilde{C}}{2\bar{\epsilon}} + \frac{\bar{\epsilon}}{2}\right)\|\phi\|_0^2 = C\|\phi\|_1^2 - b\|\phi\|_0^2.$$

This concludes the proof. ∎

This result establishes the ellipticity of the Schur complement so that the Fredholm alternative implies that the reduced operator is solvable whenever an eigenvalue of the Laplacian term, $-\nabla\cdot(\epsilon\nabla\phi)$, is not exactly cancelled by the other terms. Note that Lemma 3.1 also applies to the continuous formulation, when $W_{h,0} = \mathcal{H}^1_{\Gamma_D}$.

Following [59], the well-posedness of the discrete reduced elliptic problem is inherited from the continuous level when the mesh is sufficiently refined. This statement is formalized in Theorem 3.2 and gives a sufficient condition for the linearized PNP system to be well-posed. Define the reduced linear elliptic operator, $\hat{\mathcal{L}} : \mathcal{H}^1_{\Gamma_D} \to \mathcal{H}^{-1}$, by

$$\left(\hat{\mathcal{L}}\phi,\psi\right) \equiv (\epsilon\nabla\phi,\nabla\psi) + \sum_{i=1}^N q_i^2(\gamma_i\mathcal{K}_i^{-1}\mathcal{L}_i\phi,\psi) \quad \text{for all } \psi\in\mathcal{H}^1_{\Gamma_D}.$$

THEOREM 3.2. *Assume that the domain, $\Omega$, is convex, the inf-sup condition* (34) *holds for $i = 1,\dots,N$, and the mesh is sufficiently refined. If the continuous linear elliptic operator, $\hat{\mathcal{L}} : \mathcal{H}^1_{\Gamma_D} \to \mathcal{H}^{-1}$, is an isomorphism, then there exists a unique solution to the linearized PNP system* (32)–(33).

*Proof.* This proof exactly follows the argument for the well-posedness of discrete (reaction-)advection-diffusion systems given in [59]. Define the elliptic projection, $P_h : \mathcal{H}^1_{\Gamma_D} \to W_{h,0}$, by

$$\big(\epsilon\nabla(\phi - P_h\phi), \nabla\psi_h\big) = 0 \quad \text{for all } \psi_h \in W_{h,0}.$$

Denote the maximum element diameter $h \equiv \max_{\tau\in\mathcal{T}_h} \operatorname{diam}(\tau)$. From the convexity of the domain, a standard a priori estimate of the elliptic projection holds from a duality argument:

$$\|\psi - P_h\psi\|_0 \le C_{P_h} h \|\psi\|_1. \tag{37}$$

Since $\hat{\mathcal{L}}$ is assumed to be an isomorphism, it holds for all $\phi \in \mathcal{H}^1_{\Gamma_D}$ that

$$\sup_{\substack{\psi\in\mathcal{H}^1_{\Gamma_D}\\ \psi\neq 0}} \frac{(\hat{\mathcal{L}}\phi, \psi)}{\|\psi\|_1} \ge \hat{C}\|\phi\|_1.$$

It follows that

$$\begin{aligned}
(\hat{\mathcal{L}}\phi, P_h\psi) &= (\hat{\mathcal{L}}\phi, \psi) - \sum_{i=1}^N q_i^2(\gamma_i\mathcal{K}_i^{-1}\mathcal{L}_i\phi, \psi - P_h\psi)\\
&\ge (\hat{\mathcal{L}}\phi, \psi) - \kappa\sum_{i=1}^N q_i^2\|\gamma_i\|_\infty\|\phi\|_1\|\psi - P_h\psi\|_0\\
&\ge (\hat{\mathcal{L}}\phi, \psi) - C_{P_h}h\kappa\sum_{i=1}^N q_i^2\|\gamma_i\|_\infty\|\phi\|_1\|\psi\|_1\\
&\ge \left(\hat{C} - C_{P_h}h\kappa\sum_{i=1}^N q_i^2\|\gamma_i\|_\infty\right)\|\phi\|_1\|\psi\|_1.
\end{aligned}$$

Since $\|\psi\|_1 \ge c\|P_h\psi\|_1$ for some uniform constant, the inf-sup condition is verified for the discrete reduced elliptic operator when $h < \hat{C}/C_{P_h}\kappa\sum_{i=1}^N q_i^2\|\gamma_i\|_\infty$. The well-posedness of the linearized PNP system follows from its equivalence with the discrete reduced elliptic operator. ∎

**3.3. EAFE stabilization.** Many approaches have been proposed to improve instabilities arising from a large convection term in (30). A change of variables to the Slotboom variables [9, 55, 56] symmetrizes the drift-diffusion equations by absorbing the drift term into a nonlinear diffusion coefficient; in [35, 51] however, it is argued that this change of variables may suffer from instabilities of a different sort due to potentially large variations in the resulting diffusion coefficient. Discussions on alternative formulations can also be found in [5, 6, 28].

The Scharfetter–Gummel scheme [52] and extensions thereof constitute a large portion of the literature for stabilizing drift-diffusion equations; the stabilization scheme has variations that generalize the original work to higher dimensional finite difference and finite element schemes, for example [5, 6, 44, 54]. Finite elements may employ upwinding schemes [3, 10], mixed elements [8, 43], or edge-based schemes, as in [40]. This work employs the edge-averaged finite element (EAFE) scheme proposed in [60] because it satisfies an a priori error estimate, is general enough to accommodate up to three spatial dimensions, and places fewer constraints on the diffusion

and convection coefficients. (This approximation is shown to be equivalent to the finite volume Scharfetter–Gummel scheme of [5, 6].) One drawback of using EAFE to approximate the linearized system is that the solved linear equations do not exactly correspond to the Newton linearization as proposed above. The nonlinear solver is therefore a more stable quasi-Newton method, but potentially slower to converge.

