Multiple solutions of steady-state Poisson–Nernst–Planck equations with steric effects

Multiple solutions of steady-state Poisson–Nernst–Planck equations with steric effects

Tai-Chia Lin[1] **and Bob Eisenberg**[2]

[1] Institute of Applied Mathematical Sciences, Center for Advanced Study in Theoretical Sciences (CASTS), National Taiwan University, No.1, section4, Roosevelt Road, Taipei 106, Taiwan
[2] Department of Molecular Biophysics & Physiology Rush Medical Center, 1653 West Congress, Parkway, Chicago, IL 60612, USA

E-mail: tclin@math.ntu.edu.tw and beisenbe@rush.edu

Abstract

Experiments measuring currents through single protein channels show unstable currents. Channels switch between 'open' or 'closed' states in a spontaneous stochastic process called gating. Currents are either (nearly) zero or at a definite level, characteristic of each type of protein, independent of time, once the channel is open. The steady state Poisson–Nernst–Planck equations with steric effects (PNP-steric equations) describe steady current through the open channel quite well, in a wide variety of conditions. Here we study the existence of multiple solutions of steady state PNP-steric equations to see if they themselves, without modification or augmentation, can describe two levels of current. We prove that there are two steady state solutions of PNP-steric equations for (a) three types of ion species (two types of cations and one type of anion) with a positive constant permanent charge, and (b) four types of ion species (two types of cations and their counter-ions) with a constant permanent charge but no sign condition. The excess currents (due to steric effects) associated with these two steady state solutions are derived and expressed as two distinct formulas. Our results indicate that PNP-steric equations may become a useful model to study spontaneous gating of ion channels. Spontaneous gating is thought to involve small structural changes in the channel protein that perhaps produce large changes in the profiles of free energy that determine ion flow. Gating is known to be modulated by external structures. Both can be included in future extensions of our present analysis.

Keywords: multiple solutions, excess currents, PNP-steric equations
Mathematics Subject Classification: 76A05, 76M99, 65C30

(Some figures may appear in colour only in the online journal)

1. Introduction

The Poisson–Nernst–Planck (PNP) equations, a well-known model of ion transport, play a crucial role in the study of many physical and biological phenomena (see [3, 5, 7, 8, 11, 13, 15, 16, 25, 33, 34, 35, 38]). Such an important model can be represented by

$$
\begin{cases}
\frac{\partial c_i}{\partial t} + \nabla \cdot J_i^{\mathrm{PNP}} = 0, \quad i = 1, \cdots, N, \\[2ex]
-J_i^{\mathrm{PNP}} = D_i \left(\nabla c_i + \frac{z_i e}{k_B T} c_i \nabla \phi \right), \\[2ex]
-\nabla \cdot (\varepsilon \nabla \phi) = \rho_0 + \sum_{i=1}^{N} z_i e c_i
\end{cases}
\tag{1.1}
$$

where N is the number of ion species, c_i is the distribution function, J_i^{PNP} is the flux density, D_i is the diffusion constant, and z_i is the valence of the ith ion species, respectively. Besides, ϕ is the electrostatic potential, ε is the dielectric constant, ρ_0 is the permanent (fixed) charge density of the system, k_B is the Boltzmann constant, T is the absolute temperature and e is the elementary charge. Due to ionic sizes, steric repulsion may appear in crowded ions of several biological systems like DNAs, ribosomes and ion channels. When ions are crowded in a narrow channel, the PNP equations become unreliable because the ion-size effect becomes important, but the PNP equations represent ions as point particles without size (see [1, 6, 18, 19, 22, 26, 30, 36]).

To include ion size effects, Eisenberg and Liu modified PNP equations into a complicated system of differential-integral equations with singular integrals that simulate successfully the selectivity of important types of calcium and sodium ion channels (see [24]). The model is denoted as follows:

$$
\partial_t c_j = \nabla \cdot \left(\frac{D_j}{k_B T} c_j \nabla \frac{\delta E_{tot}}{\delta c_j} \right) \quad \text{for} \quad j = 1, \cdots, N,
\tag{1.2}
$$

where the total energy E_{tot} is given by

$$
E_{\mathrm{tot}} = E_{\mathrm{PNP}} + E_{HS},
$$

$$
E_{\mathrm{PNP}} = \int_{\mathbb{R}^d} k_B T \sum_{j=1}^{N} c_j \ln c_j + \frac{1}{2} \left(\rho_0 + \sum_{j=1}^{N} z_j e c_j \right) \phi,
$$

$$
E_{\mathrm{HS}} = \sum_{i,j=1}^{N} \int \int_{\mathbb{R}^d \times \mathbb{R}^d} \Psi_{ij}(x - y) c_i(x) c_j(y) \, dx \, dy,
$$

and $\Psi_{ij}(x - y) = \epsilon_{ij}(a_i + a_j)^{12} |x - y|^{-12}$ for ith and jth ions located at x and y with the radii a_i, a_j, respectively. Here ϵ_{ij} is an appropriately chosen energy constant, which comes from the repulsive part of Lennard-Jones potential to describe the hard sphere repulsion of ions (see [24]). However, because Ψ_{ij}'s are extremely singular, E_{HS} becomes difficult to analyze theoretically and compute numerically when forced to deal with such singularities (see [17, 24]).

To simplify model (1.2), we truncate the (spatial) frequency range of Ψ_{ij} and get the approximate potential $\Psi_{ij,\sigma}$ denoted as (see [31])

$$
\Psi_{ij,\sigma}(z) = \left(\Psi_{ij} \chi_\sigma \right) \star \varphi_\sigma(z) \quad \text{for} \quad z \in \mathbb{R}^d,
\tag{1.3}
$$

where the asterisk is the standard convolution, χ_σ is the characteristic function of the exterior ball $\{z \in \mathbb{R}^d : |z| > \sigma\}$, and φ_σ is the spatially band-limited function defined by

$$
\varphi_\sigma(x) = (1 - \chi_{\sigma^{-\gamma}}(\xi))^v,
\tag{1.4}
$$

for $x, \xi \in \mathbb{R}^d$, where v denotes the inverse Fourier transform. Note that the truncation may lose the effect of high frequencies $|\xi| > \sigma^{-\gamma} (0 < \gamma < 1)$, but still involve a large part of the steric effects because it keeps the effect of frequencies like $|\xi| \sim \sigma^{-\alpha}$ for $0 < \alpha < \gamma$ tending to infinity as σ goes to zero. Then energy E_{HS} can be approximated by $E_{\mathrm{HS},\sigma}$ as follows:

$$E_{\mathrm{HS}} \sim \sum_{i,j=1}^{N} \int \int_{\mathbb{R}^d \times \mathbb{R}^d} \Psi_{ij,\sigma}(x-y)\, c_i(x)\, c_j(y)\, \mathrm{d}x\mathrm{d}y \equiv E_{\mathrm{HS},\sigma}\,,$$

$$\sim \sum_{i,j=1}^{N} \epsilon_{ij} (a_i + a_j)^{12} S_\sigma \int_{\mathbb{R}^d} c_i(x)\, c_j(x)\, \mathrm{d}x \equiv \tilde{E}_{\mathrm{HS},\sigma}\,,$$

as $\sigma > 0$ approaches zero, where $S_\sigma \sim \sigma^{d-12}$. Note that one may use the mean-field approximation to get the energy with the same form as $\tilde{E}_{\mathrm{HS},\sigma}$ which describes hard sphere repulsion of ions (see [10]), and the excess chemical potential μ_i^{ex} in (1.8) (see [4]). Here we give another argument to derive $\tilde{E}_{\mathrm{HS},\sigma}$ which is the leading order term of the expansion of the approximated energy $E_{\mathrm{HS},\sigma}$. To derive a simplified model of (1.2), we may fix $\sigma > 0$ as a constant and replace E_{HS} by $\tilde{E}_{\mathrm{HS},\sigma}$ in (1.2) to get the Poisson–Nernst–Planck equations with steric effects called PNP-steric equations denoted as (see [31])

$$\frac{\partial c_i}{\partial t} + \nabla \cdot J_i = 0, i = 1, \cdots, N\,, \tag{1.5}$$

$$-\nabla \cdot (\varepsilon \nabla \phi) = \rho_0 + \sum_{i=1}^{N} z_i e c_i\,, \tag{1.6}$$

where flux J_i is

$$J_i = -D_i \nabla c_i - \frac{D_i c_i}{k_{\mathrm{B}} T} z_i e \nabla \phi - \frac{D_i c_i}{k_{\mathrm{B}} T} \sum_{j=1}^{N} g_{ij} \nabla c_j\,, \tag{1.7}$$

and $g_{ij} = g_{ji} \sim \epsilon_{ij}(a_i + a_j)^{12}$ is a nonnegative constant depending on ion radii a_i, a_j and the energy coupling constant ϵ_{ij} of the i-th and j-th species ions, respectively. Note that equations (1.5)–(1.7) can be regarded as a system of reaction-diffusion equations with nonlinear cross-diffusion terms being similar to [9]. Amazingly, these equations are an effective model to simulate the selectivity of ion channels (see [21]).

Comparing (1.7) with J_i^{PNP} in (1.1), the excess flux $J_i^{ex} = J_i - J_i^{\mathrm{PNP}}$ due to steric effects of ion species i is

$$-J_i^{ex} = \frac{1}{k_{\mathrm{B}} T} D_i c_i \nabla \mu_i^{ex} \quad \text{and} \quad \mu_i^{ex} = \sum_{j=1}^{N} g_{ij} c_j \tag{1.8}$$

where $\mu_i^{ex} = \sum_{j=1}^{N} g_{ij} c_j$ is the excess chemical potential of ion species i due to steric effects.

Consequently, the excess current $I^{ex} = \sum_{i=1}^{N} z_i e J_i^{ex}$ due to steric effects becomes

$$I^{ex} = -\sum_{i,j=1}^{N} \frac{z_i e}{k_{\mathrm{B}} T} D_i g_{ij} c_i \nabla c_j\,. \tag{1.9}$$

We shall use the formula (1.9) to calculate the excess currents for multiple solutions of the 1D steady-state PNP-steric equations. We are motivated by the hope–but cannot dare expect– that one solution will correspond to a closed state and the other to an open state, as found in

experiments [14] and in simulations [27]. Of course, the current measured through the open state corresponds to the total current, not just the excess currents.

The existence of multiple steady (equilibrium) states is important to study transitions between such states which may be related to the gating (switching between open and closed states) and selectivity of ion channels. Multiple steady states can be investigated by finding multiple solutions of the 1D steady-state PNP equations for two types of ion species with three regions of piecewise constant permanent charge under the assumption that the Debye number is large [15]. More general theorems related to multiple solutions of the 1D steady-state PNP equations involving multiple types of ions with multiple regions of piecewise constant permanent charge are discussed in [32]. With only a constant permanent charge, there is only a unique solution of the 1D steady-state PNP equations for multiple types of ions [34, 37]. Instead of the 1D steady-state PNP equations, here we study multiple solutions of the 1D steady-state PNP-steric equations with spatially constant permanent charges. Our use here of equilibrium states is for a particular issue showing how multiple solutions with the flavor of gating phenomena might appear in the equilibrium case. No claims are made of generality in our discussion, especially to describe permeation in ion channels which should use the full PNP equations with Dirichlet boundary conditions.

For simplicity, we consider domain as a 1D interval $(-1, 1)$ for (1.5)–(1.7) and set $J_i = 0$, $i = 1, \cdots, N$ to get the steady-state PNP-steric equations. Then by (1.7),

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\ln c_i + \frac{z_i e}{k_\mathrm{B} T}\phi + \frac{1}{k_\mathrm{B} T}\sum_{j=1}^{N} g_{ij}c_j\right) = 0 \quad \text{for } x \in (-1, 1), i = 1, \cdots, N,$$

which can be satisfied if

$$\ln c_i + \frac{z_i e}{k_\mathrm{B} T}\phi + \frac{1}{k_\mathrm{B} T}\sum_{j=1}^{N} g_{ij}c_j = 0 \quad \text{for } i = 1, \cdots, N, \tag{1.10}$$

holds true. Let $\tilde{\phi} = \frac{e}{k_\mathrm{B} T}\phi$ and $\tilde{g}_{ij} = \frac{1}{k_\mathrm{B} T}g_{ij}$ for $i, j = 1, \cdots, N$. Then (1.6) and (1.10) can be transformed into

$$\ln c_i + z_i\tilde{\phi} + \sum_{j=1}^{N} \tilde{g}_{ij}c_j = 0 \quad \text{for } i = 1, \cdots, N, \tag{1.11}$$

and

$$-\tilde{\varepsilon}\tilde{\phi}_{xx} = \tilde{\rho}_0 + \sum_{i=1}^{N} z_i c_i \quad \text{for } x \in (-1, 1), \tag{1.12}$$

where $\tilde{\varepsilon} = \frac{k_\mathrm{B} T}{e^2}\varepsilon$ and $\tilde{\rho}_0 = \frac{1}{e}\rho_0$. For notational convenience, we may remove tilde ($\sim$) and denote (1.11) and (1.12) as

$$\ln c_i + z_i\phi + \sum_{j=1}^{N} g_{ij}c_j = 0 \quad \text{for } i = 1, \cdots, N, \tag{1.13}$$

and

$$-\varepsilon\phi_{xx} = \rho_0 + \sum_{i=1}^{N} z_i c_i \quad \text{for } x \in (-1, 1). \tag{1.14}$$

Equations like (1.13) have been used to interpret bioelectric phenomena in many papers since they were adopted by Hodgkin, Huxley, and Cole (see [12,23]). Here we consider the following boundary condition given by

$$\phi(1) + \eta_\varepsilon\phi'(1) = \phi_0(1) \quad \text{and} \quad \phi(-1) - \eta_\varepsilon\phi'(-1) = \phi_0(-1), \tag{1.15}$$

where $\phi_0(1)$, $\phi_0(-1)$ are constants and η_ε is a non-negative constant. Here $\phi_0(\pm 1)$ and $\phi(\pm 1)$ are the extrachannel and intrachannel electrostatic potentials at the channel boundaries, respectively. The coefficient $\eta_\varepsilon \sim \frac{\varepsilon_0}{\varepsilon_m}$ is governed by the ratio of ε_0 the dielectric constant of the electrolyte solution and ε_m the dielectric constant of the membrane (see [39]). Note that (1.15) is of the Robin boundary condition if $\eta_\varepsilon > 0$; and of the Dirichlet boundary condition if $\eta_\varepsilon = 0$. The Robin boundary condition includes polarization (e.g. dielectric) charges in the bath and/or electrodes which the Dirichlet boundary condition does not. Such charges, induced by and dependent on the electric field play a prominent role in the art of real experiments, because they are important determinants of the background noise and stability of high speed recordings. The theoretical reasons for these practical realities have not been investigated to the best of our knowledge.

As $N = 2$, the existence, uniqueness and the solution's asymptotic behavior of (1.13)–(1.15) are investigated under non-symmetry breaking condition $0 \leqslant g_{12} = g_{21} \leqslant \sqrt{g_{11}g_{22}}$ which implies that solution (c_1, c_2) of (1.13) is uniquely determined by ϕ (see [28]). Moreover, (1.13) and (1.14) can be reduced to a single differential equation of ϕ, which is of Poisson–Boltzmann (PB) type equations (see [29]). However, as the symmetry breaking condition $g_{12} = g_{21} > \sqrt{g_{11}g_{22}}$ holds true and g_{12} is sufficiently large, solution (c_1, c_2) of (1.13) may form two branch curves of ϕ so (1.13) and (1.14) can be decomposed into two different equations (3.6) and (3.7) but not only one PB type equation. Hence we cannot regard steady state PNP-steric equations (1.13) and (1.14) as a single PB type equation as g_{12} is sufficiently large. In section 2, we introduce new variables ξ, Σ and transform (1.13) into a quadratic polynomial which can be solved precisely to get explicit formulas and represent two branches of solution curves. Using these explicit formulas, we can then define biological conductance (for that condition) as the biologists do and perform the comparison using formulas like (1.16)–(1.19). Note that the symbol g is used for conductance (units siemens) in biology and this is not equivalent to our g_{ij}. In this paper, we want to study multiple solutions of (1.13)–(1.15) for the cases of $N = 3, 4$, and $g_{12} = g_{21}$, $g_{34} = g_{43}$ sufficiently large such that symmetry breaking condition $g_{12} = g_{21} > \sqrt{g_{11}g_{22}}$, $g_{34} = g_{43} > \sqrt{g_{33}g_{44}}$ holds true.