The motivation of the EAFE scheme is to approximate an advective-diffusive flux by a piecewise constant flux vector on each element, in such a way that the modified flux inherits the monotonicity of the discrete Laplacian. Let $\tau \in \mathcal{T}_h$ and $E \subset \tau$ denote an element in the mesh and an edge of the element, respectively, where $E$ is oriented from vertex $x_i$ to $x_j$. Let $\theta_E^\tau$ be the interior angle of the element opposite to the edge $E$ in $\tau$, let $\kappa_E^\tau \subset \tau$ be the $d-2$ dimensional intersection of the two facets opposite to $E$ (in $d = 2$ dimensions, $\kappa_E^\tau$ is a vertex and $|\kappa_E^\tau| = 1$ by convention), and let $\delta_E \psi = \psi(x_i) - \psi(x_j)$ be the difference in value of a continuous function $\psi \in W$ at the endpoints of the edge $E$. Then, the weak form of the discrete Laplacian is equal to

$$(\nabla\psi, \nabla w) = \sum_{\tau\in\mathcal{T}_h} \sum_{E\subset\tau} \omega_E^\tau \delta_E\psi \delta_E w$$

for $\psi, w \in W_h$, where $\omega_E^\tau \equiv |\kappa_E^\tau| \cot\theta_E^\tau / d(d-1)$. (This identity comes directly from [60].) The harmonic average of $D_i e^{-\phi}$ on edge, $E$, is given by

$$\tilde{D}_{E,i}(\phi) \equiv \frac{|E|}{\int_E D_i^{-1} e^{\phi}\, ds},$$

which is used to define the EAFE flux

$$\begin{aligned}\tilde{a}_h(\psi, w) &\equiv \sum_{\tau\in\mathcal{T}_h} \sum_{E\subset\tau} \omega_E^\tau \tilde{D}_{E,i}(q_i\phi)\delta_E(e^{q_i\phi}\psi)\delta_E w \\ &\approx \big(D_i e^{\eta_i}\big(\nabla\psi + \nabla(\eta_i + q_i\phi)\psi\big), \nabla w\big).\end{aligned}$$

The presentation given above is specific to the given form of the convection-diffusion terms in the charge carrier residual equations and is not general; for a full discussion on the EAFE stabilization scheme, refer to [60]. For the nonlinear source term, mass-lumping is required to preserve monotonicity, which is defined in terms of the nodal interpolant $\mathcal{I}_h : W \to W_h$,

$$(\psi, w)_h \equiv \int_\Omega \mathcal{I}_h(\psi \cdot w)\, dx.$$

Then, the EAFE approximation to the linear operator is given by

$$\mathcal{J}_{i,j}^{(k)} \approx \tilde{\mathcal{J}}_{i,j}^{(k)} = \tilde{a}_h^{(k)}(\delta\eta_i, w_j) + \big(\delta R(\eta_i^{(k)})\delta\eta_i, w_j\big)_h + q_i\big(D_i e^{\eta_i^{(k)}} \nabla\delta\phi, \nabla w_j\big).$$

**3.4. The electrokinetic solver.** The solver for the electrokinetic system is an extension of the drift-diffusion solver, where the residual equations for the charge carriers contain an additional term,

$$\begin{aligned}\mathcal{R}_{i,j}(\phi, \ldots, \eta_N, \vec{u}, p) &= \big(D_i e^{\eta_i}\nabla(\eta_i + q_i\phi) - e^{\eta_i}\vec{u}, \nabla w_j\big) + \big(R(\eta_i), w_j\big) - \big(G_i, w_j\big), \\ \mathcal{J}_{i,j}(\delta\phi, \ldots, \delta\eta_N, \delta\vec{u}, \delta p) &= \big(D_i e^{\eta_i}\big(\nabla\delta\eta_i + (\nabla\eta_i + q_i\nabla\phi - D_i^{-1}\vec{u})\delta\eta_i\big), \nabla w_j\big) \\ &\quad + \big(\delta R(\eta_i)\delta\eta_i, w_j\big) + q_i\big(D_i e^{\eta_i}\nabla\delta\phi, \nabla w_j\big) - \big(e^{\eta_i}\delta\vec{u}, \nabla w_j\big),\end{aligned}$$

for $i = 1, \ldots, N$, and additional residual equations must be solved for the NS subsystem:

$$\begin{aligned}\mathcal{R}_{N+1,j}(\phi, \eta_1, \ldots, \eta_N, \vec{u}, p) &= D_{h,t}\big(\vec{u}; \vec{u}, \vec{v}\big) + A_h\big(\vec{u}, \vec{v}\big) - \big(p, \nabla \cdot \vec{v}\big) \\ &\quad - \sum_{i=1}^{N} q_i\big(e^{\eta_i}\nabla\phi, \vec{v}\big) - \big(G_u, \vec{v}\big), \\ \mathcal{R}_{N+2,j}(\phi, \eta_1, \ldots, \eta_N, \vec{u}, p) &= \big(\nabla \cdot \vec{u}, q\big),\end{aligned}$$

with their corresponding linearizations,

$$\begin{aligned}\mathcal{J}^{(k)}_{N+1,j}(\delta\phi, \delta\eta_1, \ldots, \delta\eta_N, \delta\vec{u}, \delta p) &= \Big[D_{h,t}\big(\vec{u}^{(k)}; \delta\vec{u}, \vec{v}\big) + \rho_f\big(\delta\vec{u} \cdot \nabla\vec{u}^{(k)}, \vec{v}\big)\Big] + A_h\big(\delta\vec{u}, \vec{v}\big) \\ &\quad - \big(\delta p, \nabla \cdot \vec{v}\big) - \sum_{i=1}^{N} q_i\big(e^{\eta_i^{(k)}}\nabla\delta\phi + e^{\eta_i^{(k)}}\nabla\phi^{(k)}\delta\eta_i, \vec{v}\big), \\ \mathcal{J}_{N+2,j}(\delta\phi, \delta\eta_1, \ldots, \delta\eta_N, \delta\vec{u}, \delta p) &= \big(\nabla \cdot \delta\vec{u}, q\big).\end{aligned}$$

Note that if the Stokes equations are being solved instead of the NS system, the term corresponding to the kinematic derivative in the linearized operator is dropped.

The nonlinear solver follows exactly the same form as the drift-diffusion solver, except that the EAFE approximation must now account for the additional convection from the velocity field $\vec{u}^{(k)}$:

$$\tilde{a}_h^{(k)}(\psi, w) = \sum_{\tau \in \mathcal{T}_h} \sum_{E \subset \tau} \omega_E^{\tau} \tilde{D}_{E,i}\big(q_i\phi^{(k)} + D_i^{-1}\vec{u}^{(k)}\big)\delta_E(e^{q_i\phi^{(k)} + D_i^{-1}\vec{u}^{(k)}}\psi)\delta_E w.$$

**3.5. The linear solver.** After discretization, the following linear system at each Newton iteration is given:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Longleftrightarrow \begin{pmatrix} \boldsymbol{A}_{\mathrm{PNP}} & \boldsymbol{B}_1 \\ \boldsymbol{B}_2 & \boldsymbol{A}_{\mathrm{NS}} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{\mathrm{PNP}} \\ \boldsymbol{x}_{\mathrm{NS}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_{\mathrm{PNP}} \\ \boldsymbol{b}_{\mathrm{NS}} \end{pmatrix}. \tag{38}$$

Due to the large size of the linearized problem, iterative methods are needed. In order to accelerate the convergence rate, block preconditioners are used to improve the condition number of the linearized problem. Due to the two-by-two block structure of (38), the following three different block preconditioners are considered:

$$\boldsymbol{P}_D = \begin{pmatrix} \boldsymbol{A}_{\mathrm{PNP}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_{\mathrm{NS}} \end{pmatrix}, \quad \boldsymbol{P}_L = \begin{pmatrix} \boldsymbol{A}_{\mathrm{PNP}} & \boldsymbol{0} \\ \boldsymbol{B}_2 & \boldsymbol{A}_{\mathrm{NS}} \end{pmatrix}, \quad \text{and} \quad \boldsymbol{P}_U = \begin{pmatrix} \boldsymbol{A}_{\mathrm{PNP}} & \boldsymbol{B}_1 \\ \boldsymbol{0} & \boldsymbol{A}_{\mathrm{NS}} \end{pmatrix}.$$

Note that for these three block preconditioners, the PNP subsystem, $\boldsymbol{A}_{\mathrm{PNP}}$, and NS subsystem, $\boldsymbol{A}_{\mathrm{NS}}$, are decoupled and can be solved separately. In the experiments presented in section 4, the block upper triangular preconditioner, $\boldsymbol{P}_U$, is used because it gives better performance when compared to the block diagonal preconditioner or the block lower triangular preconditioner. Schur complement block solvers are commonly used to solve linear systems arising from multiphysics problems [7, 16, 25], but were not investigated here because reasonable performance is achieved using a simpler block upper-triangular solver.

In practice, directly applying $\boldsymbol{P}_U$ requires inverting $\boldsymbol{A}_{\mathrm{PNP}}$ and $\boldsymbol{A}_{\mathrm{NS}}$ exactly, which is computationally prohibitive, especially for large-scale simulations. Therefore, these

two diagonal blocks are approximately solved, which results in the following block upper triangular preconditioner:

$$\boldsymbol{M}_U = \begin{pmatrix} \boldsymbol{M}_{\mathrm{PNP}}^{-1} & \boldsymbol{B}_1 \\ \boldsymbol{0} & \boldsymbol{M}_{\mathrm{NS}}^{-1} \end{pmatrix},$$

where

$$\boldsymbol{M}_{\mathrm{PNP}} \approx \boldsymbol{A}_{\mathrm{PNP}}^{-1} \quad \text{and} \quad \boldsymbol{M}_{\mathrm{NS}} \approx \boldsymbol{A}_{\mathrm{NS}}^{-1}.$$

This means $\boldsymbol{M}_{\mathrm{PNP}}^{-1}$ and $\boldsymbol{M}_{\mathrm{NS}}^{-1}$ should be good approximations of $\boldsymbol{A}_{\mathrm{PNP}}$ and $\boldsymbol{A}_{\mathrm{NS}}$, respectively. Algorithmic descriptions of $\boldsymbol{M}_{\mathrm{PNP}}$ and $\boldsymbol{M}_{\mathrm{NS}}$ are given below.

For the $\boldsymbol{A}_{\mathrm{PNP}}$ block, the degrees of freedom for $\phi, \eta_1, \ldots, \eta_N \in W_{h,0}$ can be ordered to follow their spatial distribution in the domain. Use this ordering for $\boldsymbol{A}_{\mathrm{PNP}}$ to keep all the degrees of freedom corresponding to a particular vertex together. This gives the following block structure:

$$\boldsymbol{A}_{\mathrm{PNP}} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1N_W} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2N_W} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{N_W 1} & \boldsymbol{A}_{N_W 2} & \cdots & \boldsymbol{A}_{N_W N_W} \end{pmatrix},$$

where $\boldsymbol{A}_{ij} \in \mathbb{R}^{(N+1)\times(N+1)}$ contains the inner-product of all basis functions corresponding to vertex $i$ with those at vertex $j$, for $N_W$ vertices. This approach is similar to that of [4]; whereas that work used a diagonal preconditioner, a block version of the unsmoothed aggregation AMG (UA-AMG) method [31, 61] is used here, which performs well for convection-diffusion problems. A block ILU smoother is used on the finest level and a block Gauss–Seidel smoother is used on all other levels. A V-cycle with one step of presmoothing and one step of postsmoothing is applied. Finally, $\boldsymbol{A}_{\mathrm{PNP}}$ is solved by the UA-AMG preconditioned GMRes method, which gives $\boldsymbol{M}_{\mathrm{PNP}}$.

Many existing block preconditioners can be used for the NS subsystem; some approaches use geometric multigrid [45, 47], an algebraic approach based on Schur complements [16, 48], or analytic approximations of differential operators [1, 36, 37]. This work follows the idea proposed in [20]; a least square commutator-type preconditioner is applied. More precisely, the least square distributive Gauss–Seidel (LSDGS) approach proposed in [13, 57] is used as a block preconditioner. This approach works well for small Reynolds numbers, as are typical in electrokinetics problems, and divergence-free discretizations [15]. For the LSDGS preconditioner, a velocity block and a pressure block need to be solved, where a standard UA-AMG method is used for both sub-blocks. An ILU smoother is used on the finest level in order to improve performance. Overall, the LSDGS-preconditioned GMRes method is used to solve $\boldsymbol{A}_{\mathrm{NS}}$ and define $\boldsymbol{M}_{\mathrm{NS}}$.

Finally, the overall iterative solver is the flexible general minimal residual (FGMRes) method with $\boldsymbol{M}_U$ as the preconditioner. The flexible version of the GMRes method is used because the diagonal blocks of $\boldsymbol{M}_U$, $\boldsymbol{M}_{\mathrm{PNP}}$, and $\boldsymbol{M}_{\mathrm{NS}}$ are both defined by preconditioned GMRes. This implies that the overall block preconditioner $\boldsymbol{M}_U$ is actually changing from one step to another; therefore, FGMRes has to be used.