1.1. Main Results

System (1.13) can be regarded as a coupled system of algebraic equations. Because $g_{ij} = 0$ for $i, j = 1, \cdots, N$, a solution of system (1.13) can be expressed as $c_i = e^{-z_i\phi}$ for $i = 1, \cdots, N$. However, it seems impossible to solve system (1.13) explicitly for the general case of $g_{ij} > 0$ for $i, j = 1, \cdots, N$. To overcome such difficulty, we may set $N = 2$, $z_2 = -z_1 = q \geqslant 1$, $g_{11} = g_{22} = g > 0$, and introduce new variables $\xi = c_1 c_2$ and $\Sigma = c_1 + c_2$. Then (1.13) can be transformed into a quadratic polynomial that can be solved explicitly (see section 2). For $g_{12} = g_{21} = z$ large (see theorem 2.4 in section 2), system (1.13) has two branches of solutions $(c_1, c_2) = (c_1(\Sigma_{A_1}(\phi)), c_2(\Sigma_{A_1}(\phi)))$ and $(c_1, c_2) = (c_1(\Sigma_{B_1}(\phi)), c_2(\Sigma_{B_1}(\phi)))$ such that $(c_1 - c_2) \circ \Sigma_{A_1} : [-\phi_{A,c}, \infty) \to \mathbb{R}$ and $(c_1 - c_2) \circ \Sigma_{B_1} : (-\infty, \phi_{A,c}] \to \mathbb{R}$ are monotone increasing functions to ϕ, where $\phi_{A,c} > 0$ is a constant, Σ_{A_1} and Σ_{B_1} are two functions satisfying

$$(c_1 - c_2) \circ \Sigma_{A_1}(-\phi_{A,c}) = (c_1 - c_2)(\Sigma_c) > 0,$$

$$(c_1 - c_2) \circ \Sigma_{B_1}(\phi_{A,c}) = -(c_1 - c_2)(\Sigma_c) < 0,$$

$$\lim_{\phi \to \infty} (c_1 - c_2) \circ \Sigma_{A_1}(\phi) = \infty \quad \text{and} \quad \lim_{\phi \to -\infty} (c_1 - c_2) \circ \Sigma_{B_1}(\phi) = -\infty.$$

Here $\circ$ denotes the function $(c_1 - c_2)$ acting on the function $\Sigma_{A_1}(\phi)$, i.e. the function composition and g_c is the positive constant defined in proposition 2.2. Besides, $\phi_{A,c}$ satisfies

$\phi_{A,c} \to +\infty$ and $(c_1 - c_2)(\Sigma_c) \to 0$ as $z \to +\infty$ and $g > 0$ is fixed. Hence (1.13) and (1.14) can be decomposed into two differential equations like (3.6) and (3.7) but they can not have uniformly bounded solutions to $\varepsilon > 0$ (see lemma 4.5). This fact motivates us to add one extra species c_3 and assume that $N = 3$, $g_{12} = g_{21} = z$ is sufficiently large, $g_{11} = g_{22} = g > 0$, $z_2 = -z_1 = q \geqslant 1$, $z_3 > 0$, $g_{i3} = g_{3i} = 0$, $i = 1, 2, 3$ (which implies $c_3 = e^{-z_3\phi}$). Then (1.13) and (1.14) may be reduced to two differential equations (3.6) and (3.7) having uniformly bounded solutions, respectively. This may provide multiple solutions of (1.13)–(1.15).

Natural biological solutions always contain at least three species (sodium, potassium, and chloride, and usually calcium). Experiments are often done, however, with just two species (say sodium chloride) along with traces of hydrogen ion, and perhaps other contaminants. Gating occurs in simplified unnatural situations and so we hope to study mathematical solutions in corresponding situations in a separate paper.

Now we state the main result of this paper as follows:

Theorem 1.1. *Let $N = 3$, $z_2 = -z_1 = q \geqslant 1$, $z_3 > 0$ and $\rho_0 > 0$ be a constant. Assume that $g_{11} = g_{22} = g > 0$ is fixed and $g_{i3} = g_{3i} = 0$ for $i = 1, 2, 3$. Then as $g_{12} = g_{21} = z > 0$ is sufficiently large, the system of equations (1.13)–(1.15) has two uniformly bounded (to ε) solutions ϕ_ε^A and ϕ_ε^B such that $\phi_\varepsilon^A(x) \to \phi_{A_1,0}$ and $\phi_\varepsilon^B(x) \to \phi_{B_1,0}$ for $x \in (-1, 1)$ as $\varepsilon \to 0$. Here $\phi_{A_1,0}$ and $\phi_{B_1,0}$ are two distinct constants satisfying $f_A(\phi_{A_1,0}) = f_B(\phi_{B_1,0}) = 0$, where*

$$f_A(\phi) = q(c_1 - c_2)\left(\Sigma_{A_1}(\phi)\right) - z_3 e^{-z_3\phi} - \rho_0,$$

and

$$f_B(\phi) = q(c_1 - c_2)\left(\Sigma_{B_1}(\phi)\right) - z_3 e^{-z_3\phi} - \rho_0.$$

Remark 1.2. For the case of $N = 2$, solutions ϕ_ε^A and ϕ_ε^B may be unbounded as ε approaches zero (see lemma 4.5). To get the uniform boundedness of solutions ϕ_ε^A and ϕ_ε^B, we need to consider the case of $N = 3$. Without condition $g_{i3} = g_{3i} = 0$ for $i = 1, 2, 3$ in theorem 1.1, equation (1.13) becomes very complicated and it seems impossible to get the multiplicity and monotonicity of $\sum_{i=1}^{3} z_i c_i + \rho_0$ (to ϕ) which is important to prove the multiple solutions of (1.13)–(1.15). There is no general theorem for the multiplicity and monotonicity of $\sum_{i=1}^{3} z_i c_i + \rho_0$ (to ϕ), especially for all g_{ij}'s are nonzero. Here we use such a condition ($g_{i3} = g_{3i} = 0$ for $i = 1, 2, 3$) to reduce the problem to the $N = 2$ case so we may apply results of section 2 to get two branches of solutions $(c_1, c_2) = (c_1(\Sigma_{A_1}(\phi)), c_2(\Sigma_{A_1}(\phi)))$ and $(c_1, c_2) = (c_1(\Sigma_{B_1}(\phi)), c_2(\Sigma_{B_1}(\phi)))$, and the monotonicity of f_A and f_B to ϕ, provided that $g_{11} = g_{22} = g > 0$ is fixed and $g_{12} = g_{21} = z > 0$ is sufficiently large. Moreover, $g_{i3} = g_{3i} = 0$ implies $c_3 = e^{-z_3\phi}$ and the profile of f_A and f_B may touch zero at $\phi_{A_1,0}$ and $\phi_{B_1,0}$ (see figure 4 in page 15), which is important to prove the uniform boundedness of solutions ϕ_ε^A and ϕ_ε^B using lemma 4.1.

In most of the 'cation' (e.g. sodium, potassium, and calcium) channels, ρ_0 is a negative number. There are regions ('rings') of negative charge and some channels (sodium channel DEKA) have a ring of positive charge as well. Here we assume the positive sign of ρ_0 which may produce the values $\phi_{A_1,0}$ and $\phi_{B_1,0}$ (see figure 4 in section 3.1), and the proof of theorem 1.1 is given in section 3.1.

To remove the sign condition on ρ_0, we may consider four ion species composed of two cations and counterions (like the mixture of Na^+, Ca^{+2}, Cl^- and CO_3^{-2}) and study multiple solutions of (1.13)–(1.15) with $N = 4$, $z_2 = -z_1 = q_1 \geqslant 1$, $z_4 = -z_3 = q_2 \geqslant 1$, $g_{11} = g_{22} = g > 0$, and $g_{33} = g_{44} = \tilde{g} > 0$. Using the assumption $g_{ij} = g_{ji} = 0$ for $i = 1, 2$ and $j = 3, 4$, we may decompose system (1.13) with $N = 4$ into two independent

systems having the same form as (1.13) with $N = 2$. Hence theorem 2.4 (in section 2) implies that as $g_{12} = g_{21} = z$ and $g_{34} = g_{43} = \tilde{z} > 0$ sufficiently large, system (1.13) has four branches of solutions

$$(c_1, c_2) = \left(c_1\left(\Sigma_{A_1}(\phi)\right), c_2\left(\Sigma_{A_1}(\phi)\right)\right), \quad (c_1, c_2) = \left(c_1\left(\Sigma_{B_1}(\phi)\right), c_2\left(\Sigma_{B_1}(\phi)\right)\right),$$
$$(c_3, c_4) = \left(c_3\left(\Sigma_{M_1}(\phi)\right), c_4\left(\Sigma_{M_1}(\phi)\right)\right), \quad (c_3, c_4) = \left(c_3\left(\Sigma_{N_1}(\phi)\right), c_4\left(\Sigma_{N_1}(\phi)\right)\right),$$

such that $(c_1 - c_2) \circ \Sigma_{A_1} : [-\phi_{A,c}, \infty) \rightarrow \mathbb{R}$, $(c_1 - c_2) \circ \Sigma_{B_1} : (-\infty, \phi_{A,c}] \rightarrow \mathbb{R}$, $(c_3 - c_4) \circ \Sigma_{N_1} : [-\phi_{M,c}, \infty) \rightarrow \mathbb{R}$ and $(c_3 - c_4) \circ \Sigma_{M_1} : (-\infty, \phi_{M,c}] \rightarrow \mathbb{R}$, are monotone increasing functions of ϕ, where $\phi_{A,c}, \phi_{M,c} > 0$ are constants, $\Sigma_{A_1}, \Sigma_{B_1}, \Sigma_{M_1}$ and Σ_{N_1} are functions satisfying

$$(c_1 - c_2) \circ \Sigma_{A_1}(-\phi_{A,c}), (c_3 - c_4) \circ \Sigma_{N_1}(-\phi_{M,c}) > 0,$$
$$(c_1 - c_2) \circ \Sigma_{B_1}(\phi_{A,c}), (c_3 - c_4) \circ \Sigma_{M_1}(\phi_{M,c}) < 0,$$
$$\lim_{\phi \to \infty}(c_1 - c_2) \circ \Sigma_{A_1}(\phi) = \lim_{\phi \to \infty}(c_3 - c_4) \circ \Sigma_{N_1}(\phi) = \infty,$$
$$\lim_{\phi \to -\infty}(c_1 - c_2) \circ \Sigma_{B_1}(\phi) = \lim_{\phi \to -\infty}(c_3 - c_4) \circ \Sigma_{M_1}(\phi) = -\infty.$$

Here $\circ$ denotes function composition. Moreover, $\phi_{A,c}, \phi_{M,c} \to +\infty$ and $(c_1 - c_2) \circ \Sigma_{A_1}(-\phi_{A,c}), (c_1 - c_2) \circ \Sigma_{B_1}(\phi_{A,c}), (c_3 - c_4) \circ \Sigma_{M_1}(\phi_{M,c})$ and $(c_3 - c_4) \circ \Sigma_{N_1}(-\phi_{M,c})$ tend to zero as $z, \tilde{z} \to +\infty$ and $g, \tilde{g} > 0$ are fixed.

Without loss of generality, we may assume $\phi_{M,c} < \phi_{A,c}$. Then the graphs of functions $(c_1 - c_2) \circ \Sigma_{A_1}$ and $(c_4 - c_3) \circ \Sigma_{M_1}$ may intersect at $\phi = \phi_{A_1,0}$ as z and $\tilde{z}$ sufficiently large (see figure 5 in section 3.2). Similarly, the graphs of functions $(c_1 - c_2) \circ \Sigma_{B_1}$ and $(c_4 - c_3) \circ \Sigma_{N_1}$ may intersect at $\phi = \phi_{B_1,0}$ as z and $\tilde{z}$ sufficiently large. Hence (1.13) and (1.14) may be reduced to two differential equations with the same forms as (3.6) and (3.7) having uniformly bounded solutions, respectively. This may provide the following result for multiple solutions of (1.13)–(1.15).

Theorem 1.3. *Let $N = 4$, $z_2 = -z_1 = q_1 \geqslant 1$, $z_4 = -z_3 = q_2 \geqslant 1$ and $\rho_0 \neq 0$ be a constant. Assume that $g_{11} = g_{22} = g > 0$, $g_{33} = g_{44} = \tilde{g} > 0$ are fixed and $g_{ij} = g_{ji} = 0$ for $i = 1, 2$ and $j = 3, 4$. Then as $g_{12} = g_{21} = z > 0$ and $g_{34} = g_{43} = \tilde{z} > 0$ are sufficiently large, the system of equations (1.13)–(1.15) has two uniformly bounded (to ε) solutions ϕ_ε^A and ϕ_ε^B such that $\phi_\varepsilon^A(x) \to \phi_{A_1,0}$ and $\phi_\varepsilon^B(x) \to \phi_{B_1,0}$ for $x \in (-1, 1)$ as $\varepsilon \to 0$. Here $\phi_{A_1,0}$ and $\phi_{B_1,0}$ are two distinct constants satisfying $f_A(\phi_{A_1,0}) = f_B(\phi_{B_1,0}) = 0$, where*

$$f_A(\phi) = q_1(c_1 - c_2)\left(\Sigma_{A_1}(\phi)\right) + q_2(c_3 - c_4)\left(\Sigma_{M_1}(\phi)\right) - \rho_0,$$

and

$$f_B(\phi) = q_1(c_1 - c_2)\left(\Sigma_{B_1}(\phi)\right) + q_2(c_3 - c_4)\left(\Sigma_{N_1}(\phi)\right) - \rho_0.$$

The proof of theorem 1.3 is given in section 3.2.

Remark 1.4. As for remark 1.2, we use condition $g_{ij} = g_{ji} = 0$ for $i = 1, 2$ and $j = 3, 4$ to reduce (1.13) with $N = 4$ into two independent systems of (c_1, c_2) and (c_3, c_4) with the same form as (1.13) with $N = 2$ (studied in section 2) which give two branches of solutions

$$(c_1, c_2, c_3, c_4) = \left(c_1\left(\Sigma_{A_1}(\phi)\right), c_2\left(\Sigma_{A_1}(\phi)\right), c_3\left(\Sigma_{M_1}(\phi)\right), c_4\left(\Sigma_{M_1}(\phi)\right)\right),$$

and

$$(c_1, c_2, c_3, c_4) = \left(c_1\left(\Sigma_{B_1}(\phi)\right), c_2\left(\Sigma_{B_1}(\phi)\right), c_3\left(\Sigma_{N_1}(\phi)\right), c_4\left(\Sigma_{N_1}(\phi)\right)\right),$$

and the monotonicity of f_A and f_B to ϕ, provided that $g_{11} = g_{22} = g > 0$, $g_{33} = g_{44} = \tilde{g} > 0$ are fixed and $g_{12} = g_{21} = z > 0$, $g_{34} = g_{43} = \tilde{z} > 0$ are sufficiently large. As for theorem 1.1, we may use lemma 4.1 to prove the uniform boundedness of solutions ϕ_ε^A and ϕ_ε^B.

For solutions ϕ_ε^A and ϕ_ε^B, the corresponding excess currents defined in (1.9) may be denoted as I_A^{ex} and I_B^{ex}, respectively. Under the same hypotheses of theorem 1.1 for three ion species, we may use the new variable Σ to derive the following formulas (see section 5.1):

$$\int_{x_1}^{x_2} I_A^{ex}\mathrm{d}x = q\,e \int_{\Sigma_1^A}^{\Sigma_2^A} \frac{D_2 - D_1}{2} \left\{ (1-q) - q\left[g\Sigma + \left(g^2 - z^2\right)\mathrm{e}^{-(g+z)\Sigma}\right]\right\}\mathrm{d}\Sigma$$

$$-q\,e \int_{\Sigma_1^A}^{\Sigma_2^A} \frac{D_1 + D_2}{2\sqrt{\Sigma^2 - 4\mathrm{e}^{-(g+z)\Sigma}}} \left\{ (1-q)\Sigma - q\,g\,\Sigma^2 + (g+z)\left[2 - q\,(g-z)\,\Sigma\right]\mathrm{e}^{-(g+z)\Sigma}\right\}\mathrm{d}\Sigma,$$

$$(1.16)$$

and

$$\int_{x_1}^{x_2} I_B^{ex}\mathrm{d}x = q\,e \int_{\Sigma_1^B}^{\Sigma_2^B} \frac{D_2 - D_1}{2} \left\{ (1-q) - q\left[g\Sigma + \left(g^2 - z^2\right)\mathrm{e}^{-(g+z)\Sigma}\right]\right\}\mathrm{d}\Sigma$$

$$+q\,e \int_{\Sigma_1^B}^{\Sigma_2^B} \frac{D_1 + D_2}{2\sqrt{\Sigma^2 - 4\mathrm{e}^{-(g+z)\Sigma}}} \left\{ (1-q)\Sigma - q\,g\,\Sigma^2 + (g+z)\left[2 - q\,(g-z)\,\Sigma\right]\mathrm{e}^{-(g+z)\Sigma}\right\}\mathrm{d}\Sigma,$$

$$(1.17)$$

for $-1 < x_1 < x_2 < 1$, where $\Sigma_j^A = \Sigma_{A_1}\left(\phi_\varepsilon^A(x_j)\right)$ and $\Sigma_j^B = \Sigma_{B_1}\left(\phi_\varepsilon^B(x_j)\right)$ for $j = 1, 2$. From (1.16) and (1.17), it is clear that the difference between I_A^{ex} and I_B^{ex} which may give various ion flows related to currents observed in channels as they switch (i.e. gate) from one level of current to another.