**4. Numerical experiments.** A solver has been implemented in C++ using the FEniCS software package [34] for discretization and assembling the linear equations corresponding to linearized finite element equations. The linear equations are solved

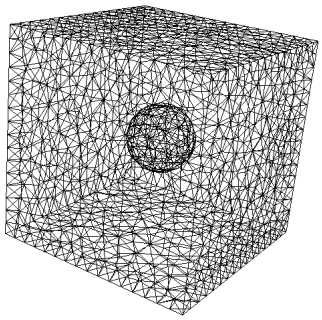

FIG. 1. *A mesh of* 2,805 *vertices, used in computing the PNP convergence experiments.*

using the FASP solver package [12] as described above. The release of the software is pending a public release of the FASP solver package.

All experiments solve the steady state equations since computing the solutions at steady state requires a more robust solver. The experiments here aim to accomplish two goals. The first set of experiments verifies that the PNP solver converges to a known solution, and the second set of experiments demonstrates the convergence of the proposed Newton solver for the electrokinetic equations, where coefficients of the equations are chosen to emulate a biological system.

**4.1. Convergence of the PNP solver.** In this experiment, the PNP equations are solved where a fixed charge is added to the Poisson equation so that the electrostatic potential is equal to the solution of a radially symmetric Poisson–Boltzmann equation.

The domain for this problem is given by a cube containing a spherical cavity: let $L = 0.3$ and $r_0 = 0.1$;

$$\Omega = \{x \in \mathbb{R}^3 \mid \|x\|_\infty \leq L \text{ and } \|x\|_2 \geq r_0 \}.$$

One of the meshes used for this experiment is depicted in Figure 1.

The spherical cavity is intended to model a charged colloid; so, the fixed charge in the Poisson equation is selected so that the electrostatic potential is given by the solution to the radially symmetric Poisson–Boltzmann equation.

The one-dimensional Poisson–Boltzmann equation can be written as

$$-\epsilon\phi''_{\mathrm{pb}}(r) = -2\sinh(\phi_{\mathrm{pb}}(r)) = e^{-\phi_{\mathrm{pb}}(r)} - e^{\phi_{\mathrm{pb}}(r)}, \qquad r \geq r_0, \tag{39}$$

with Dirichlet boundary condition

$$\phi_{\mathrm{pb}}(r_0) = -1.$$

It is straightforward to verify that the solution to this differential equation is given by

$$\phi_{\mathrm{pb}}(r) = 2\ln\left(\frac{1-\gamma e^{-\kappa r}}{1+\gamma e^{-\kappa r}}\right), \qquad \text{where} \quad \gamma = e^{\kappa r_0}\frac{e^{1/2}-1}{e^{1/2}+1} \quad \text{and} \quad \kappa = \sqrt{2/\epsilon}.$$

In this experiment, the system is modeled by a single cation and anion so that $q_1 = 1$ and $q_2 = -1$. The analytic solution for this experiment is given by

$$\phi(x) = \phi_{\mathrm{pb}}(\|x\|_2) \qquad \text{and} \qquad \eta_i(x) = -q_i\phi_{\mathrm{pb}}(\|x\|_2), \quad i = 1, 2. \tag{40}$$

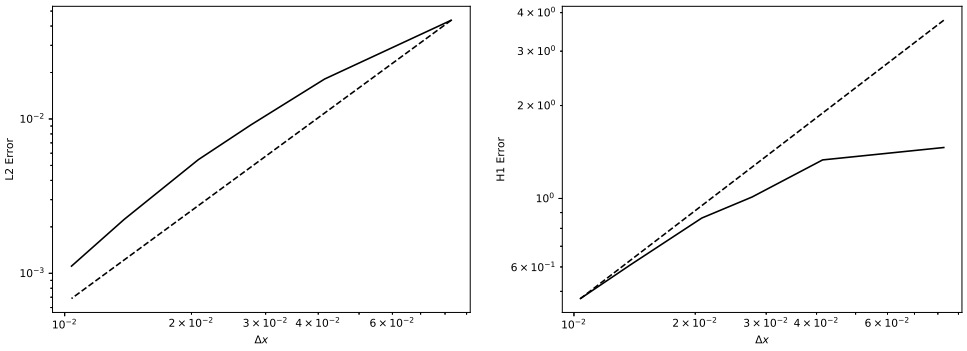


FIG. 2. *On the left is the $L_2$-error of the computed solution plotted against the maximal element diameter. A dashed line of slope two is also plotted as a reference for second order convergence. The figure on the right shows the $H^1$-error plotted against the maximal element diameter, along with a dashed line of unit slope to show asymptotically linear convergence. Poor resolution of the boundary layer near the colloid prevents asymptotic rates until the mesh is sufficiently refined.*

It is clear that this choice of the solution readily satisfies the drift-diffusion equations (6) with time derivatives set to zero. For the Poisson equation, a fixed charge is added to the equation:

$$-\nabla \cdot (\epsilon \nabla \phi) - e^{\eta_1} + e^{\eta_2} = \rho_{\text{fixed}},$$

with

$$\rho_{\text{fixed}}(x) \equiv -\epsilon \Delta \phi_{\text{pb}}(\|x\|_2) - e^{-\phi_{\text{pb}}(\|x\|_2)} + e^{\phi_{\text{pb}}(\|x\|_2)},$$

so that the solution to this PNP system is given by (40).

For this system, Dirichlet boundary conditions are imposed on opposite sides of the domain, for $x_1 = \pm L$, to match the analytic solution, and the value of the permittivity is set to $\epsilon = 10^{-4}$ to ensure a steep boundary layer near the spherical cavity. (Since the permittivity is so small, the Dirichlet boundary conditions are nearly constant $\phi = -\eta_1 = \eta_2 \approx 0$.)

To establish the rate of convergence, the proposed Newton solver is applied to a set of successively refined meshes and the error of the computed solution is measured on each mesh. To keep results consistent, the Newton solver is run for ten iterations, regardless of the size of the computed residual. Typically, the relative residual does not change after the second or third Newton iterate for this problem as the error of the initial guess for the iterative linear solver is below its error tolerance. The results of the experiments are displayed in Figure 2.

While uniform mesh refinement provides valuable information regarding the rate of convergence for the solver, it is not a practical approach to computing solutions when narrow boundary layers are present. The last part of this experiment uses adaptive refinement to refine the mesh more intelligently than in the experiment above, where the refinement strategy simply refines elements with the largest computed entropy:

$$\sum_{i=1,2} \int_{\Omega} D_i e^{\eta_i} |\nabla(\eta_i + q_i \phi)|^2 \, dx > 0.$$

Upon convergence of the Newton solver, the elements with the largest computed entropy are marked for refinement. The count of marked elements is chosen so that

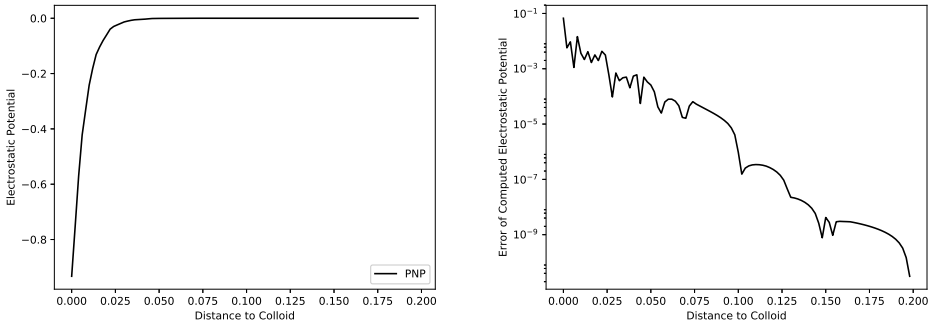


FIG. 3. *The figure on the left plots the electrostatic potential of the computed solution (PNP) and the exact solution (PB), as given in* (40). *To generate this figure, the computed and analytic solutions are sampled along a line that extends out from the spherical cavity radially and the dependent variable in this plot is the distance to the colloid. Only a small portion of the domain near the spherical cavity is shown to illustrate the boundary layer, as the solutions are nearly constant and in agreement further from the colloid. The $L_2$-error of the computed solution along this line is plotted on the right.*

the mesh resulting from the refinement has nearly twice as many elements as its parent mesh.

This marking strategy is motivated by the stability estimate in [42], where it is shown that the exact finite element solution should have a computed entropy of zero. Using this simple adaptive strategy, the solver is able to target elements near the spherical cavity for refinement. The solution computed using this adaptive approach is compared to the analytical solution in Figure 3, where it is shown that very small elements are needed to resolve the boundary layer near the cavity, while much larger elements can be used elsewhere.

**4.2. Electrophoresis experiment.** This experiment emulates a biological system with a charged colloid suspended in an electrolyte of a single cation and anion. The modeled behavior is that an electrical field is applied across the domain and, consequently, the charged colloid moves along the electric field. The relationship between the strength of this applied electric field and the resulting speed of the charged colloid is computed and verified to qualitatively match physical experiments.

These simulations again employ a cube-shaped domain with a spherical cavity in its center. The cube is $2L = 20$nm long in each direction and the cavity in the center of the domain has a radius of $r_0 = 4$nm. Using more compact notation,

$$\Omega = \{x \in \mathbb{R}^3 \mid \|x\|_\infty \leq L \text{ and } \|x\|_2 \geq r_0\}.$$

The cavity again models a charged colloid, where the surface charge of this colloid is modeled by a nonhomogeneous Neumann boundary condition on the electrostatic potential and no-flux boundary conditions for the log-charge densities and fluid velocity: on $\Gamma_N = \{x \in \Omega \mid \|x\|_2 = r_0\}$,

$$\epsilon \nabla\phi \cdot \vec{n} = S, \qquad \nabla(\eta_i + q_i\phi) \cdot \vec{n} = 0, \qquad \text{and} \qquad \vec{u} \cdot \vec{n} = 0$$

for $i = 1, 2$ with $q_1 = +1$ and $q_2 = -1$, where $S$ denotes the surface charge of the colloid. The electric field is applied along the first coordinate dimension and

is modeled as Dirichlet boundary conditions on opposite sides of the domain: on $\Gamma_D = \{x \in \Omega \mid x_1 = \pm L\}$,

$$\phi(x) = \pm\frac{1}{2}\delta V \quad \text{and} \quad \eta_i(x) = \log(\rho_{\text{bath}}) \qquad \text{for } x_1 = \pm L,$$

where $\delta V$ is the total voltage drop of the electrostatic potential and the positive value $\rho_{\text{bath}}$ is the bath (bulk) concentration of the charge carriers in the region surrounding the domain. On $\Gamma_D$, natural boundary conditions are imposed for the fluid velocity that correspond to no surface tension,

$$(2\mu\varepsilon(\vec{u}) - pI)\,\vec{n} = 0,$$

as this allows for the fluid to freely flow in and out of the domain along this direction. For the remaining boundary surfaces of the domain, natural boundary conditions are used for all variables.

The coefficients of the differential equations are chosen to match biological conditions:

$$\rho_{\text{bath}} = 10 \text{ millimolar (mM)} \qquad \text{and} \qquad S = 10\ e_C/\text{nm}^2,$$

where $e_C \approx 1.6 \times 10^{-19} C$ is the elementary charge constant, and the voltage drop, measured in millivolts (mV), is sampled along a discrete set of values:

$$\delta V \in \{0\text{mV}, 50\text{mV}, 100\text{mV}, 150\text{mV}, 200\text{mV}, 250\text{mV}\}.$$

The remaining coefficients are chosen such that $\vec{u}$ models the velocity of water near room-temperature, $300K$, and the charge carriers represent the sodium cation and chloride anion. For brevity, only physical units are reported along with the resulting values after dimensional analysis. The units are summarized:

$$\begin{aligned}
q_1 &= (1e_C)/q_{\text{ref}} = 1,\\
q_2 &= (-1e_C)/q_{\text{ref}} = -1,\\
D_1 &= (1.334 \times 10^{-5}\ \text{cm}^2/\text{s})/D_{\text{ref}} = 1,\\
D_2 &= (2.032 \times 10^{-5}\ \text{cm}^2/\text{s})/D_{\text{ref}} = 1.523,\\
\epsilon &= (80\epsilon_0\ \text{C V}^{-1}\text{m}^{-1})/\epsilon_{\text{ref}} = 0.021,\\
\rho_f &= (9.966 \times 10^2\ \text{kg m}^{-3})/\rho_{f,\text{ref}} = 0.0036,\\
\mu &= (8.6 \times 10^{-4}\ \text{kg m}^{-1}\text{s}^{-1})/\mu_{\text{ref}} = 1.034.
\end{aligned}$$

The values for $\epsilon_{\text{ref}}$, $\rho_{f,\text{ref}}$, and $\mu_{\text{ref}}$ are not free parameters, but are determined by the scaling of the charge concentration, the electrostatic potential, and the fluidic velocity. (For more details regarding how to carry out the dimensional analysis, the reader is referred to [19].)