The method of section 5.1 can be generalized to four ion species with the same hypotheses of theorem 1.3. As for (1.16) and (1.17), we may derive (see section 5.2)

$$\int_{x_1}^{x_2} I_{A,M}^{ex}\mathrm{d}x$$

$$= \int_{x_1}^{x_2} I_A^{ex} + I_M^{ex}\mathrm{d}x = q_1 e \int_{\Sigma_1^A}^{\Sigma_2^A} \frac{D_2 - D_1}{2} \left\{ (1-q_1) - q_1\left[g\Sigma + \left(g^2 - z^2\right)\mathrm{e}^{-(g+z)\Sigma}\right]\right\}\mathrm{d}\Sigma$$

$$-q_1 e \int_{\Sigma_1^A}^{\Sigma_2^A} \frac{D_1 + D_2}{2\sqrt{\Sigma^2 - 4\mathrm{e}^{-(g+z)\Sigma}}} \left\{ (1-q_1)\Sigma - q_1\,g\,\Sigma^2 + (g+z)\left[2 - q_1\,(g-z)\,\Sigma\right]\mathrm{e}^{-(g+z)\Sigma}\right\}\mathrm{d}\Sigma$$

$$+q_2 e \int_{\Sigma_1^M}^{\Sigma_2^M} \frac{D_4 - D_3}{2} \left\{ (1-q_2) - q_2\left[\tilde{g}\Sigma + \left(\tilde{g}^2 - \tilde{z}^2\right)\mathrm{e}^{-(\tilde{g}+\tilde{z})\Sigma}\right]\right\}\mathrm{d}\Sigma$$

$$-q_2 e \int_{\Sigma_1^M}^{\Sigma_2^M} \frac{D_3 + D_4}{2\sqrt{\Sigma^2 - 4\mathrm{e}^{-(\tilde{g}+\tilde{z})\Sigma}}} \left\{ (1-q_2)\Sigma - q_2\,\tilde{g}\,\Sigma^2 + (\tilde{g}+\tilde{z})\left[2 - q_2\,(\tilde{g}-\tilde{z})\,\Sigma\right]\mathrm{e}^{-(\tilde{g}+\tilde{z})\Sigma}\right\}\mathrm{d}\Sigma,$$

$$(1.18)$$

and

$$\int_{x_1}^{x_2} I_{B,N}^{ex}\mathrm{d}x$$

$$= int_{x_1}^{x_2} I_B^{ex} + I_N^{ex}\mathrm{d}x = q_1 e \int_{\Sigma_1^B}^{\Sigma_2^B} \frac{D_2 - D_1}{2} \left\{ (1-q_1) - q_1\left[g\Sigma + \left(g^2 - z^2\right)\mathrm{e}^{-(g+z)\Sigma}\right]\right\}\mathrm{d}\Sigma$$

$$+q_1 e \int_{\Sigma_1^B}^{\Sigma_2^B} \frac{D_1 + D_2}{2\sqrt{\Sigma^2 - 4\mathrm{e}^{-(g+z)\Sigma}}} \left\{ (1-q_1)\Sigma - q_1\,g\,\Sigma^2 + (g+z)\left[2 - q_1\,(g-z)\,\Sigma\right]\mathrm{e}^{-(g+z)\Sigma}\right\}\mathrm{d}\Sigma$$

$$+q_2 e \int_{\Sigma_1^N}^{\Sigma_2^N} \frac{D_4 - D_3}{2} \left\{ (1-q_2) - q_2\left[\tilde{g}\Sigma + \left(\tilde{g}^2 - \tilde{z}^2\right)\mathrm{e}^{-(\tilde{g}+\tilde{z})\Sigma}\right]\right\}\mathrm{d}\Sigma$$

$$+q_2 e \int_{\Sigma_1^N}^{\Sigma_2^N} \frac{D_3 + D_4}{2\sqrt{\Sigma^2 - 4e^{-(\tilde{g}+\tilde{z})\Sigma}}} \left\{ (1-q_2)\Sigma - q_2\,\tilde{g}\,\Sigma^2 + (\tilde{g}+\tilde{z}) \left[2 - q_2\,(\tilde{g}-\tilde{z})\,\Sigma \right] e^{-(\tilde{g}+\tilde{z})\Sigma} \right\} d\Sigma,$$

(1.19)

where $\Sigma_j^A = \Sigma_{A_1}(\phi_\varepsilon^A(x_j))$, $\Sigma_j^M = \Sigma_{M_1}(\phi_\varepsilon^A(x_j))$, $\Sigma_j^B = \Sigma_{B_1}(\phi_\varepsilon^B(x_j))$, and $\Sigma_j^N = \Sigma_{N_1}(\phi_\varepsilon^B(x_j))$ for $j = 1, 2$. The difference between $I_{A,M}^{ex}$ and $I_{B,N}^{ex}$ may also give various ion flows related to currents observed in channels as they switch (i.e. gate) from one level of current to another. Formulas (1.16)–(1.19) show how multiple solutions with the flavor of gating phenomena might appear in the equilibrium case.

The rest of this paper is organized as follows: We may solve system (1.13) of algebraic equations with $N = 2$, $z_2 = -z_1 = q \geqslant 1$ and $g_{11} = g_{22} > 0$ in section 2. Theorem 1.1 and 1.3 are proven in section 3. The proofs of lemma 4.1 and 4.5 are given in section 4, and formulas (1.16)–(1.19) are derived in section 5.

2. Solutions of (1.13) with $N = 2$, $z_2 = -z_1 = q \geqslant 1$ and $g_{11} = g_{22}$

In this section, we study equation (1.13) with $N = 2$, $z_2 = -z_1 = q \geqslant 1$ and $g_{11} = g_{22} = g$ which can be denoted as follows:

$$(\ln c_1 - q\,\phi) + (g\,c_1 + z\,c_2) = 0,$$

(2.1)

$$(\ln c_2 + q\,\phi) + (g\,c_2 + z\,c_1) = 0,$$

(2.2)

where $z = g_{12}$ and $g = g_{11} = g_{22}$ are positive constants. Physically, $g_{ij} \sim \epsilon_{ij}(a_i + a_j)^{12}$, where a_i is the ion radius of i-th ion species with concentration c_i, and $\epsilon_{ij} > 0$ is the energy coupling constant between i-th and j-th ion species for $i = 1, 2$. Note that (2.1) and (2.2) are formulated as a system of algebraic equations. We want to solve these equations and get solutions for (c_1, c_2) as a function of ϕ. Adding (2.1) and (2.2), we get

$$\ln(c_1 c_2) + (g + z)(c_1 + c_2) = 0.$$

(2.3)

Now we introduce new variables as follows:

$$\xi = c_1 c_2 \quad \text{and} \quad \Sigma = c_1 + c_2.$$

Multiplying Σ by c_1, we get a quadratic polynomial of c_1 as follows:

$$\Sigma c_1 = c_1^2 + \xi$$

which gives $c_1 = \frac{\Sigma \pm \sqrt{\Sigma^2 - 4\xi}}{2}$ and hence by $c_1 c_2 = \xi$, (c_1, c_2) can be expressed as

$$(c_1, c_2) = \left(\frac{\Sigma + \sqrt{\Sigma^2 - 4\xi}}{2}, \frac{\Sigma - \sqrt{\Sigma^2 - 4\xi}}{2} \right),$$

or

$$(c_1, c_2) = \left(\frac{\Sigma - \sqrt{\Sigma^2 - 4\xi}}{2}, \frac{\Sigma + \sqrt{\Sigma^2 - 4\xi}}{2} \right),$$

(2.4)

for $\Sigma \geqslant 2\sqrt{\xi} > 0$. Moreover, (2.3) can be transformed into $\ln \xi = -(g + z)\Sigma$ i.e.

$$\xi = e^{-(g+z)\Sigma}.$$

(2.5)

Hence the solution (c_1, c_2) of (2.1) and (2.2) may be described by two curves A and B parameterized by the total concentration Σ and denoted as

$$A = \left\{ (c_1, c_2) = \left(\frac{\Sigma + \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}}{2}, \frac{\Sigma - \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}}{2} \right) : \Sigma \geqslant \Sigma_z > 0 \right\},$$

(2.6)

and

$$B = \left\{ (c_1, c_2) = \left(\frac{\Sigma - \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}}{2}, \frac{\Sigma + \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}}{2} \right) : \Sigma \geqslant \Sigma_z > 0 \right\}.$$

(2.7)

Here $\Sigma_z > 0$ a critical total concentration is the unique positive solution of $\Sigma^2 = 4e^{-(g+z)\Sigma}$ such that concentrations c_1 and c_2 are equal to $\frac{1}{2}\Sigma_z$ as the total concentration $\Sigma = \Sigma_z$. Then

$$c_1 - c_2 = \begin{cases} \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}} & \text{on } A, \\ -\sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}} & \text{on } B. \end{cases}$$

(2.8)

Take (2.6) and (2.7) into (2.1), and let $\phi_A = \phi$ on curve A, and $\phi_B = \phi$ on curve B, respectively. Then

$$\begin{cases} q\,\phi_A(\Sigma) = \ln\left[\frac{1}{2}\left(\Sigma + \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}\right)\right] + \frac{g+z}{2}\Sigma + \frac{g-z}{2}\sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}, \\[2mm] q\,\phi_B(\Sigma) = \ln\left[\frac{1}{2}\left(\Sigma - \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}\right)\right] + \frac{g+z}{2}\Sigma + \frac{z-g}{2}\sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}, \end{cases}$$

(2.9)

for $\Sigma \geqslant \Sigma_z$. Consequently,

$$\phi_A + \phi_B = 0,$$

(2.10)

$$q\,\frac{\mathrm{d}\phi_A}{\mathrm{d}\Sigma} = \frac{(1 + g\Sigma)e^{(g+z)\Sigma} + g^2 - z^2}{e^{(g+z)\Sigma}\sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}},$$

(2.11)

and

$$q\,\frac{\mathrm{d}\phi_B}{\mathrm{d}\Sigma} = -\frac{(1 + g\Sigma)e^{(g+z)\Sigma} + g^2 - z^2}{e^{(g+z)\Sigma}\sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}},$$

(2.12)

when the total concentration Σ is larger than Σ_z. Note that curve A and B are joined only at a single point $(c_1, c_2) = (\frac{1}{2}\Sigma_z, \frac{1}{2}\Sigma_z)$ which is located only at $\Sigma = \Sigma_z$. Moreover, $\phi_A(\Sigma_z) = \phi_B(\Sigma_z) = 0$ and $(c_1 - c_2)(\Sigma_z) = 0$.

Suppose $0 < z \leqslant g$. Then ϕ_A and ϕ_B can be regarded as one variable ϕ and $c_1 - c_2$ may become a strictly monotone increasing function of ϕ. The result is stated as follows:

Proposition 2.1. *Suppose $0 < z \leqslant g$. Then $\Sigma = \Sigma(\phi)$ can be a single-valued function of ϕ with domain being the entire space $\mathbb{R}$ and range $[\Sigma_z, \infty)$ such that $\Sigma(0) = \Sigma_z$,*

$$\begin{cases} \phi_A(\Sigma(\phi)) = \phi & \text{if } \phi \geqslant 0, \\[2mm] \phi_B(\Sigma(\phi)) = \phi & \text{if } \phi \leqslant 0, \end{cases}$$

(2.13)

and $c_1 - c_2 = (c_1 - c_2)(\Sigma(\phi))$ is a strictly monotone increasing function of ϕ from $-\infty$ to ∞.

Proof. Suppose $0 < z \leqslant g$. Then by (2.11) and (2.12), we have

$$\frac{\mathrm{d}}{\mathrm{d}\Sigma}\phi_A(\Sigma) > 0 \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}\Sigma}\phi_B(\Sigma) < 0 \quad \text{for } \Sigma \geqslant \Sigma_z.$$

(2.14)

Here we have used $0 < z \leqslant g$. Thus $\phi_A(\Sigma) > 0$ and $\phi_B(\Sigma) < 0$ for $\Sigma > \Sigma_z$. Besides, the range of ϕ_A is $[0, \infty)$ and the range of ϕ_B is $(-\infty, 0]$. Note that $\phi_A(\Sigma_z) = \phi_B(\Sigma_z) = 0$. We may combine ϕ_A and ϕ_B as one variable ϕ (see figure 1) defined as follows:

$$\begin{cases} \phi = \phi_A(\Sigma) \geqslant 0 & \text{on } A, \\[2mm] \phi = \phi_B(\Sigma) \leqslant 0 & \text{on } B. \end{cases}$$
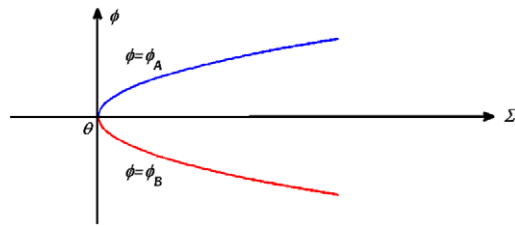
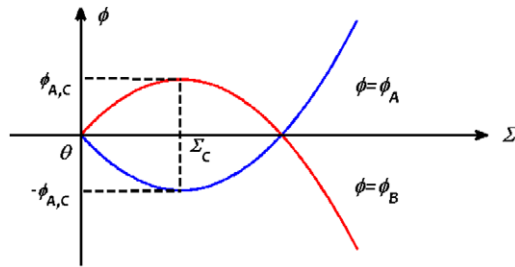**Figure 1.** $\theta = (\Sigma_z, 0)$ in (Σ, ϕ) coordinates.

Figure 2. $\theta = (\Sigma_z, 0)$ in (Σ, ϕ) coordinates.

Hence by (2.14) and inverse function theorem, Σ can be denoted as $\Sigma = \Sigma(\phi)$ and become a single-valued function of ϕ with domain being the entire space $\mathbb{R}$ and range $[\Sigma_z, \infty)$ such that $\Sigma(0) = \Sigma_z$ and (2.13) hold true. The derivative of Σ with respect to ϕ is

$$\frac{d\Sigma}{d\phi} = \frac{1}{\frac{d\phi}{d\Sigma}} = \begin{cases} q \, \frac{e^{(g+z)\Sigma} \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}}{(1+g\Sigma)e^{(g+z)\Sigma} + g^2 - z^2} & \text{if} \quad \phi \geqslant 0, \\[4mm] -q \, \frac{e^{(g+z)\Sigma} \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}}}{(1+g\Sigma)e^{(g+z)\Sigma} + g^2 - z^2} & \text{if} \quad \phi \leqslant 0. \end{cases} \tag{2.15}$$

Moreover, $c_1 - c_2 = (c_1 - c_2)(\Sigma(\phi))$ is also a function of ϕ. Note that $\Sigma(0) = \Sigma_z$, $\Sigma'(0) = 0$ and $(c_1 - c_2)(\Sigma(0)) = (c_1 - c_2)(\Sigma_z) = 0$. Then (2.8) and (2.15) imply

$$\frac{d}{d\phi}(c_1 - c_2) = \frac{d}{d\Sigma}(c_1 - c_2)\frac{d\Sigma}{d\phi} = q \, \frac{\Sigma e^{(g+z)\Sigma} + 2(g+z)}{(1+g\Sigma)e^{(g+z)\Sigma} + g^2 - z^2} > 0 \quad \text{for } \phi \in \mathbb{R}.$$

Therefore, $c_1 - c_2$ is strictly monotone increasing to ϕ and we complete the proof. $\qquad\square$

When $z = g_{12}$ is increased, for example when the ion is divalent like calcium, the profiles of ϕ_A and ϕ_B may lose monotonicity and become oscillatory. It is well known in experiments that calcium has profound and complex effects on the current voltage relations of channels (see [2, 20]). Suppose $z > \sqrt{1 + g^2} > 0$. Then $z^2 - g^2 > 1$ and there exists a unique $\Sigma_c > 0$ (because $(1 + g\Sigma)e^{(g+z)\Sigma}$ is strictly monotone increasing to $\Sigma > 0$) depending on Σ_z such that

$$(1 + g\Sigma_c)e^{(g+z)\Sigma_c} + g^2 - z^2 = 0.$$

Note that $\frac{d\phi_A}{d\Sigma}(\Sigma_c) = \frac{d\phi_B}{d\Sigma}(\Sigma_c) = 0$ if $\Sigma_c > \Sigma_z > 0$. We shall prove that Σ_c may be located in the domain of ϕ_A and ϕ_B i.e. $\Sigma_c > \Sigma_z > 0$ if z is sufficiently large (see proposition 2.2). By (2.11) and (2.12), $\frac{d\phi_A}{d\Sigma} < 0$ on (Σ_z, Σ_c), $\frac{d\phi_A}{d\Sigma} > 0$ on (Σ_c, ∞), $\frac{d\phi_B}{d\Sigma} > 0$ on (Σ_z, Σ_c), $\frac{d\phi_B}{d\Sigma} < 0$ on (Σ_c, ∞). Then Σ_c is a unique (global) minimal point of ϕ_A and a unique (global) maximal point of ϕ_B, respectively (see figure 2).
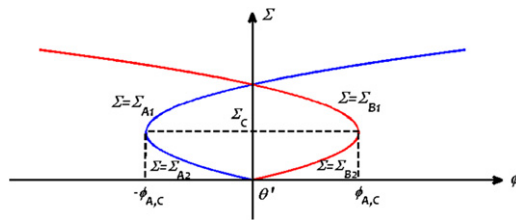
Figure 3. $\theta' = (0, \Sigma_z)$.