Choosing the length scale so that $L = 1$, the authors are careful to note that the size of the permittivity, $\epsilon = 0.021$, is somewhat larger than in the previous experiment, which implies the effects of the charged colloid propagate further through the domain. To ensure that the Dirichlet boundary conditions do not compromise the qualitative behavior of the computed solution that comes from the charged colloid, PNP simulations without the fluid are run at each voltage drop on a larger domain, where $|x_1| \leq 10L$. Figure 4 depicts computed solutions from some of these simulations and demonstrates that the qualitative effects of the charged colloid on the computed solution are well-insulated from the boundary.

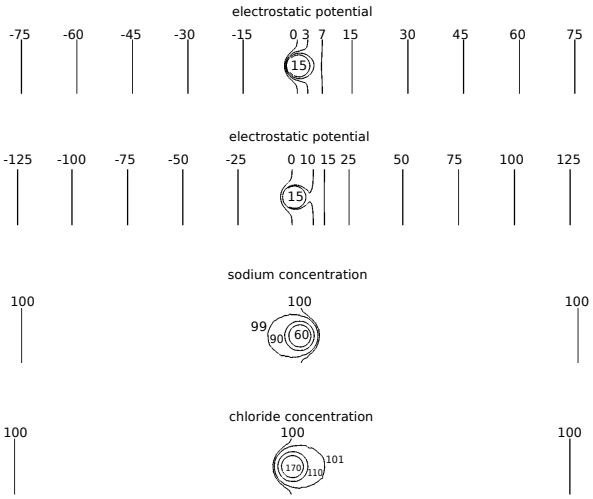


FIG. 4. *Plots above display contour lines of a cross-section of the computed solution to the PNP system on the extended domain. Displayed from top to bottom are the electrostatic potential at* $\delta V = 150\,\mathrm{mV}$*, the electrostatic potential at* $\delta V = 250\,\mathrm{mV}$*, and the sodium and chloride concentrations for* $\delta V = 250\,\mathrm{mV}$*. The top figures show that the voltage is linear throughout the cross-section, aside from very near the charged colloid. The computed concentrations are constant except for near the surface of the colloid.*

For each voltage drop, the solution for the electrokinetic system is solved using the proposed Newton solver and convergence is achieved for each simulation in about three iterations, as shown in Figure 5. A cross-section of the computed fluidic velocity at 250mV is depicted in Figure 6, which shows how the applied electric field causes fluid flow around the colloid, as expected.

The relationship between the average fluid outflow rate on the boundary $x_1 = 1$ and the voltage drop is linear, as illustrated in Figure 7, and typical in electrophoretic systems [32]. The coefficient in this linear relationship is the *electrophoretic mobility*, which characterizes the induced speed of the charged colloid per unit of electric field applied. The electrophoretic mobility is computed by averaging the fluidic velocity at the outflow boundary and dividing by the strength of the electric field; the computed value is

$$\mu_e = \frac{|\vec{u}_{\mathrm{outflow}}|}{|\nabla\phi|} = 5.31 \times 10^{-8}\mathrm{m}^2/\mathrm{V\,s},$$

which is in a reasonable range when compared to other experiments [46].

**5. Summary and closing remarks.** The authors proposed a general framework for solving two standard PNP models, the drift-diffusion and the electrokinetic models, based on a discretization using a logarithmic transformation on the charge carrier densities and Stokes-stable and div-conforming finite elements for the fluidic velocity and pressure. The framework is based on a monolithic Newton method, applied to the entire (multiphysics) system, to resolve the nonlinear couplings between the variables. Theorem 3.2 shows that the linearized PNP subsystem is well-posed when the mesh is sufficiently refined, and the proof is general enough to cover alternate formulations of the PDE as well, such as the commonly used primitive formulation.

Furthermore, fast linear solvers are designed to approximately solve the linearized PNP system. The approach of this PNP solver aims to avoid unnecessary numeri-

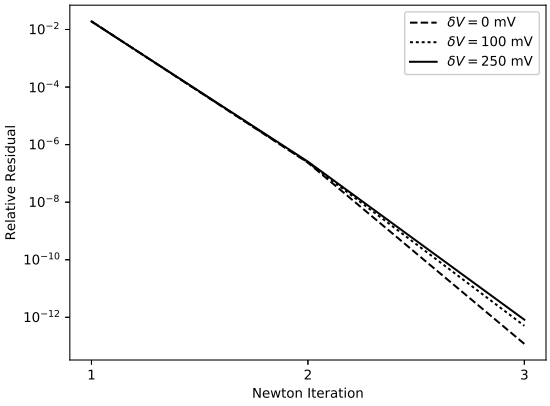


FIG. 5. *This graph of the relative residual of the Newton solver for the electrokinetic system demonstrates the convergence to a satisfactory solution for some of the tested voltage drops. Graphs for other tested voltages are nearly indistinguishable from those shown here, indicating that convergence is insensitive to the strength of the tested voltages.*

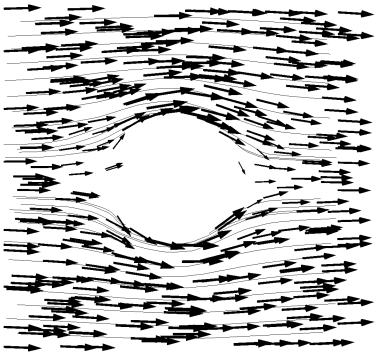

FIG. 6. *Streamlines and sampling of the fluidic velocity field on a cross section, with* 250 mV *applied over* 200 nm. *The fluid flows around the colloid due to the applied electric field, reaching speeds up to roughly* 82 mm/s. *Lower applied voltage leads to slower fluid flow.*

cal approximations to the nonlinear PDE, except for the stabilizing effects of the EAFE scheme. The PNP solver is tested for convergence and efficiency, in addition to its ability to solve physically motivated problems. The numerical examples presented above focused on simulations of a charged particle; however, there are many more applications that would benefit from an efficient and robust solver for charge transport; further applications can be found in the fields studying semiconductors, electrochemistry, or electrophysiology.

Future work includes more advanced applications of the proposed solver, as well as error analysis of the proposed finite element scheme. Such analysis has remained a challenge, due to the nonlinear and strongly coupled nature of the PDE. The authors are also considering a reformulation of the stress tensor in the Navier–Stokes equations to make the linear system sparser.