Moreover, by (2.10),

$$\phi_{A,c} \equiv -\phi_A(\Sigma_c) = -\min_{\Sigma > \Sigma_z} \phi_A(\Sigma) = \max_{\Sigma > \Sigma_z} \phi_B(\Sigma) = \phi_B(\Sigma_c) > 0. \qquad (2.16)$$

By figure 2, the inverse image of function ϕ_A consists of two functions $\Sigma_{A_1} : (-\phi_{A,c}, \infty) \to (\Sigma_c, \infty)$ and $\Sigma_{A_2} : [-\phi_{A,c}, 0] \to [\Sigma_z, \Sigma_c]$ such that $\frac{d\Sigma_{A_1}}{d\phi} > 0$ on $(-\phi_{A,c}, \infty)$ and $\frac{d\Sigma_{A_2}}{d\phi} < 0$ on $(-\phi_{A,c}, 0)$ (see figure 3).

Moreover, by (2.11),

$$\frac{d\Sigma_{A_1}}{d\phi} = q \, \frac{e^{(g+z)\Sigma_{A_1}} \sqrt{\Sigma_{A_1}^2 - 4e^{-(g+z)\Sigma_{A_1}}}}{(1 + g\,\Sigma_{A_1})e^{(g+z)\Sigma_{A_1}} + g^2 - z^2} > 0 \quad \text{for} \quad \phi > -\phi_{A,c}, \qquad (2.17)$$

and

$$\frac{d\Sigma_{A_2}}{d\phi} = q \, \frac{e^{(g+z)\Sigma_{A_2}} \sqrt{\Sigma_{A_2}^2 - 4e^{-(g+z)\Sigma_{A_2}}}}{(1 + g\,\Sigma_{A_2})e^{(g+z)\Sigma_{A_2}} + g^2 - z^2} < 0 \quad \text{for} \quad -\phi_{A,c} < \phi < 0.$$

Similarly, the inverse image of function ϕ_B consists of another two functions $\Sigma_{B_1} : (-\infty, \phi_{A,c}) \to (\Sigma_c, \infty)$ and $\Sigma_{B_2} : [0, \phi_{A,c}] \to [\Sigma_z, \Sigma_c]$ such that $\frac{d\Sigma_{B_1}}{d\phi} < 0$ on $(-\infty, \phi_{A,c})$ and $\frac{d\Sigma_{B_2}}{d\phi} > 0$ on $(0, \phi_{A,c})$. Moreover, by (2.12),

$$\frac{d\Sigma_{B_1}}{d\phi} = -q \, \frac{e^{(g+z)\Sigma_{B_1}} \sqrt{\Sigma_{B_1}^2 - 4e^{-(g+z)\Sigma_{B_1}}}}{(1 + g\,\Sigma_{B_1})e^{(g+z)\Sigma_{B_1}} + g^2 - z^2} < 0 \quad \text{for} \quad \phi < \phi_{A,c}, \qquad (2.18)$$

and

$$\frac{d\Sigma_{B_2}}{d\phi} = -q \, \frac{e^{(g+z)\Sigma_{B_2}} \sqrt{\Sigma_{B_2}^2 - 4e^{-(g+z)\Sigma_{B_2}}}}{(1 + g\,\Sigma_{B_2})e^{(g+z)\Sigma_{B_2}} + g^2 - z^2} > 0 \quad \text{for} \quad 0 < \phi < \phi_{A,c}.$$

Thus by (2.8), we may consider two functions of $(c_1 - c_2) \circ \Sigma_{A_1}$ and $(c_1 - c_2) \circ \Sigma_{B_1}$ as follows:

$$(c_1 - c_2)(\Sigma_{A_1}(\phi)) = \sqrt{\Sigma_{A_1}^2 - 4e^{-(g+z)\Sigma_{A_1}}} \quad \text{for} \quad \phi \geqslant -\phi_{A,c}, \qquad (2.19)$$

and

$$(c_1 - c_2)(\Sigma_{B_1}(\phi)) = -\sqrt{\Sigma_{B_1}^2 - 4e^{-(g+z)\Sigma_{B_1}}} \quad \text{for} \quad \phi \leqslant \phi_{A,c}. \qquad (2.20)$$

Note that $(c_1 - c_2)(\Sigma_{A_1}(\cdot))$ and $(c_1 - c_2)(\Sigma_{B_1}(\cdot))$ are continuous functions on $[-\phi_{A,c}, \phi_{A,c}]$. Moreover, by (2.17)–(2.20), we have

$$\frac{d}{d\phi}(c_1 - c_2)(\Sigma_{A_1}(\phi)) = q \, \frac{e^{(g+z)\Sigma_{A_1}}[\Sigma_{A_1} + 2(g + z)e^{-(g+z)\Sigma_{A_1}}]}{(1 + g\,\Sigma_{A_1})e^{(g+z)\Sigma_{A_1}} + g^2 - z^2} > 0 \quad \text{for} \quad \phi > -\phi_{A,c},$$

$$(2.21)$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\phi}(c_1 - c_2)(\Sigma_{B_1}(\phi)) = q \, \frac{\mathrm{e}^{(g+z)\Sigma_{B_1}}[\Sigma_{B_1} + 2(g+z)\mathrm{e}^{-(g+z)\Sigma_{B_1}}]}{(1 + g\Sigma_{B_1})\mathrm{e}^{(g+z)\Sigma_{B_1}} + g^2 - z^2} > 0 \quad \text{for} \quad \phi < \phi_{A,c}.$$

(2.22)

Here we have used (2.38) and (2.39). Consequently, $(c_1 - c_2)(\Sigma_{A_1}(\cdot))$ and $(c_1 - c_2)(\Sigma_{B_1}(\cdot))$ are smooth functions on $(-\phi_{A,c}, \phi_{A,c})$. Since $(c_1 - c_2)(\Sigma_{A_1}(\cdot))$ and $(c_1 - c_2)(\Sigma_{B_1}(\cdot))$ are strictly monotone increasing to ϕ (see (2.21) and (2.22)), then we may use (2.8) to get

$$(c_1 - c_2)(\Sigma_{A_1}(\phi)) \geqslant (c_1 - c_2)(\Sigma_{A_1}(-\phi_{A,c})) = \sqrt{\Sigma_c^2 - 4\mathrm{e}^{-(g+z)\Sigma_c}} > 0,$$

(2.23)

$$(c_1 - c_2)(\Sigma_{B_1}(\phi)) \leqslant (c_1 - c_2)(\Sigma_{B_1}(\phi_{A,c})) = -\sqrt{\Sigma_c^2 - 4\mathrm{e}^{-(g+z)\Sigma_c}} < 0,$$

(2.24)

for $\phi \in (-\phi_{A,c}, \phi_{A,c})$.

Now we claim that if z is sufficiently large, then $\Sigma_c > \Sigma_z > 0$ i.e. Σ_c is located in the domain of ϕ_A and ϕ_B as follows:

Proposition 2.2. *Let*

$$g_c = \inf \Big\{ z > \sqrt{1 + g^2} : \text{there exists } \Sigma_{c,z} > \Sigma_z > 0 \text{ such that } (1 + g\Sigma_{c,z})\mathrm{e}^{(g+z)\Sigma_{c,z}}$$
$$+ g^2 - z^2 = 0 \Big\},$$

(2.25)

where $\Sigma_z > 0$ is the unique solution of $\Sigma = 2\mathrm{e}^{-\frac{1}{2}(g+z)\Sigma}$ for $z > 0$. Then for $z > g_c$, there exists a unique $\Sigma_c = \Sigma_{c,z} > \Sigma_z$ depending on z such that $(1 + g\Sigma_c)\mathrm{e}^{(g+z)\Sigma_c} + g^2 - z^2 = 0$. Conversely, for $0 < z < g_c$, no such Σ_c exists and $(1 + g\Sigma)\mathrm{e}^{(g+z)\Sigma} + g^2 - z^2 > 0$ for $\Sigma \geqslant \Sigma_z > 0$.

Proof. Firstly, we claim that g_c is well-defined. For any $z > 0$, we may define a function $f_z = f_z(\Sigma)$ by

$$f_z(\Sigma) = (1 + g\Sigma)\mathrm{e}^{(g+z)\Sigma} + g^2 - z^2 \quad \text{for} \quad \Sigma > 0.$$

(2.26)

Then it is obvious that $f_z(+\infty) = \infty$,

$$f_z'(\Sigma) = [g + (1 + g\Sigma)(g + z)]\mathrm{e}^{(g+z)\Sigma} > 0 \quad \text{for} \quad \Sigma, z > 0,$$

(2.27)

and $f_z(0) = 1 + g^2 - z^2 < 0$ if $z > \sqrt{1 + g^2}$. Hence there exists a unique $\Sigma_{c,z} > 0$ such that $f_z(\Sigma_{c,z}) = 0$. Let $\Sigma{///}_z > 0$ be the unique solution of

$$\Sigma_z = 2\mathrm{e}^{-\frac{1}{2}(g+z)\Sigma_z} \quad \text{for} \quad z > 0.$$

(2.28)

Now we prove $\Sigma_{c,z} > \Sigma_z$ as z sufficiently large. By (2.28), Σ_z is decreasing to z (differentiate (2.28) to z) and $z = -\left(g + \frac{2\ln\Sigma_z - \ln 4}{\Sigma_z}\right)$. Thus $\Sigma_z \to 0$ as $z \to \infty$ and

$$f_z(\Sigma_z) = (1 + g\Sigma_z)\mathrm{e}^{(g+z)\Sigma_z} + g^2 - z^2$$
$$= [4(1 + g\Sigma_z) + (g^2 - z^2)\Sigma_z^2]/\Sigma_z^2 \quad \text{by (2.28)}$$
$$= [4(1 + g\Sigma_z) - 2g\Sigma_z(2\ln\Sigma_z - \ln 4) - (2\ln\Sigma_z - \ln 4)^2]/\Sigma_z^2 \to -\infty \quad \text{as } z \to \infty,$$

and then $f_z(\Sigma_z) < 0$ as z sufficiently large. Since $f_z(\Sigma_{c,z}) = 0$ and $f_z(\Sigma_z) < 0$ as z sufficiently large, then by (2.27), we have $\Sigma_{c,z} > \Sigma_z$ as z sufficiently large. Consequently, the set

$$\mathcal{Z} = \{z > \sqrt{1 + g^2} : \exists \Sigma_{c,z} > \Sigma_z > 0 \text{ such that } f_z(\Sigma_{c,z}) = 0\}$$

(2.29)

$$= \{z > \sqrt{1 + g^2} : f_z(\Sigma_z) < 0\}$$

is nonempty and the value $g_c = \inf_{z \in \mathcal{Z}} z$ (defined in (2.25)) is well-defined. Note that the existence of $\Sigma_{c,z}$ with $f_z(\Sigma_{c,z}) = 0$ is guaranteed due to $z > \sqrt{1 + g^2}$, so (2.27) implies $\Sigma_{c,z} > \Sigma_z$ if $f_z(\Sigma_z) < 0$ holds true.

   To complete the proof of proposition 2.2, we need the following result:

Claim 1.   Suppose $f_{z_0}(\Sigma_{z_0}) = 0$ and $\Sigma_{z_0} > 0$ for some $z_0 > \sqrt{1 + g^2}$. Then there exist $z_l, z_r > \sqrt{1 + g^2}$ and $z_l < z_0 < z_r$ such that $f_z(\Sigma_z) > 0$ for $z \in (z_l, z_0)$ and $f_z(\Sigma_z) < 0$ for $z \in (z_0, z_r)$.

Proof. By (2.26) and (2.28),

$$f(\Sigma_z) = \frac{4(1 + g\Sigma_z)}{\Sigma_z^2} + g^2 - z^2. \tag{2.30}$$

Then $f_{z_0}(\Sigma_{z_0}) = 0$ gives

$$4\frac{1 + g\Sigma_{z_0}}{\Sigma_{z_0}^2} = z_0^2 - g^2,$$

and Σ_{z_0} satisfies $(z_0^2 - g^2)\Sigma_{z_0}^2 - 4g\Sigma_{z_0} - 4 = 0$ having solutions as $\Sigma_{z_0} = \frac{2}{z_0 - g}$ and $\Sigma_{z_0} = -\frac{2}{z_0 + g}$. Hence due to $\Sigma_{z_0} > 0$,

$$\Sigma_{z_0} = \frac{2}{z_0 - g}. \tag{2.31}$$

Note that $z_0 > \sqrt{1 + g^2} > \pm g$. Differentiating (2.28) and (2.30) to z, we have

$$\frac{\mathrm{d}}{\mathrm{d}z} f_z(\Sigma_z) = -4\frac{2 + g\Sigma_z}{\Sigma_z^3}\frac{\mathrm{d}\Sigma_z}{\mathrm{d}z} - 2z,$$

$$\frac{\mathrm{d}\Sigma_z}{\mathrm{d}z} = \frac{-\Sigma_z^2}{(g + z)\Sigma_z + 2}.$$

Thus by (2.31), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}z} f_z(\Sigma_z)|_{z=z_0} = -z_0 - g < 0. \tag{2.32}$$

Therefore, by (2.32), we may complete the proof of claim 1.   □

   It is obvious that

$$f_z(\Sigma) > 0 \quad \text{for} \quad \Sigma > 0 \quad \text{and} \quad 0 < z \leqslant \sqrt{1 + g^2}. \tag{2.33}$$

Now we want to prove that

$$\mathcal{Z} = (g_c, \infty), \tag{2.34}$$

where $g_c = \inf_{z \in \mathcal{Z}} z$. Due to the continuity of f_z, (2.29) implies that the set $\mathcal{Z}$ is open. Suppose the set $\mathcal{Z}$ has two components. Then without loss of generality, we may assume that there exists $z_a > g_c$ such that $\mathcal{Z} = (g_c, z_a) \cup (z_a, \infty)$. Hence $f_{z_a}(\Sigma_{z_a}) = 0$ and $f_z(\Sigma_z) < 0$ for $z \in (g_c, z_a) \cup (z_a, \infty)$. However, claim 1 implies that $f_z(\Sigma_z) > 0$ for $z \in (z_l, z_a)$ which contradicts to $f_z(\Sigma_z) < 0$ for $z \in (g_c, z_a)$. Thus the proof of (2.34) is done. On the other hand, claim 1 also implies that

$$f_z(\Sigma_z) > 0 \quad \text{for} \quad 0 < z < g_c. \tag{2.35}$$

Otherwise, by (2.33), there exists $z_b \in (\sqrt{1 + g^2}, g_c)$ such that $f_{z_b}(\Sigma_{z_b}) = 0$. Then as for (2.32), we have $\frac{\mathrm{d}}{\mathrm{d}z} f_z(\Sigma_z)|_{z=z_b} = -z_b - g < 0$ and hence there exists $z_c \in (z_b, g_c)$ such that $f_{z_c}(\Sigma_{z_c}) < 0$ which contradicts to (2.34). Therefore, by (2.27) and (2.35), we complete the proof of proposition 2.2.   □

Remark 2.3.

(i) The proof of proposition 2.2 shows that $f_z(\Sigma_z) > 0$ for $0 < z < g_c$ and $f_z(\Sigma_z) < 0$ for $z > g_c$ (see (2.34) and (2.35)). Hence by the continuity of f_z, $f_{g_c}(\Sigma_{g_c}) = 0$.

(ii) By (2.27) and (2.35), we have

$$f_z(\Sigma) = (1 + g\Sigma)\mathrm{e}^{(g+z)\Sigma} + g^2 - z^2 > 0 \quad \text{for} \quad \Sigma \geqslant \Sigma_z \quad \text{and} \quad 0 < z < g_c. \qquad (2.36)$$

(iii) By (2.26), $f_z(\Sigma_z) > 0$ as $z = \sqrt{1+g^2}$ but $f_{g_c}(\Sigma_{g_c}) = 0$. Hence remark 2.3 (i) implies $g_c > \sqrt{1+g^2}$.

Suppose $0 < z < g_c$. Then (2.36) gives $f_z(\Sigma) > 0$ for $\Sigma \geqslant \Sigma_z$. Hence by (2.11) and (2.12), $\frac{\mathrm{d}\phi_A}{\mathrm{d}\Sigma} > 0$ and $\frac{\mathrm{d}\phi_B}{\mathrm{d}\Sigma} < 0$ for $\Sigma \geqslant \Sigma_z$ which gives $\phi_A(\Sigma) > \phi_A(\Sigma_z) = 0 = \phi_B(\Sigma_z) > \phi_B(\Sigma)$ for $\Sigma > \Sigma_z$. Thus as for proposition 2.1, $\Sigma = \Sigma(\phi)$ can be a single-valued function of ϕ with domain as the entire space $\mathbb{R}$ and range $[\Sigma_z, \infty)$ such that $\Sigma(0) = \Sigma_z$ and $c_1 - c_2$ is strictly monotone increasing to ϕ. Moreover, $\Sigma \to \infty$ as $\phi \to \pm\infty$ and $(c_1 - c_2)(\Sigma(\phi)) \to \pm\infty$ as $\phi \to \pm\infty$.