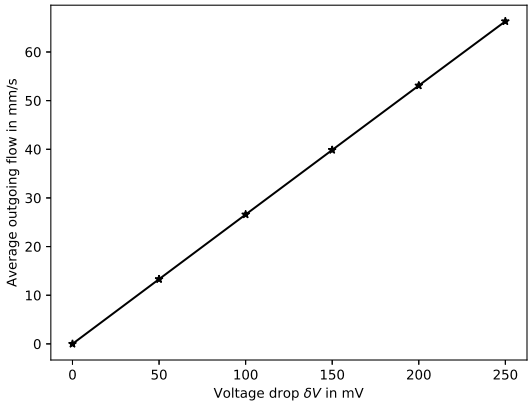


FIG. 7. *This plot shows that the computed relationship of the applied electric field strength and the average speed of the computed fluid is linear with a proportionality constant of* $5.31\times10^{-8}\mathrm{m}^2/\mathrm{V\,s}$.

## REFERENCES


[1] B. Ayuso de Dios, F. Brezzi, L. D. Marini, J. Xu, and L. Zikatanov, *A simple preconditioner for a discontinuous Galerkin method for the Stokes problem*, J. Sci. Comput., 58 (2014), pp. 517–547.

[2] H. P. Baltes and R. S. Popovic, *Integrated semiconductor magnetic field sensors*, Proc. IEEE, 74 (1986), pp. 1107–1132.

[3] R. E. Bank, J. F. Bürgler, W. Fichtner, and R. K. Smith, *Some upwinding techniques for finite element approximations of convection-diffusion equations*, Numer. Math., 58 (1990), pp. 185–202.

[4] R. E. Bank, T. F. Chan, W. M. Coughran, Jr., and R. K. Smith, *The alternate-block-factorization procedure for systems of partial differential equations*, BIT, 29 (1989), pp. 938–954.

[5] R. E. Bank, W. Coughran, Jr., and L. C. Cowsar, *The finite volume Scharfetter-Gummel method for steady convection diffusion equations*, Comput. Vis. Sci., 1 (1998), pp. 123–136.

[6] R. E. Bank, D. J. Rose, and W. Fichtner, *Numerical methods for semiconductor device simulation*, IEEE Trans. Electron Devices, 30 (1983), pp. 1031–1041.

[7] M. Benzi and G. H. Golub, *A preconditioner for generalized saddle point problems*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 20–41, https://doi.org/10.1137/S0895479802417106.

[8] F. Brezzi, L. D. Marini, and P. Pietra, *Numerical simulation of semiconductor devices*, Comput. Methods Appl. Mech. Engrg., 75 (1989), pp. 493–514.

[9] F. Brezzi, L. D. Marini, and P. Pietra, *Two-dimensional exponential fitting and applications to drift-diffusion models*, SIAM J. Numer. Anal., 26 (1989), pp. 1342–1355, https://doi.org/10.1137/0726078.

[10] A. N. Brooks and T. J. Hughes, *Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations*, Comput. Methods Appl. Mech. Engrg., 32 (1982), pp. 199–259, https://doi.org/10.1016/0045-7825(82)90071-8.

[11] J. F. Burgler, R. E. Bank, W. Fichtner, and R. K. Smith, *A new discretization scheme for the semiconductor current continuity equations*, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst., 8 (1989), pp. 479–489.

[12] Center for Computational Mathematics and Applications, *FASP: Fast Auxiliary Space Preconditioning*, Mathematics Department, Pennsylvania State University, fasp.sourceforge.net/, 2015. Version 1.6.8.

[13] L. Chen, X. Hu, M. Wang, and J. Xu, *A multigrid solver based on distributive smoother and residual overweighting for Oseen problems*, Numer. Math. Theory Methods Appl., 8 (2015), pp. 237–252.

[14] B. Cockburn, G. Kanschat, and D. Schötzau, *A locally conservative LDG method for the incompressible Navier-Stokes equations*, Math. Comp., 74 (2005), pp. 1067–1095.

[15] F. Colombini, V. Petkov, and J. Rauch, *Spectral problems for non-elliptic symmetric systems with dissipative boundary conditions*, J. Funct. Anal., 267 (2014), pp. 1637–1661.
[16] E. C. Cyr, J. N. Shadid, R. S. Tuminaro, R. P. Pawlowski, and L. Chacón, *A new approximate block factorization preconditioner for two-dimensional incompressible (reduced) resistive MHD*, SIAM J. Sci. Comput., 35 (2013), pp. B701–B730, https://doi.org/10.1137/12088879X.
[17] S. Datta, *Electronic Transport in Mesoscopic Systems*, Cambridge University Press, Cambridge, UK, 1997.
[18] A. De Mari, *An accurate numerical steady-state one-dimensional solution of the pn junction*, Solid-State Electronics, 11 (1968), pp. 33–58.
[19] B. Eisenberg, Y. Hyon, and C. Liu, *Energy variational analysis of ions in water and channels: Field theory for primitive models of complex ionic fluids*, J. Chem. Phys., 133 (2010), 104104.
[20] H. Elman, D. Silvester, and A. Wathen, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, 2005.
[21] D. Erickson and D. Li, *Influence of surface heterogeneity on electrokinetically driven microfluidic mixing*, Langmuir, 18 (2002), pp. 1883–1892.
[22] H. K. Gummel, *A self-consistent iterative scheme for one-dimensional steady state transistor calculations*, IEEE Trans. Electron Devices, 11 (1964), pp. 455–465.
[23] T.-L. Horng, T.-C. Lin, C. Liu, and R. Eisenberg, *PNP equations with steric effects: A model of ion flow through channels*, J. Phys. Chem. B, 116 (2012), pp. 11422–11441.
[24] C.-Y. Hsieh, Y. Hyon, H. Lee, T.-C. Lin, and C. Liu, *Transport of charged particles: Entropy production and maximum dissipation principle*, J. Comput. Chem., 34 (2013), pp. 2065–2078.
[25] X. Hu, J. Xu, and C. Zhang, *Application of auxiliary space preconditioning in field-scale reservoir simulation*, Sci. China Math., 56 (2013), pp. 2737–2751.
[26] Y. Hyon, B. Eisenberg, and C. Liu, *A mathematical model for the hard sphere repulsion in ionic solutions*, Commun. Math. Sci., 9 (2011), pp. 459–475.
[27] J. W. Jerome, *Consistency of semiconductor modeling: An existence/stability analysis for the stationary van Roosbroeck system*, SIAM J. Appl. Math., 45 (1985), pp. 565–590, https://doi.org/10.1137/0145034.
[28] J. W. Jerome, *Analysis of Charge Transport: A Mathematical Study of Semiconductor Devices*, Springer, Berlin, 1996.
[29] R. Karnik, C. Duan, K. Castelino, H. Daiguji, and A. Majumdar, *Rectification of ionic current in a nanofluidic diode*, Nano Lett., 7 (2007), pp. 547–551.
[30] T. Kerkhoven and J. W. Jerome, *$L_\infty$ stability of finite element approximations to elliptic gradient equations*, Numer. Math., 57 (1990), pp. 561–575.
[31] H. Kim, J. Xu, and L. Zikatanov, *A multigrid method based on graph matching for convection-diffusion equations*, Numer. Linear Algebra Appl., 10 (2003), pp. 181–195, https://doi.org/10.1002/nla.317.
[32] D. Li, *Electrokinetics in Microfluidics*, Vol. 2, Academic Press, New York, 2004.
[33] S. Liu, J. Savage, and G. A. Voth, *Mesoscale study of proton transport in proton exchange membranes: Role of morphology*, J. Phys. Chem. C, 119 (2015), pp. 1753–1762.
[34] A. Logg, K. A. Mardal, and G. Wells, *Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book*, Vol. 84, Springer Science & Business Media, Berlin, 2012.
[35] B. Lu, M. J. Holst, J. A. McCammon, and Y. Zhou, *Poisson–Nernst–Planck equations for simulating biomolecular diffusion–reaction processes* I*: Finite element solutions*, J. Comput. Phys., 229 (2010), pp. 6979–6994.
[36] K.-A. Mardal and R. Winther, *Uniform preconditioners for the time dependent Stokes problem*, Numer. Math., 98 (2004), pp. 305–327.
[37] K.-A. Mardal and R. Winther, *Preconditioning discretizations of systems of partial differential equations*, Numer. Linear Algebra Appl., 18 (2011), pp. 1–40.
[38] P. A. Markowich, *The Stationary Semiconductor Device Equations*, Springer-Verlag, Vienna, 1986.
[39] P. A. Markowich, C. A. Ringhofer, and C. Schmeiser, *Semiconductor Equations*, Springer-Verlag, Vienna, 1990.
[40] P. A. Markowich and M. A. Zlámal, *Inverse-average-type finite element discretizations of selfadjoint second-order elliptic problems*, Math. Comp., 51 (1988), pp. 431–449.
[41] D. Meng, B. Zheng, G. Lin, and M. L. Sushko, *Numerical solution of 3D Poisson-Nernst-Planck Equations coupled with classical density functional theory for modeling ion and elec-*

*tron transport in a confined environment*, Commun. Comput. Phys., 16 (2014), pp. 1298–1322.

[42] M. S. Metti, J. Xu, and C. Liu, *Energetically stable discretizations for charge transport and electrokinetic models*, J. Comput. Phys., 306 (2016), pp. 1–18.

[43] J. J. H. Miller and S. Wang, *A triangular mixed finite element method for the stationary semiconductor device equations*, RAIRO Modél. Math. Anal. Numér., 25 (1991), pp. 441–463.

[44] M. S. Mock, *Analysis of Mathematical Models of Semiconductor Devices*, Vol. 3, Boole Press, Dún Laoghaire, Ireland, 1983.

[45] M. Olshanskii, *Multigrid analysis for the time dependent Stokes problem*, Math. Comp., 81 (2012), pp. 57–79.

[46] W. Pan, K. Kim, M. Perego, A. M. Tartakovsky, and M. L. Parks, *Modeling electrokinetic flows by consistent implicit incompressible smoothed particle hydrodynamics*, J. Comput. Phys., 334 (2017), pp. 125–144.

[47] M. Pernice and M. D. Tocci, *A multigrid-preconditioned Newton–Krylov Method for the incompressible Navier–Stokes equations*, SIAM J. Sci. Comput., 23 (2001), pp. 398–418, https://doi.org/10.1137/S1064827500372250.

[48] E. G. Phillips, H. C. Elman, E. C. Cyr, J. N. Shadid, and R. P. Pawlowski, *A block preconditioner for an exact penalty formulation for stationary MHD*, SIAM J. Sci. Comput., 36 (2014), pp. B930–B951, https://doi.org/10.1137/140955082.

[49] A. Prohl and M. Schmuck, *Convergent discretizations for the Nernst–Planck–Poisson system*, Numer. Math., 111 (2009), pp. 591–630.

[50] A. Prohl and M. Schmuck, *Convergent finite element discretizations of the Navier-Stokes-Nernst-Planck-Poisson system*, M2AN Math. Model. Numer. Anal., 44 (2010), pp. 531–571.

[51] R. Sacco and M. Stynes, *Finite element methods for convection-diffusion problems using exponential splines on triangles*, Comput. Math. Appl., 35 (1998), pp. 35–45.

[52] D. Scharfetter and H. Gummel, *Large signal analysis of a silicon Read diode oscillator*, IEEE Trans. Electron Devices, 16 (1969), pp. 64–77.

[53] A. H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp., 28 (1974), pp. 959–962.

[54] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer-Verlag, Vienna, 2012.

[55] J. W. Slotboom, *Computer-aided two-dimensional analysis of bipolar transistors*, IEEE Trans. Electron Devices, 20 (1973), pp. 669–679.

[56] B. Tu, M. Chen, Y. Xie, L. Zhang, B. Eisenberg, and B. Lu, *A parallel finite element simulator for ion transport through three-dimensional ion channel systems*, J. Comput. Chem., 34 (2013), pp. 2065–2078.

[57] M. Wang and L. Chen, *Multigrid methods for the Stokes equations using distributive Gauss–Seidel relaxations based on the least squares commutator*, J. Sci. Comput., 56 (2013), pp. 409–431.

[58] J. Wu, V. Srinivasan, J. Xu, and C. Wang, *Newton-Krylov-Multigrid algorithms for battery simulation*, J. Electrochem. Soc., 149 (2002), pp. A1342–A1348.

[59] J. Xu, *Two-grid discretization techniques for linear and nonlinear PDEs*, SIAM J. Numer. Anal., 33 (1996), pp. 1759–1777, https://doi.org/10.1137/S0036142992232949.

[60] J. Xu and L. Zikatanov, *A monotone finite element scheme for convection-diffusion equations*, Math. Comp., 68 (1999), pp. 1429–1446.

[61] J. Xu and L. Zikatanov, *Algebraic multigrid methods*, Acta Numer., 26 (2017), pp. 591–721, https://doi.org/10.1017/S0962492917000083.

[62] S. Xu, M. Chen, S. Majd, X. Yue, and C. Liu, *Modeling and simulating asymmetrical conductance changes in gramicidin pores*, Mol. Based Math. Biol., 2 (2014), pp. 34–55.