Suppose $z > g_c > 0$. Then proposition 2.2 gives that there exists a unique $\Sigma_c \in (\Sigma_z, \infty)$ such that

$$(1 + g\Sigma_c)\mathrm{e}^{(g+z)\Sigma_c} + g^2 - z^2 = 0, \qquad (2.37)$$

which implies

$$(1 + g\Sigma)\mathrm{e}^{(g+z)\Sigma} + g^2 - z^2 > 0, \quad \text{for} \quad \Sigma > \Sigma_c, \qquad (2.38)$$

and

$$(1 + g\Sigma)\mathrm{e}^{(g+z)\Sigma} + g^2 - z^2 < 0, \quad \text{for} \quad \Sigma_z \leqslant \Sigma < \Sigma_c, \qquad (2.39)$$

By (2.9) and (2.11), we have $\frac{\mathrm{d}\phi_A}{\mathrm{d}\Sigma} > 0$ for $\Sigma > \Sigma_c$; $\frac{\mathrm{d}\phi_A}{\mathrm{d}\Sigma} < 0$ for $\Sigma_z < \Sigma < \Sigma_c$, and ϕ_A tends to $+\infty$ as Σ goes to $+\infty$. Hence Σ_c is the unique minimum point of ϕ_A. Since $\Sigma_z^2 = 4\mathrm{e}^{-(g+z)\Sigma_z}$, then $\phi_A(\Sigma_z) = 0$ which implies $-\phi_{A,c} = \phi_A(\Sigma_c) < 0$. Since Σ_c satisfies $(1 + g\Sigma_c)\mathrm{e}^{(g+z)\Sigma_c} = z^2 - g^2$ i.e. $\Sigma_c + \frac{\ln(1+g\Sigma_c)}{g+z} = \frac{\ln(z^2-g^2)}{g+z}$, then Σ_c must tend to zero as z goes to infinity. Note that $g > 0$ is a fixed constant. Consequently, $-\ln\left[\frac{1}{2}\left(\Sigma_c + \sqrt{\Sigma_c^2 - 4\mathrm{e}^{-(g+z)\Sigma_c}}\right)\right] \to +\infty$ as $z \to +\infty$, and then

$$q\,\phi_{A,c} = q\,\phi_A(\Sigma_c)$$
$$= -\ln\left[\frac{1}{2}\left(\Sigma_c + \sqrt{\Sigma_c^2 - 4\mathrm{e}^{-(g+z)\Sigma_c}}\right)\right] + \frac{g+z}{2}\Sigma_c + \frac{g-z}{2}\sqrt{\Sigma_c^2 - 4\mathrm{e}^{-(g+z)\Sigma_c}}$$
$$= -\ln\left[\frac{1}{2}\left(\Sigma_c + \sqrt{\Sigma_c^2 - 4\mathrm{e}^{-(g+z)\Sigma_c}}\right)\right] + \frac{g}{2}\left(\Sigma_c + \sqrt{\Sigma_c^2 - 4\mathrm{e}^{-(g+z)\Sigma_c}}\right)$$
$$\quad + \frac{z}{2}\left(\Sigma_c - \sqrt{\Sigma_c^2 - 4\mathrm{e}^{-(g+z)\Sigma_c}}\right)$$
$$\geqslant -\ln\left[\frac{1}{2}\left(\Sigma_c + \sqrt{\Sigma_c^2 - 4\mathrm{e}^{-(g+z)\Sigma_c}}\right)\right] \to +\infty$$

as $z \to +\infty$. Thus $\phi_{A,c} \to +\infty$ as $z \to +\infty$ and $g > 0$ is fixed. Besides, since $\mathrm{e}^{-(g+z)\Sigma_c} = (1 + g\Sigma_c)/(z^2 - g^2)$ and $\Sigma_c \to 0$ as $z \to \infty$, then by (2.8), we have

$(c_1 - c_2)(\Sigma_c) \to 0$ as $z \to +\infty$ and $g > 0$ is fixed. Therefore, we may summarize the above results as follows:

Theorem 2.4.

(i)   *Suppose $0 < z < g_c$. Then $(c_1 - c_2) \circ \Sigma$ is a monotone increasing function to $\phi \in \mathbb{R}$ satisfying $(c_1 - c_2)(\Sigma(\phi)) \to \pm\infty$ as $\phi \to \pm\infty$, respectively.*

(ii) *Suppose $z > g_c$. Then there are two functions Σ_{A_1} and Σ_{B_1} such that $(c_1 - c_2) \circ \Sigma_{A_1} : [-\phi_{A,c}, \infty) \to \mathbb{R}$ and $(c_1 - c_2) \circ \Sigma_{B_1} : (-\infty, \phi_{A,c}] \to \mathbb{R}$ are monotone increasing functions of ϕ, where $\phi_{A,c}$ satisfies $\phi_{A,c} \to +\infty$ and $(c_1 - c_2)(\Sigma_c) \to 0$ as $z \to +\infty$ and $g > 0$ is fixed. Moreover,*

$$(c_1 - c_2) \circ \Sigma_{A_1}(-\phi_{A,c}) = (c_1 - c_2)(\Sigma_c) > 0,$$
$$(c_1 - c_2) \circ \Sigma_{B_1}(\phi_{A,c}) = -(c_1 - c_2)(\Sigma_c) < 0,$$
$$\lim_{\phi \to \infty} (c_1 - c_2) \circ \Sigma_{A_1}(\phi) = \infty \quad \text{and} \quad \lim_{\phi \to -\infty} (c_1 - c_2) \circ \Sigma_{B_1}(\phi) = -\infty.$$

Here $\circ$ denotes function composition and g_c is the positive constant defined in proposition 2.2.

3. Proof of theorem 1.1 and 1.3

3.1. Proof of theorem 1.1

In this section, we study multiple solutions of the system of equations (1.13)–(1.15) with $N = 3$ and the following assumptions:

$$\rho_0 > 0, \quad g_{i3} = g_{3i} = 0 \quad \text{for} \quad i = 1, 2, 3. \tag{3.1}$$

Then we may get solutions of (1.13) by solving

$$\ln c_i + z_i \phi + \sum_{j=1}^{2} g_{ij} c_j = 0 \quad \text{for} \quad i = 1, 2, \tag{3.2}$$

and let

$$c_3 = \mathrm{e}^{-z_3 \phi}. \tag{3.3}$$

Note that (3.2) is same as (1.13) with $N = 2$. Assume

$$z_2 = -z_1 = q \geqslant 1, \quad g_{11} = g_{22} > 0 \quad \text{and} \quad g_{12} > g_c > 0, \tag{3.4}$$

where $g_c > 0$ is a sufficiently large constant defined in proposition 2.2. We shall use (3.4) and set $\rho_0 > 0$ in order to apply theorem 2.4 (ii) (in section 2) and lemma 4.1 (in section 4) for the proof of theorem 1.1 which gives multiple solutions of (1.13)–(1.15) with $N = 3$ and $\rho_0 > 0$.

By theorem 2.4 (ii), equation (3.2) has multiple solutions

$$(c_1, c_2) = \left(c_1\left(\Sigma_{A_1}(\phi)\right), c_2\left(\Sigma_{A_1}(\phi)\right)\right) \quad \text{and} \quad (c_1, c_2) = \left(c_1\left(\Sigma_{B_1}(\phi)\right), c_2\left(\Sigma_{B_1}(\phi)\right)\right) \tag{3.5}$$

such that $f_{A_1}(\phi) = q(c_1 - c_2)\left(\Sigma_{A_1}(\phi)\right)$ and $f_{B_1}(\phi) = q(c_1 - c_2)\left(\Sigma_{B_1}(\phi)\right)$ are monotone increasing to ϕ but the values of f_{A_1} and f_{B_1} are away from zero (see figure 4). By lemma 4.5, it is impossible to get uniformly bounded solution by solving either $\varepsilon \phi''(x) = f_{A_1}(\phi(x))$ or $\varepsilon \phi''(x) = f_{B_1}(\phi(x))$ for $x \in (-1, 1)$. This motivates us to develop lemma 4.1 (in section 4), and use (3.3) to transform (1.14) into the following equations:

$$\varepsilon \phi''(x) = f_A(\phi(x)) \quad \text{for} \quad x \in (-1, 1), \tag{3.6}$$
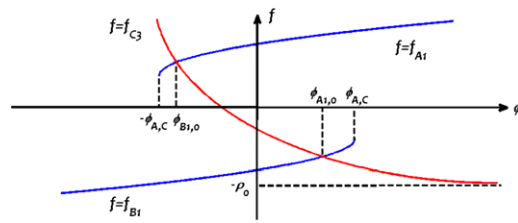
Figure 4. Figures of f_{A_1}, f_{B_1} and f_{c_3}.

and

$$\varepsilon\phi''(x) = f_B(\phi(x)) \quad \text{for} \quad x \in (-1, 1), \tag{3.7}$$

where

$$f_A(\phi) = q(c_1 - c_2)\left(\Sigma_{A_1}(\phi)\right) - z_3 e^{-z_3\phi} + \rho_0,$$

and

$$f_B(\phi) = q(c_1 - c_2)\left(\Sigma_{B_1}(\phi)\right) - z_3 e^{-z_3\phi} + \rho_0.$$

We may denote f_A and f_B as follows: $f_A(\phi) = f_{A_1}(\phi) - f_{c_3}(\phi)$ and $f_B(\phi) = f_{B_1}(\phi) - f_{c_3}(\phi)$, where $f_{A_1}(\phi) = q(c_1 - c_2)\left(\Sigma_{A_1}(\phi)\right)$, $f_{B_1}(\phi) = q(c_1 - c_2)\left(\Sigma_{B_1}(\phi)\right)$, and $f_{c_3}(\phi) = z_3 e^{-z_3\phi} - \rho_0$.

Let $\rho_0 > 0$. Then theorem 2.4 (ii) (in section 2) implies that as $g_{12} = z \geqslant g_{\rho_0} > g_c > 0$ (g_{ρ_0} is a large constant depending on ρ_0), both functions f_{A_1} and f_{B_1} intersect with the function f_{c_3} at $\phi_{A_1,0}$ and $\phi_{B_1,0}$, respectively (see figure 4). Note that the assumption $\rho_0 > 0$ is necessary for the existence of $\phi_{A_1,0}$ and $\phi_{B_1,0}$. Moreover, $f_A = f_{A_1} - f_{c_3}$ and $f_B = f_{B_1} - f_{c_3}$ satisfy

(1) $f_A : [-\phi_{A,c}, \infty) \to \mathbb{R}$ is smooth and strictly monotone increasing, $-\phi_{A,c} < 0$, $f_A(-\phi_{A,c}) < 0$, $f_A(\infty) > 0$ and $f_A(\phi_{A_1,0}) = 0$ for some $\phi_{A_1,0} > -\phi_{A,c}$.
(2) $f_B : (-\infty, \phi_{A,c}] \to \mathbb{R}$ is smooth and strictly monotone increasing, $\phi_{A,c} > 0$, $f_B(\phi_{A,c}) > 0$, $f_B(-\infty) < 0$ and $f_B(\phi_{B_1,0}) = 0$ for some $\phi_{B_1,0} < \phi_{A,c}$.

Hence by lemma 4.1, we may get uniformly bounded solutions ϕ_ε^A and ϕ_ε^B of (3.6) and (3.7), respectively. Moreover, $\phi_\varepsilon^A(x) \to \phi_{A_1,0}$ and $\phi_\varepsilon^B(x) \to \phi_{B_1,0}$ for $x \in (-1, 1)$ as $\varepsilon \to 0$. Therefore, we complete the proof of theorem 1.1.

3.2. Proof of theorem 1.3

Let $N = 4$, $z_2 = -z_1 = q_1 \geqslant 1$ and $z_4 = -z_3 = q_2 \geqslant 1$. Assume $g_{11} = g_{22} = g > 0$, $g_{33} = g_{44} = \tilde{g} > 0$ and $g_{ij} = g_{ji} = 0$ for $i = 1, 2$ and $j = 3, 4$. Then (1.13) may be represented as

$$\ln c_i + z_i\phi + \sum_{j=1}^{2} g_{ij}c_j = 0 \quad \text{for } i = 1, 2, \tag{3.8}$$

and

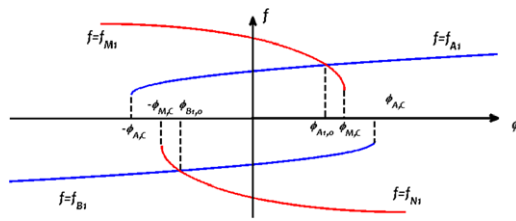$$\ln c_i + z_i\phi + \sum_{j=3}^{4} g_{ij}c_j = 0 \quad \text{for } i = 3, 4. \tag{3.9}$$

Figure 5. Figures of f_{A_1}, f_{B_1}, f_{M_1} and f_{N_1}.

Note that both (3.8) and (3.9) have the same form as (3.2) with (3.4) which can be solved explicitly. As for theorem 2.4 in section 2, both (3.8) and (3.9) have two branches of solutions, respectively. We may denote these solutions as follows:

$$(c_1, c_2) = \left(c_1 \left(\Sigma_{A_1}(\phi)\right), c_2 \left(\Sigma_{A_1}(\phi)\right)\right),$$
$$(c_1, c_2) = \left(c_1 \left(\Sigma_{B_1}(\phi)\right), c_2 \left(\Sigma_{B_1}(\phi)\right)\right),$$
$$(c_3, c_4) = \left(c_3 \left(\Sigma_{M_1}(\phi)\right), c_4 \left(\Sigma_{M_1}(\phi)\right)\right),$$
$$(c_3, c_4) = \left(c_3 \left(\Sigma_{N_1}(\phi)\right), c_4 \left(\Sigma_{N_1}(\phi)\right)\right),$$

such that $(c_1 - c_2) \circ \Sigma_{A_1} : [-\phi_{A,c}, \infty) \to \mathbb{R}$, $(c_1 - c_2) \circ \Sigma_{B_1} : (-\infty, \phi_{A,c}] \to \mathbb{R}$, $(c_3 - c_4) \circ \Sigma_{N_1} : [-\phi_{M,c}, \infty) \to \mathbb{R}$ and $(c_3 - c_4) \circ \Sigma_{M_1} : (-\infty, \phi_{M,c}] \to \mathbb{R}$, are monotone increasing functions of ϕ, where $\phi_{A,c}, \phi_{M,c} > 0$ are constants, Σ_{A_1}, Σ_{B_1}, Σ_{M_1} and Σ_{N_1} are functions satisfying

$$(c_1 - c_2) \circ \Sigma_{A_1}(-\phi_{A,c}), (c_3 - c_4) \circ \Sigma_{N_1}(-\phi_{M,c}) > 0,$$
$$(c_1 - c_2) \circ \Sigma_{B_1}(\phi_{A,c}), (c_3 - c_4) \circ \Sigma_{M_1}(\phi_{M,c}) < 0,$$
$$\lim_{\phi \to \infty} (c_1 - c_2) \circ \Sigma_{A_1}(\phi) = \lim_{\phi \to \infty} (c_3 - c_4) \circ \Sigma_{N_1}(\phi) = \infty,$$
$$\lim_{\phi \to -\infty} (c_1 - c_2) \circ \Sigma_{B_1}(\phi) = \lim_{\phi \to -\infty} (c_3 - c_4) \circ \Sigma_{M_1}(\phi) = -\infty.$$

Here $\circ$ denotes function composition. Moreover, theorem 2.4 gives $\phi_{A,c}, \phi_{M,c} \to +\infty$ and $(c_1 - c_2) \circ \Sigma_{A_1}(-\phi_{A,c})$, $(c_1 - c_2) \circ \Sigma_{B_1}(\phi_{A,c})$, $(c_3 - c_4) \circ \Sigma_{M_1}(\phi_{M,c})$ and $(c_3 - c_4) \circ \Sigma_{N_1}(-\phi_{M,c})$ tend to zero as $z, \tilde{z} \to +\infty$ and $g, \tilde{g} > 0$ are fixed.

Without loss of generality, we may assume $\phi_{M,c} < \phi_{A,c}$. Fix $\rho_0 \in \mathbb{R}$ arbitrarily. Then as for (3.2), we may solve (3.8) and get functions $f_{A_1}(\phi) = q_1(c_1 - c_2)\left(\Sigma_{A_1}(\phi)\right) - \rho_0$ and $f_{B_1}(\phi) = q_1(c_1 - c_2)\left(\Sigma_{B_1}(\phi)\right) - \rho_0$ which are sketched in figure 5 (up to a shift by ρ_0), provided that $g_{12} = g_{21} = z > 0$ is sufficiently large. Similarly, we may solve (3.9) and get functions $f_{M_1}(\phi) = q_2(c_4 - c_3)\left(\Sigma_{M_1}(\phi)\right)$ and $f_{N_1}(\phi) = q_2(c_4 - c_3)\left(\Sigma_{N_1}(\phi)\right)$ as $g_{34} = g_{43} = \tilde{z} > 0$ sufficiently large (see figure 5).

Because function $q_2(c_3 - c_4) \circ \Sigma_{M_1}$ is negative and increasing to ϕ, function $q_2(c_4 - c_3) \circ \Sigma_{M_1}$ becomes positive and decreasing to ϕ. On the other hand, function $q_1(c_1 - c_2) \circ \Sigma_{A_1} - \rho_0$ is positive and increasing to ϕ. This implies that as z and $\tilde{z}$ sufficiently large, functions $q_1(c_1 - c_2) \circ \Sigma_{A_1} - \rho_0$ and $q_2(c_3 - c_4) \circ \Sigma_{M_1}$ may intersect at $\phi = \phi_{A_1,0}$. Similarly, functions $q_1(c_1 - c_2) \circ \Sigma_{B_1} - \rho_0$ and $q_2(c_4 - c_3) \circ \Sigma_{N_1}$ may intersect at $\phi = \phi_{B_1,0}$ as z and $\tilde{z}$ sufficiently large. Generically, values $\phi_{A_1,0}$ and $\phi_{B_1,0}$ can be different by choosing z and $\tilde{z}$ suitably e.g. z and $\tilde{z}$ sufficiently large.

Let $f_A = f_{A_1} - f_{M_1}$ and $f_B = f_{B_1} - f_{N_1}$. Then

$$f_A(\phi) = q_1(c_1 - c_2)\left(\Sigma_{A_1}(\phi)\right) + q_2(c_3 - c_4)\left(\Sigma_{M_1}(\phi)\right) - \rho_0,$$

and

$$f_B(\phi) = q_1(c_1 - c_2)\left(\Sigma_{B_1}(\phi)\right) + q_2(c_3 - c_4)\left(\Sigma_{N_1}(\phi)\right) - \rho_0,$$

satisfy

(1) $f_A : [-\phi_{A,c}, \phi_{M,c}] \to \mathbb{R}$ is smooth and strictly monotone increasing, $-\phi_{A,c} < 0$, $f_A(-\phi_{A,c}) < 0$, $f_A(\phi_{M,c}) > 0$ and $f_A(\phi_{A_1,0}) = 0$ for some $\phi_{A_1,0} > -\phi_{A,c}$.

(2) $f_B : [-\phi_{M,c}, \phi_{A,c}] \to \mathbb{R}$ is smooth and strictly monotone increasing, $\phi_{A,c} > 0$, $f_B(\phi_{A,c}) > 0$, $f_B(-\phi_{M,c}) < 0$ and $f_B(\phi_{B_1,0}) = 0$ for some $\phi_{B_1,0} < \phi_{A,c}$.

Moreover, equation (1.14) can be expressed as $\varepsilon\phi_{xx} = f_A(\phi)$ and $\varepsilon\phi_{xx} = f_B(\phi)$ for $x \in (-1, 1)$ which have the same forms as equations (3.6) and (3.7), respectively. Therefore by lemma 4.1, we may complete the proof of theorem 1.3.

4. Uniformly bounded solutions

In this section, we consider the equation

$$\varepsilon\phi''(x) = f(\phi(x)) \quad \text{for} \quad x \in (-1, 1), \tag{4.1}$$

with the Robin boundary condition

$$\phi(1) + \eta_\varepsilon\phi'(1) = \phi_0(1) \quad \text{and} \quad \phi(-1) - \eta_\varepsilon\phi'(-1) = \phi_0(-1), \tag{4.2}$$

where $\phi_0(1), \phi_0(-1)$ are constants and η_ε is a non-negative constant. Note that the solution ϕ_ε of (4.1)–(4.2) may depend on the parameter ε. For notational convenience, we omit ε and denote ϕ as the solution of (4.1)–(4.2). To get uniform boundedness of ϕ, we assume the function f satisfies one of the following conditions:

(F1) $f : [A, M] \to \mathbb{R}$ is smooth and strictly monotone increasing, $A < 0$, $f(A) < 0$, $0 < M \leqslant \infty$, $f(M) > 0$ and $f(\phi_A) = 0$ for some $A < \phi_A < M$.

(F2) $f : [-M, B] \to \mathbb{R}$ is is smooth and strictly monotone increasing, $B > 0$, $f(B) > 0$, $0 < M \leqslant \infty$, $f(-M) < 0$ and $f(\phi_B) = 0$ for some $-M < \phi_B < B$.

Then we have

Lemma 4.1. *Assume the function f satisfies either* (F1) *or* (F2)*, and the constants $A \leqslant \phi_0(-1), \phi_0(1) \leqslant M$ as* (F1) *holds, and $-M \leqslant \phi_0(-1), \phi_0(1) \leqslant B$ as (F2) holds. Let $c = \phi_A$ if* (F1) *holds, and $c = \phi_B$ if* (F2) *holds. Let ϕ be a nonconstant solution of (4.1) with the Robin boundary condition (4.2). Then*

(i)   *If $\phi_0(1), \phi_0(-1) > c$, then there exists $x_1 \in (-1, 1)$ such that $\phi'(x_1) = 0$, $\phi(x_1) > c$, and ϕ is strictly monotone decreasing in $(-1, x_1)$ and increasing in $(x_1, 1)$.*

(ii) *If $\phi_0(1), \phi_0(-1) < c$, then there exists $x_2 \in (-1, 1)$ such that $\phi'(x_2) = 0$, $\phi(x_2) < c$, and ϕ is strictly monotone increasing in $(-1, x_2)$ and decreasing in $(x_2, 1)$.*

(iii) *If $\phi_0(1) \geqslant c \geqslant \phi_0(-1)$, then ϕ is monotone increasing in $(-1, 1)$.*

(iv) *If $\phi_0(1) \leqslant c \leqslant \phi_0(-1)$, then ϕ is monotone decreasing in $(-1, 1)$.*

(v) $\min\{\phi_0(-1), \phi_0(1), 0\} \leqslant \phi(x) \leqslant \max\{\phi_0(-1), \phi_0(1), 0\}$ *for $x \in (-1, 1)$.*

(vi) $\phi(x) \to c$ *as $\varepsilon \to 0+$, where $c = \phi_A$ if* (F1) *holds, and $c = \phi_B$ if* (F2) *holds.*

Proof. Without loss of generality, we may assume the function f satisfying (F1). Replacing ϕ by $\phi + c$, we may assume $c = 0$ and $f(0) = 0$ in the whole proof for notational convenience. Since the domain of the function f is only $[A, M]$, then we firstly extend it smoothly to the entire real line $\mathbb{R}$ in order to use the standard direct method to get the existence of solution ϕ. Hence we may temporarily assume the function f as a smooth and strictly monotone increasing function on $\mathbb{R}$. Actually, such an assumption can be ignored because of (4.5).

To prove lemma 4.1, we need the following proposition:

Proposition 4.2.

(a) *If $x_a \in (-1, 1)$ is a local minimum point of ϕ, then $\phi(x_a) > 0$, ϕ is monotone decreasing in $(-1, x_a)$ and increasing in $(x_a, 1)$.*

(b) *If $x_b \in (-1, 1)$ is a local maximum point of ϕ, then $\phi(x_b) < 0$, ϕ is monotone increasing in $(-1, x_b)$ and decreasing in $(x_b, 1)$.*

The proof of proposition 4.2 (b) is quite similar to that of proposition 4.2 (a) so we only state the proof of proposition 4.2 (a) as follows: Suppose $x_a \in (-1, 1)$ is a local minimum point of ϕ. Then $\phi'(x_a) = 0$ and $\phi''(x_a) \geqslant 0$. If $\phi''(x_a) = 0$, then the equation $\varepsilon\phi'' = f(\phi)$ gives $f(\phi(x_a)) = \varepsilon\phi''(x_a) = 0$ which implies $\phi(x_a) = 0$ and then by the uniqueness of ordinary differential equations and $\phi(x_a) = \phi'(x_a) = 0$, we have $\phi \equiv 0$ which contradicts to ϕ is nonconstant. Hence $\phi''(x_a) > 0$ and $f(\phi(x_a)) = \varepsilon\phi''(x_a) > 0$ i.e. $\phi(x_a) > 0$. Now we prove that ϕ is decreasing in $(-1, x_a)$ and increasing in $(x_a, 1)$. Suppose not. Then there exists $x_c \in (-1, 1)$ and $x_c \neq x_a$ such that x_c is a local maximum point of ϕ i.e. $\phi'(x_c) = 0, \phi''(x_c) \leqslant 0$ and $\phi(x_c) > \phi(x_1) > 0$ but $\varepsilon\phi''(x_c) = f(\phi(x_c)) > 0$ which contradicts to $\phi''(x_c) \leqslant 0$. Therefore, we may complete the proof of proposition 4.2.

For the proof lemma 4.1 (i), we need

Claim 1. Assume $\phi_0(1), \phi_0(-1) > 0$. Then $\phi(-1), \phi(1) > 0, \phi'(-1) < 0$ and $\phi'(1) > 0$. We may prove claim 1 by contradiction. Suppose one of the following cases holds:

Case I. $\phi(-1) > 0$ and $\phi'(-1) \geqslant 0$.

Case II. $\phi(-1) \leqslant 0$.

For the Case I, we may use $\phi(-1) > 0$ and the continuity of ϕ to obtain that as $x \in (-1, 1)$ sufficiently close to -1, $\phi(x) > 0$ and $\varepsilon\phi''(x) = f(\phi(x)) > 0$ which implies $\lim_{x \to -1+} \varepsilon\phi''(x) = f(\phi(-1)) > 0$. Since $\phi'(-1) \geqslant 0$ and $\lim_{x \to -1+} \phi''(x) > 0$, then ϕ is monotone increasing in $(-1, -1 + \delta_0)$, where $\delta_0 > 0$ is a constant. Now we may show that ϕ is monotone increasing in $(-1, 1)$ by contradiction. Suppose ϕ has a local maximum point at $x_0 \in (-1, 1)$ such that $\phi'(x_0) = 0, \phi''(x_0) \leqslant 0$ and ϕ is monotone increasing in $(-1, x_0)$. However, $\varepsilon\phi''(x_0) = f(\phi(x_0)) \geqslant f(\phi(-1)) > 0$ contradicts to $\phi''(x_0) \leqslant 0$. Hence ϕ is monotone increasing in $(-1, 1)$ which provides $\phi''(x) = \frac{1}{\varepsilon}f(\phi(x)) \geqslant \frac{1}{\varepsilon}f(\phi(-1))$ i.e. $\phi''(x) \geqslant \frac{1}{\varepsilon}f(\phi(-1))$ for $x \in (-1, 1)$. Integrating the inequality from -1 to x, we have $\phi'(x) - \phi'(-1) \geqslant \frac{1}{\varepsilon}f(\phi(-1))(x+1)$ i.e. $\phi'(x) \geqslant \phi'(-1) + \frac{1}{\varepsilon}f(\phi(-1))(x+1)$ for $x \in (-1, 1)$ which implies

$$\phi(1) - \phi(-1) = \int_{-1}^{1} \phi'(x)\,\mathrm{d}x \geqslant \int_{-1}^{1}\left[\phi'(-1) + \frac{1}{\varepsilon}f(\phi(-1))(x+1)\right]\mathrm{d}x$$

$$= 2\left[\phi'(-1) + \frac{1}{\varepsilon}f(\phi(-1))\right],$$

i.e. $\phi(1) \geqslant \phi(-1) + 2\left[\phi'(-1) + \frac{1}{\varepsilon}f(\phi(-1))\right] \geqslant \frac{2}{\varepsilon}f(\phi(-1))$. On the other hand, the Robin boundary condition (4.2) gives $\phi_0(1) = \phi(1) + \eta_\varepsilon\phi'(1) \geqslant \phi(1)$ and $\phi(-1) = \phi_0(-1) + \eta_\varepsilon\phi'(-1) \geqslant \phi_0(-1) > 0$ since ϕ is monotone increasing in $(-1, 1)$. Thus

$$\phi_0(1) \geqslant \phi(1) \geqslant \frac{2}{\varepsilon}f(\phi(-1)) \geqslant \frac{2}{\varepsilon}f(\phi_0(-1)),$$

which contradicts to the hypothesis that $\phi_0(1), \phi_0(-1)$ are independent to ε.

For the Case II, we first use the Robin boundary condition (4.2) to get $\eta_\varepsilon \phi'(-1) = \phi(-1) - \phi_0(-1) \leqslant -\phi_0(-1) < 0$ which implies $\eta_\varepsilon > 0$ and $\phi'(-1) < 0$. Then $\phi(x) < 0$ for $x \in (-1, -1+\delta_1)$ and ϕ is monotone decreasing in $(-1, -1+\delta_1)$, where $\delta_1 > 0$ is a constant. Hence ϕ is negative and monotone decreasing in $(-1, 1)$. Otherwise, there exists $x_3 \in (-1, 1)$ a local minimum point of ϕ such that $\phi(x_3) < 0$ and $\phi''(x_3) \geqslant 0$ but $\phi''(x_3) = \frac{1}{\varepsilon} f(\phi(x_3)) < 0$ which contradicts to $\phi''(x_3) \geqslant 0$. Such a contradiction shows that ϕ is negative and monotone decreasing in $(-1, 1)$. However, $0 > \phi(1) = \phi_0(1) - \eta_\varepsilon \phi'(1) \geqslant \phi_0(1)$ contradicts to $\phi_0(1) > 0$. Notice that both Case I and II produce contradiction. Similarly, the condition $\phi(1) > 0$ and $\phi'(1) \leqslant 0$ and the other condition $\phi(1) \leqslant 0$ also result in contradiction, respectively. Therefore, we may complete the proof of claim I.

By claim I, there exists $x_1 \in (-1, 1)$ a local minimum point of ϕ, and then by proposition 4.2 (a), we may complete the proof of lemma 4.1 (i). On the other hand, we may also use the similar argument of claim I to prove that there exists $x_2 \in (-1, 1)$ a local maximum point of ϕ. Hence by proposition 4.2 (b), we complete the proof of lemma 4.1 (ii).

Now we prove lemma 4.1 (iii) by contradiction. Suppose ϕ is not monotone increasing. By proposition 4.2, it is sufficient to consider two cases as follows: $\phi(-1) < 0$ and $\phi(-1) > 0$. If $\phi(-1) < 0$, then proposition 4.2 implies that there exists $x_2 \in (-1, 1)$ a maximum point of ϕ such that $\phi(x_2) < 0$, ϕ is monotone increasing in $(-1, x_2)$ and decreasing in $(x_2, 1)$ so $\phi'(1) \leqslant 0$. However, the boundary condition $\phi(1) + \eta_\varepsilon \phi'(1) = \phi_0(1)$ and $\phi'(1) \leqslant 0$ give $\phi_0(1) \leqslant \phi_0(1) - \eta_\varepsilon \phi'(1) = \phi(1) \leqslant \phi(x_2) < 0$ which contradicts to $\phi_0(1) \geqslant c = 0$. On the other hand, if $\phi(-1) > 0$, then proposition 4.2 implies that there exists $x_1 \in (-1, 1)$ a minimum point of ϕ such that $\phi(x_1) > 0$, ϕ is monotone decreasing in $(-1, x_1)$ and increasing in $(x_1, 1)$ so $\phi'(-1) \leqslant 0$. However, the boundary condition $\phi(-1) - \eta_\varepsilon \phi'(-1) = \phi_0(-1)$ and $\phi'(-1) \leqslant 0$ give $\phi_0(-1) = \phi(-1) - \eta_\varepsilon \phi'(-1) \geqslant \phi(-1) > 0$ which contradicts to $\phi_0(-1) \leqslant c = 0$. Therefore, we complete the proof of lemma 4.1 (iii). Similar argument of lemma 4.1 (iii) can be applied to prove lemma 4.1 (iv) and we omit the detail here.

Using lemma 4.1 (i)–(iv), we may prove $\min\{\phi_0(-1), \phi_0(1), 0\} \leqslant \phi(x) \leqslant \max\{\phi_0(-1), \phi_0(1), 0\}$ for $x \in (-1, 1)$. The proof is stated as follows: By lemma 4.1 (i) and the boundary condition (4.2), we have $\phi(-1) = \phi_0(-1) + \eta_\varepsilon \phi'(-1) \leqslant \phi_0(-1)$, $\phi(1) = \phi_0(1) - \eta_\varepsilon \phi'(1) \leqslant \phi_0(1)$ and $c = 0 < \phi(x_1) \leqslant \phi(x) \leqslant \max\{\phi(1), \phi(-1)\} \leqslant \max\{\phi_0(1), \phi_0(-1)\}$ for $x \in (-1, 1)$. Similarly, lemma 4.1 (ii) and the boundary condition (4.2) imply $\phi(-1) = \phi_0(-1) + \eta_\varepsilon \phi'(-1) \geqslant \phi_0(-1)$, $\phi(1) = \phi_0(1) - \eta_\varepsilon \phi'(1) \geqslant \phi_0(1)$ and $c = 0 > \phi(x_2) \geqslant \phi(x) \geqslant \min\{\phi(1), \phi(-1)\} \geqslant \min\{\phi_0(1), \phi_0(-1)\}$ for $x \in (-1, 1)$. On the other hand, we may apply lemma 4.1 (iii) and the boundary condition (4.2) to get $\phi(-1) = \phi_0(-1) + \eta_\varepsilon \phi'(-1) \geqslant \phi_0(-1)$, $\phi(1) = \phi_0(1) - \eta_\varepsilon \phi'(1) \leqslant \phi_0(1)$ and $\phi_0(-1) \leqslant \phi(-1) \leqslant \phi(x) \leqslant \phi(1) \leqslant \phi_0(1)$ for $x \in (-1, 1)$. Similarly, lemma 4.1 (iv) and the boundary condition (4.2) give $\phi(-1) = \phi_0(-1) + \eta_\varepsilon \phi'(-1) \leqslant \phi_0(-1)$, $\phi(1) = \phi_0(1) - \eta_\varepsilon \phi'(1) \geqslant \phi_0(1)$ and $\phi_0(-1) \geqslant \phi(-1) \geqslant \phi(x) \geqslant \phi(1) \geqslant \phi_0(1)$ for $x \in (-1, 1)$. Hence we complete the proof of lemma 4.1 (v) i.e.

$$\min\{\phi_0(-1), \phi_0(1), 0\} \leqslant \phi(x) \leqslant \max\{\phi_0(-1), \phi_0(1), 0\} \quad \text{for} \quad x \in (-1, 1). \tag{4.3}$$

Let $A_0 = \min\{\phi_0(-1), \phi_0(1), 0\}$ and $A_1 = \max\{\phi_0(-1), \phi_0(1), 0\}$. Then (4.3) implies

$$\|\phi\|_{L^\infty} \leqslant A_2 = \max\{-A_0, A_1\}. \tag{4.4}$$

Since $A \leqslant \phi_0(-1), \phi_0(1) \leqslant M$ and $A < 0$, then (4.3) gives

$$\phi(x) \in [A_0, A_1] \subset [A, M] \quad \text{for} \quad x \in (-1, 1), \tag{4.5}$$

i.e. each value of $\phi(x)$ must be contained in the original domain of the function f. Thus we may neglect the extension of the function f and regard ϕ as a well-defined solution of equation (4.1) with boundary condition (4.2).

Now we claim that $\phi(x) \to 0$ as $\varepsilon \to 0+$ for $x \in (-1, 1)$. To prove this, we remark that

$$\frac{\varepsilon}{2}(\phi^2)''(x) = \varepsilon(\phi\phi'' + (\phi')^2)(x) \geqslant \varepsilon\phi\phi''(x) = \phi(x)f(\phi(x))$$

$$= \phi(x)\int_0^{\phi(x)} f'(s)\,\mathrm{d}s \geqslant \alpha_0\phi^2(x),$$

for $x \in (-1, 1)$, where $\alpha_0 = \min_{z \in [A_0, A_1]} f'(z) > 0$ is a constant coming from the strictly monotone increasing of f. Note that if $\phi(x) < 0$, then

$$\phi(x)\int_0^{\phi(x)} f'(s)\,\mathrm{d}s = (-\phi(x))\int_{\phi(x)}^0 f'(s)\,\mathrm{d}s \geqslant (-\phi(x))\int_{\phi(x)}^0 \alpha_0\,\mathrm{d}s = \alpha_0\phi^2(x).$$

Since $\frac{\varepsilon}{2}(\phi^2)''(x) \geqslant \alpha_0\phi^2(x)$ for $x \in (-1, 1)$, then by (4.4) and the standard comparison theorem, we have $\phi^2(x) \leqslant A_2^2\left(\mathrm{e}^{-(1+x)\sqrt{2\alpha_0/\varepsilon}} + \mathrm{e}^{-(1-x)\sqrt{2\alpha_0/\varepsilon}}\right)$ for $x \in (-1, 1)$. Therefore, $\phi(x) \to 0$ as $\varepsilon \to 0+$ for $x \in (-1, 1)$, and we may complete the proof of lemma 4.1 (vi) and completed the proof of lemma 4.1.   □

Remark 4.3. The equation (4.1) with the boundary condition (4.2) has a unique solution.

The uniqueness comes from the strictly monotone increasing of the function f. The proof is sketched as follows: Suppose ϕ_1 and ϕ_2 are solutions of (4.1) and (4.2). We may subtract the equation of ϕ_1 by that of ϕ_2, and multiply the resulting equation by $u = \phi_1 - \phi_2$ and integrate it over $(-1, 1)$. Then using integration by part, we have $u'(1)u(1) - u'(-1)u(-1) - \int_{-1}^1 (u'(x))^2\,\mathrm{d}x = \int_{-1}^1 c(x)u^2\,\mathrm{d}x$, where $c(x) = \frac{f(\phi_1(x)) - f(\phi_2(x))}{\phi_1(x) - \phi_2(x)}$ is positive since the function f is strictly monotone increasing. On the other hand, the Robin boundary condition (4.2) gives $u(-1) = \eta_\varepsilon u'(-1)$, $u(1) = -\eta_\varepsilon u'(1)$ and $u'(1)u(1) - u'(-1)u(-1) = -\eta_\varepsilon\left[(u'(-1))^2 + (u'(1))^2\right]$. Hence

$$0 \leqslant \int_{-1}^1 c(x)u^2\,\mathrm{d}x = -\eta_\varepsilon\left[(u'(-1))^2 + (u'(1))^2\right] - \int_{-1}^1 (u'(x))^2\,\mathrm{d}x \leqslant 0$$

which implies $u \equiv 0$ i.e. $\phi_1 \equiv \phi_2$ and the uniqueness proof of ϕ is complete.

Remark 4.4. The solution ϕ of the equation (4.1) with the boundary condition (4.2) has linear stability.

To get the linear stability of the solution ϕ of the equation (4.1) with the boundary condition (4.2), we study the eigenvalue problem $Lv = \lambda v$ of the corresponding linearized operator $Lv = -\varepsilon v'' + f'(\phi)v$ with the boundary condition $v(\pm 1) \pm \eta_\varepsilon v(\pm 1) = 0$. Using integration by part, it is obvious that

$$\lambda\int_{-1}^1 v^2\,\mathrm{d}x = \int_{-1}^1 v\,Lv\,\mathrm{d}x = \int_{-1}^1 \varepsilon v''v\,\mathrm{d}x + \int_{-1}^1 f'(\phi)\,v^2\,\mathrm{d}x$$

$$= \eta_\varepsilon\left[(v'(-1))^2 + (v'(1))^2\right] + \int_{-1}^1 \left[\varepsilon(v')^2 + f'(\phi)v^2\right]\,\mathrm{d}x \geqslant \mu_0\int_{-1}^1 v^2\,\mathrm{d}x, \quad (4.6)$$

and hence $\lambda \geqslant \mu_0 > 0$, where $\mu_0 = \min_{s \in [\min\phi, \max\phi]} f'(s)$ is a positive constant arising from the strictly monotone increasing of the function f.

In lemma 4.1, the existence of zero point ϕ_A (or ϕ_B) of f is essential. If the function f has not any zero point like ϕ_A (or ϕ_B) i.e. the value of f is away from zero, then the equation (4.1) may not have uniformly bounded solutions $\{\phi\}_{\varepsilon>0}$. Such a result is stated as follows:

Lemma 4.5. *Assume f is a function satisfying one of the following conditions:*

(a) $f : [A, \infty) \to \mathbb{R}$ is monotone increasing, $A < 0$ and $f(A) > 0$.
(b) $f : (-\infty, B] \to \mathbb{R}$ is monotone increasing, $B > 0$ and $f(B) < 0$.

For each $\varepsilon > 0$, let ϕ be a solution of the equation (4.1). Then $\sup\limits_{\varepsilon>0} \|\phi\|_{L^\infty} = \infty$.

Proof. Without loss of generality, we may assume the function f satisfies the condition (a). Now we prove lemma 4.5 by contradiction. Suppose $\{\phi\}_{\varepsilon>0}$ is uniformly bounded i.e. $\sup\limits_{\varepsilon>0} \|\phi\|_{L^\infty} < \infty$. We divide three cases to complete the proof as follows:

Case I. The solution $\phi = \phi(x)$ is monotone decreasing to x i.e. $\phi'(x) \leqslant 0$ for $x \in (-1, 1)$.
Using the equation $\varepsilon\phi'' = f(\phi)$ and the condition (a), we have

$$
\begin{aligned}
-\phi'(x) \geqslant \phi'(1) - \phi'(x) &= \int_x^1 \phi''(\tau)\,d\tau \\
&= \varepsilon^{-1} \int_x^1 f(\phi(\tau))\,d\tau \\
&\geqslant \varepsilon^{-1} \int_x^1 f(A)\,d\tau = \varepsilon^{-1} f(A)(1 - x), \quad \forall x \in (-1, 1),
\end{aligned}
$$

and hence

$$
-2\|\phi\|_{L^\infty} \leqslant \phi(1) - \phi(-1) = \int_{-1}^1 \phi'(x)\,dx \leqslant -\varepsilon^{-1} f(A) \int_{-1}^1 (1 - x)\,dx = -2\varepsilon^{-1} f(A),
$$

i.e. $\|\phi\|_{L^\infty} \geqslant \varepsilon^{-1} f(A) \to \infty$ as $\varepsilon \to 0+$ which contradicts to the hypothesis $\sup\limits_{\varepsilon>0} \|\phi\|_{L^\infty} < \infty$.

Case II. The solution $\phi = \phi(x)$ is monotone increasing to x i.e. $\phi'(x) \geqslant 0$ for $x \in (-1, 1)$.
As for the argument of Case I, we obtain

$$
\begin{aligned}
\phi'(x) \geqslant \phi'(x) - \phi'(-1) &= \int_{-1}^x \phi''(\tau)\,d\tau \\
&= \varepsilon^{-1} \int_{-1}^x f(\phi(\tau))\,d\tau \\
&\geqslant \varepsilon^{-1} \int_{-1}^x f(A)\,d\tau = \varepsilon^{-1} f(A)(1 + x), \quad \forall x \in (-1, 1),
\end{aligned}
$$

and hence

$$
2\|\phi\|_{L^\infty} \geqslant \phi(1) - \phi(-1) = \int_{-1}^1 \phi'(x)\,dx \geqslant \varepsilon^{-1} f(A) \int_{-1}^1 (1 + x)\,dx = 2\varepsilon^{-1} f(A),
$$

i.e. $\|\phi\|_{L^\infty} \geqslant \varepsilon^{-1} f(A) \to \infty$ as $\varepsilon \to 0+$ which contradicts to the hypothesis $\sup\limits_{\varepsilon>0} \|\phi\|_{L^\infty} < \infty$.

Case III. The solution $\phi = \phi(x)$ has a local minimum point at $x_0 \in (-1, 1)$ such that $\phi'(x_0) = 0$ and $\phi''(x_0) > 0$.

Note that since $\varepsilon\phi'' = f(\phi) \geqslant f(A) > 0$, it is impossible to have any local maximum point in $(-1, 1)$. By the equation $\varepsilon\phi'' = f(\phi)$ and the condition (a), we have

$$
\begin{aligned}
-\phi'(x) = \phi'(x_0) - \phi'(x) &= \int_x^{x_0} \phi''(\tau)\,\mathrm{d}\tau \\
&= \varepsilon^{-1} \int_x^{x_0} f(\phi(\tau))\,\mathrm{d}\tau \\
&\geqslant \varepsilon^{-1} \int_x^{x_0} f(A)\,\mathrm{d}\tau = \varepsilon^{-1} f(A)(x_0 - x), \quad \forall x \in (-1, x_0),
\end{aligned}
$$

and hence

$$
\begin{aligned}
-2\|\phi\|_{L^\infty} \leqslant \phi(x_0) - \phi(-1) = \int_{-1}^{x_0} \phi'(x)\,\mathrm{d}x &\leqslant -\varepsilon^{-1} f(A) \int_{-1}^{x_0} (x_0 - x)\,\mathrm{d}x \\
&= -\frac{1}{2}\varepsilon^{-1} f(A)(x_0 + 1)^2,
\end{aligned}
$$

i.e.

$$
|x_0 + 1| \leqslant 2\varepsilon^{1/2}\sqrt{\|\phi\|_{L^\infty}/f(A)}. \tag{4.7}
$$

On the other hand,

$$
\begin{aligned}
\phi'(x) = \phi'(x) - \phi'(x_0) &= \int_{x_0}^x \phi''(\tau)\,\mathrm{d}\tau \\
&= \varepsilon^{-1} \int_{x_0}^x f(\phi(\tau))\,\mathrm{d}\tau \\
&\geqslant \varepsilon^{-1} \int_{x_0}^x f(A)\,\mathrm{d}\tau = \varepsilon^{-1} f(A)(x - x_0), \quad \forall x \in (x_0, 1),
\end{aligned}
$$

and hence

$$
2\|\phi\|_{L^\infty} \geqslant \phi(1) - \phi(x_0) = \int_{x_0}^1 \phi'(x)\,\mathrm{d}x \geqslant \varepsilon^{-1} f(A) \int_{x_0}^1 (x - x_0)\,\mathrm{d}x = \frac{1}{2}\varepsilon^{-1} f(A)(x_0 - 1)^2,
$$

i.e.

$$
|x_0 - 1| \leqslant 2\varepsilon^{1/2}\sqrt{\|\phi\|_{L^\infty}/f(A)}. \tag{4.8}
$$

Therefore, as $\varepsilon > 0$ sufficiently small, (4.7) and (4.8) provide a contradiction and we may complete the proof of lemma 4.5. $\qquad\square$

5. Excess currents due to steric effects

Here we want to use solutions ϕ_ε^A and ϕ_ε^B of (1.13)–(1.15) (see theorem 1.1 and 1.3) to calculate excess currents (due to steric effects) represented by formula (1.9). By (1.10),

$$
\sum_{j=1}^N g_{ij} c_j = -k_{\mathrm{B}}T \ln c_i - z_i e\phi,
$$

and then formula (1.9) becomes

$$
I^{ex} = \sum_{i=1}^N z_i e D_i \left(\nabla c_i + z_i c_i \nabla\tilde{\phi}\right), \tag{5.1}
$$

where $\tilde{\phi} = \frac{e}{k_{\mathrm{B}}T}\phi$.

5.1. Under the same hypotheses of theorem 1.1

Here we set $N = 3$, $z_2 = -z_1 = q \geqslant 1$, $z_3 > 0$, $\rho_0 > 0$, and assume that $g_{11} = g_{22} = g > 0$ is fixed, $g_{12} = g_{21} = z > 0$ is sufficiently large, and $g_{i3} = g_{3i} = 0$ for $i = 1, 2, 3$. By (3.3), we have $c_3 = e^{-z_3 \tilde{\phi}}$ which implies $\nabla c_3 + z_3 c_3 \nabla \tilde{\phi} = 0$. Hence (5.1) becomes

$$
\begin{aligned}
I^{ex} &= \sum_{i=1}^{2} z_i e D_i \left(\frac{\mathrm{d}c_i}{\mathrm{d}x} + z_i c_i \frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}x} \right) \\
&= -q\, e D_1 \left(\frac{\mathrm{d}c_1}{\mathrm{d}x} - q\, c_1 \frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}x} \right) + q\, e D_2 \left(\frac{\mathrm{d}c_2}{\mathrm{d}x} + q\, c_2 \frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}x} \right) \\
&= q\, e \left[\frac{\mathrm{d}}{\mathrm{d}x} \left(-D_1 c_1 + D_2 c_2 \right) + q\, \left(D_1 c_1 + D_2 c_2 \right) \frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}x} \right]
\end{aligned}
$$

i.e.

$$
I^{ex} = q\, e \left[\frac{\mathrm{d}}{\mathrm{d}x} \left(-D_1 c_1 + D_2 c_2 \right) + q\, \left(D_1 c_1 + D_2 c_2 \right) \frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}x} \right] \tag{5.2}
$$

Using $c_1 = \frac{c_1 - c_2}{2} + \frac{c_1 + c_2}{2}$ and $c_2 = \frac{c_1 + c_2}{2} - \frac{c_1 - c_2}{2}$, formula (5.2) can be expressed as

$$
\begin{aligned}
\frac{1}{q\,e} I^{ex} =\ & \frac{\mathrm{d}}{\mathrm{d}x} \left[\frac{D_2 - D_1}{2} \left(c_1 + c_2 \right) - \frac{D_1 + D_2}{2} \left(c_1 - c_2 \right) \right] \\
& + q \left[\frac{D_1 + D_2}{2} \left(c_1 + c_2 \right) - \frac{D_2 - D_1}{2} \left(c_1 - c_2 \right) \right] \frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}x}
\end{aligned} \tag{5.3}
$$

Note that $c_1 + c_2 = \Sigma$ and $c_1 - c_2 = \begin{cases} \sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}} & \text{on } A, \\ -\sqrt{\Sigma^2 - 4e^{-(g+z)\Sigma}} & \text{on } B. \end{cases}$ (see (2.8) in section 2). As for (3.5)–(3.7), we may set $\left(\Sigma, \tilde{\phi} \right) = \left(\Sigma_{A_1} \left(\tilde{\phi} \right), \phi_\varepsilon^A (x) \right)$ and $\left(\Sigma, \tilde{\phi} \right) = \left(\Sigma_{B_1} \left(\tilde{\phi} \right), \phi_\varepsilon^B (x) \right)$, respectively. Then along $c_1 + c_2 = \Sigma = \Sigma_{A_1}$ and $c_1 - c_2 = \sqrt{\Sigma_{A_1}^2 - 4e^{-(g+z)\Sigma_{A_1}}}$, we may use (2.17), (2.21) and Chain Rule to get

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x} \left(c_1 + c_2 \right) &= \frac{\mathrm{d}}{\mathrm{d}x} \Sigma_{A_1} \left(\phi_\varepsilon^A (x) \right) \\
&= \frac{\mathrm{d}\Sigma_{A_1}}{\mathrm{d}\phi} \left(\phi_\varepsilon^A (x) \right) \frac{\mathrm{d}\phi_\varepsilon^A}{\mathrm{d}x} (x) \\
&= \frac{\sqrt{\Sigma_{A_1}^2 \left(\phi_\varepsilon^A(x) \right) - 4e^{-(g+z)\Sigma_{A_1} \left(\phi_\varepsilon^A(x) \right)}}}{1 + g \Sigma_{A_1} \left(\phi_\varepsilon^A(x) \right) + \left(g^2 - z^2 \right) e^{-(g+z)\Sigma_{A_1} \left(\phi_\varepsilon^A(x) \right)}} \frac{\mathrm{d}\phi_\varepsilon^A}{\mathrm{d}x} (x) ,
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x} \left(c_1 - c_2 \right) &= \frac{\mathrm{d}}{\mathrm{d}x} \left(c_1 - c_2 \right) \left(\Sigma_{A_1} \left(\phi_\varepsilon^A (x) \right) \right) \\
&= \frac{\mathrm{d}}{\mathrm{d}\phi} \left(c_1 - c_2 \right) \left(\Sigma_{A_1} \left(\phi_\varepsilon^A (x) \right) \right) \frac{\mathrm{d}\phi_\varepsilon^A}{\mathrm{d}x} (x) \\
&= \frac{\Sigma_{A_1} \left(\phi_\varepsilon^A(x) \right) + 2(g+z) e^{-(g+z)\Sigma_{A_1} \left(\phi_\varepsilon^A(x) \right)}}{1 + g \Sigma_{A_1} \left(\phi_\varepsilon^A(x) \right) + \left(g^2 - z^2 \right) e^{-(g+z)\Sigma_{A_1} \left(\phi_\varepsilon^A(x) \right)}} \frac{\mathrm{d}\phi_\varepsilon^A}{\mathrm{d}x} (x) .
\end{aligned}
$$

For simplicity, we may set $\hat{\Sigma}_{A_1} = \Sigma_{A_1} \left(\phi_\varepsilon^A (x) \right)$ and denote $\frac{\mathrm{d}}{\mathrm{d}x} \left(c_1 \pm c_2 \right)$ as follows:

$$
\frac{\mathrm{d}}{\mathrm{d}x} \left(c_1 + c_2 \right) = \frac{\sqrt{\hat{\Sigma}_{A_1}^2 - 4e^{-(g+z)\hat{\Sigma}_{A_1}}}}{1 + g \hat{\Sigma}_{A_1} + \left(g^2 - z^2 \right) e^{-(g+z)\hat{\Sigma}_{A_1}}} \frac{\mathrm{d}\phi_\varepsilon^A}{\mathrm{d}x} (x) ,
$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(c_1 - c_2\right) = \frac{\hat{\Sigma}_{A_1} + 2(g+z)\mathrm{e}^{-(g+z)\hat{\Sigma}_{A_1}}}{1 + g\hat{\Sigma}_{A_1} + (g^2 - z^2)\mathrm{e}^{-(g+z)\hat{\Sigma}_{A_1}}}\frac{\mathrm{d}\phi_\varepsilon^A}{\mathrm{d}x}(x).$$

Consequently, by setting $I_A^{ex} = I^{ex}$ along A_1 and $I_B^{ex} = I^{ex}$ along B_1, (5.3) becomes

$$I_A^{ex} = q\,e\,i_A\left(\hat{\Sigma}_{A_1}\right)\frac{\mathrm{d}\phi_\varepsilon^A}{\mathrm{d}x}(x),\tag{5.4}$$

and

$$I_B^{ex} = q\,e\,i_B\left(\hat{\Sigma}_{B_1}\right)\frac{\mathrm{d}\phi_\varepsilon^B}{\mathrm{d}x}(x),\tag{5.5}$$

where

$$i_A\left(\Sigma\right) = \frac{\frac{D_2 - D_1}{2}\sqrt{\Sigma^2 - 4\mathrm{e}^{-(g+z)\Sigma}} - \frac{D_1 + D_2}{2}\left[\Sigma + 2\left(g+z\right)\mathrm{e}^{-(g+z)\Sigma}\right]}{1 + g\Sigma + \left(g^2 - z^2\right)\mathrm{e}^{-(g+z)\Sigma}},$$

$$+q\left[\frac{D_1 + D_2}{2}\Sigma - \frac{D_2 - D_1}{2}\sqrt{\Sigma^2 - 4\mathrm{e}^{-(g+z)\Sigma}}\right]\tag{5.6}$$

and

$$i_B\left(\Sigma\right) = \frac{\frac{D_1 - D_2}{2}\sqrt{\Sigma^2 - 4\mathrm{e}^{-(g+z)\Sigma}} - \frac{D_1 + D_2}{2}\left[\Sigma + 2\left(g+z\right)\mathrm{e}^{-(g+z)\Sigma}\right]}{1 + g\Sigma + \left(g^2 - z^2\right)\mathrm{e}^{-(g+z)\Sigma}}.$$

$$+q\left[\frac{D_1 + D_2}{2}\Sigma - \frac{D_1 - D_2}{2}\sqrt{\Sigma^2 - 4\mathrm{e}^{-(g+z)\Sigma}}\right]\tag{5.7}$$

Without loss of generality, ϕ_ε^A can be assumed as a monotone increasing function. Such an assumption can be fulfilled by setting $\phi_0(-1) < \phi_0(1)$ and using lemma 4.1 (iii). Integrating I_A^{ex} from x_1 to x_2, we have

$$\int_{x_1}^{x_2} I_A^{ex}\mathrm{d}x = e\int_{x_1}^{x_2} i_A\left(\Sigma_{A_1}\left(\phi_\varepsilon^A(x)\right)\right)\frac{\mathrm{d}\phi_\varepsilon^A}{\mathrm{d}x}\mathrm{d}x = e\int_{\phi_1^A}^{\phi_2^A} i_A\left(\Sigma_{A_1}(\phi)\right)\mathrm{d}\phi,\tag{5.8}$$

for $-1 < x_1 < x_2 < 1$, where $\phi_1 \leqslant \phi_2$ and $\phi_j^A = \phi_\varepsilon^A\left(x_j\right)$, $j = 1, 2$. Setting $\Sigma = \Sigma_{A_1}$ and using change of variables, Inverse Function Theorem and (2.17), we have

$$\mathrm{d}\phi = \frac{\mathrm{d}\phi}{\mathrm{d}\Sigma}\mathrm{d}\Sigma = \frac{1}{\frac{\mathrm{d}\Sigma}{\mathrm{d}\phi}}\mathrm{d}\Sigma = \frac{1 + g\Sigma + \left(g^2 - z^2\right)\mathrm{e}^{-(g+z)\Sigma}}{\sqrt{\Sigma^2 - 4\mathrm{e}^{-(g+z)\Sigma}}}\mathrm{d}\Sigma.$$

Therefore, by (5.4)–(5.8), we may complete the proof of (1.16) and (1.17).

5.2. Under the same hypotheses of theorem 1.3

Here we set $N = 4$, $z_2 = -z_1 = q_1 \geqslant 1$, $z_4 = -z_3 = q_2 \geqslant 1$, $\rho_0 \neq 0$, and assume that $g_{11} = g_{22} = g > 0$, $g_{33} = g_{44} = \tilde{g} > 0$ are fixed, $g_{12} = g_{21} = z > 0$, $g_{34} = g_{43} = \tilde{z} > 0$ are sufficiently large, and $g_{ij} = g_{ji} = 0$ for $i = 1, 2$ and $j = 3, 4$. As for section 3.2, these hypotheses imply that (1.13) can be decomposed into two independent equations (3.8) and (3.9) which have the same form as (3.2) with (3.4). Solving equations (3.8) and (3.9), we may get (c_1, c_2) (with branches A_1, B_1) and (c_3, c_4) (with branches M_1, N_1) as functions of ϕ, respectively. By (5.1), the excess currents of (c_1, c_2) and (c_3, c_4) can be represented as $\sum_{i=1}^{2} z_i e D_i\left(\frac{\mathrm{d}c_i}{\mathrm{d}x} + z_i c_i\frac{\mathrm{d}\phi}{\mathrm{d}x}\right)$ and $\sum_{i=3}^{4} z_i e D_i\left(\frac{\mathrm{d}c_i}{\mathrm{d}x} + z_i c_i\frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}x}\right)$ which can be calculated by the same method as section 5.1. We may denote the total excess current as $I_{A,M}^{ex} = I_A^{ex} + I_M^{ex}$, where I_A^{ex}

and I_M^{ex} are the excess currents along branches A_1 for (c_1, c_2) and M_1 for (c_3, c_4), respectively. Similarly, another total excess current can be denoted as $I_{B,N}^{ex} = I_B^{ex} + I_N^{ex}$, where I_B^{ex} and I_N^{ex} are the excess currents along branches B_1 for (c_1, c_2) and N_1 for (c_3, c_4), respectively. Therefore, as for (5.4)–(5.8), we may complete the proof of (1.18) and (1.19).

Acknowledgments

References

[1] Abbas Z, Gunnarsson M, Ahlberg E and Nordholm S 2002 Corrected DebyeHuckel theory of salt solutions: size asymmetry and effective diameters *J. Phys. Chem.* B **106** 1403–20

[2] Armstrong C M and Miller C 1990 Do voltage-dependent K+ channels require Ca2+? A critical test employing a heterologous expression system *Proc. Natl Acad. Sci. USA* **87** 7579–82

[3] Barcilon V, Chen D-P, Eisenberg R S and Jerome J W 1997 Qualitative properties of steady-state Poisson–Nernst–Planck systems: perturbation and simulation study *SIAM J. Appl. Math.* **57** 631–48

[4] Bazant M Z, Kilic M S, Storey B D and Ajdari A 2009 Towards an understanding of induced-charge electrokinetics at large applied voltages in concentrated solutions *Adv. Colloid Interface Sci.* **152** 48–88

[5] Bazant M Z, Thornton K and Ajdari A 2004 Diffuse-charge dynamics in electrochemical systems *Phys. Rev.* E **70** 021506-1-24

[6] Bazant M Z, Storey B D and Kornyshev A A 2011 Double layer in ionic liquids: overscreening versus crowding *Phys. Rev. Lett.* **106** 046102

[7] Brezzi F, Marini L D, Micheletti S, Pietra P, Sacco R and Wang S 2002 Finite element and finite volume discretizations of Drift-Diffusion type fluid models for semiconductors *Centre National de la Recherche Scientifique Paris, France, Technical Report* 2002-1302

[8] Burger M 2011 Inverse problems in ion channel modelling *Inverse Problems* **27** 083001

[9] Burger M, Francesco M D, Pietschmann J F and Schlake B 2010 Nonlinear cross-diffusion with size exclusion *SIAM J. Math. Anal.* **42** 2842–71

[10] Caprio D, Borkowska Z and Stafiej J 2004 Specific ionic interactions within a simple extension of the Gouy–Chapman theory including hard sphere effects *J. Electroanal. Chem.* **572** 51–9

[11] Coalson R D and Kurnikova M G 2005 Poisson–Nernst–Planck theory approach to the calculation of current through biological ion channels *IEEE Trans. Nanobiosci.* **4** 81–93

[12] Hodgkin A L and Huxley A F 1952 A quantitative description of membrane current and its application to conduction and excitation in nerve *J. Physiol.* **117** 500–44

[13] Eisenberg R and Chen D 1993 Poisson–Nernst–Planck (PNP) theory of an open ionic channel *Biophys. J.* **64** A22

[14] Eisenberg R S 1996 Atomic biology, electrostatics and ionic channels *New Developments and Theoretical Studies of Proteins* vol 7, ed R Elber (Philadelphia: World Scientific) pp 269–357 (arXiv:0807.0715)

[15] Eisenberg B and Liu W 2007 Poisson–Nernst–Planck systems for ion channels with permanent charges *SIAM J. Math. Anal.* **38** 1932–66

[16] Eisenberg B 2012 Ion in fluctuating channels: transistors alive *Fluctuation Noise Lett.* **11** 1240001

[17] Eisenberg B 2011 Mass action in ionic solutions *Chem. Phys. Lett.* **511** 1–6

[18] Eisenberg B 2011 *Crowded Charges in Ion Channels, Advances in Chemical Physics* ed S A Rice (New York: Wiley) pp 77–223 http://arxiv.org/abs/1009.1786v1001

[19] Fraenkel D 2010 Monoprotic mineral acids analyzed by the smaller-ion shell model of strong electrolyte solutions *J. Phys. Chem.* B **115** 557–68

[20] Frankenhaeuser B and Hodgkin A L 1957 The action of calcium on the electrical properties of squid axons *J. Physiol.* **137** 218–44

[21] Horng T L, Lin T C, Liu C and Eisenberg B 2012 PNP equations with steric effects: a model of ion flow through channels *J. Phys. Chem.* B **116** 11422–41

[22] Howard J J, Perkyns J S and Pettitt B M 2010 The behavior of ions near a charged wall-dependence on ion size, concentration, and surface charge *J. Phys. Chem.* B **114** 6074–83

[23] Huxley A F 2002 From overshoot to voltage clamp *Trends Neurosci.* **25** 553–8

[24] Hyon Y, Eisenberg B and Liu C 2011 A Mathematical model for the hard sphere repulsion in ionic solutions *Commun. Math. Sci.* **9** 459–75

[25] Johannesson B 2010 Development of a generalized version of the Poisson–Nernst–Planck equations using the hybrid mixture theory: presentation of 2D numerical examples *Transp. Porous Media* **85** 565–92

[26] Kalcher I, Schulz J C F and Dzubiella J 2010 Ion-specific excluded-volume correlations and solvation forces *Phys. Rev. Lett.* **104** 097802

[27] Kaufman I, Luchinsky D G, Tindjong R, McClintock P V and Eisenberg R S 2013 Multi-ion conduction bands in a simple model of calcium ion channels *Phys. Biol.* **10** 026007

[28] Lee C C, Lin T C and Liu C 2013 Electrostatics for concentrated solutions: Poisson–Boltzmann equations with steric effects *in preparation*

[29] Lee C C, Lin T C and Liu C Finite size effects and fluctuation phenomena: new modified Poisson–Boltzmann equations *in preparation*

[30] Li B 2009 Continuum electrostatics for ionic solutions with nonuniform ionic sizes *Nonlinearity* **22** 811–33

[31] Lin T C and Eisenberg B 2014 A new approach to the Lennard-Jones potential and a new model: PNP-steric equations *Commun. Math. Sci.* **12** 149–73

[32] Liu W 2009 One-dimensional steady-state Poisson–Nernst–Planck systems for ion channels with multiple ion species *J. Differ. Equ.* **246** 428–51

[33] Lu B and Zhou Y C 2012 Poisson–Nernst–Planck equations for simulating biomolecular diffusion-reaction processes II: size effects on ionic distributions and diffusion-reaction rates *Biophys. J.* **100** 2475–85

[34] Markowich P A, Ringhofer C A and Schmeiser C 1990 *Semiconductor Equations* (New York: Springer)

[35] Selberherr S 1984 *Analysis and Simulation of Semiconductor Devices* (New York: Springer)

[36] Vincze J, Valisko M and Boda D 2010 The nonmonotonic concentration dependence of the mean activity coefficient of electrolytes is a result of a balance between solvation and ion–ion correlations *J. Chem. Phys.* **133** 154507

[37] Wang X S, He D, Wylie J J and Huang H 2014 Singular perturbation solutions of steady-state Poisson–Nernst–Planck systems *Phys. Rev.* E **89** 022722

[38] Wei G, Zheng Q, Chen Z and Xia K 2012 Variational multiscale models for charge transport *SIAM Rev.* **54** 699–754

[39] Ziebert F, Bazant M Z and Lacoste D 2010 Effective zero-thickness model for a conductive membrane driven by an electric field *Phys. Rev.* E **81** 031912