

Hydration-shell boundary conditions enable accurate continuum models for solvent-shell response

Jaydeep P. Bardhan

Dept. of Mechanical and Industrial Engineering



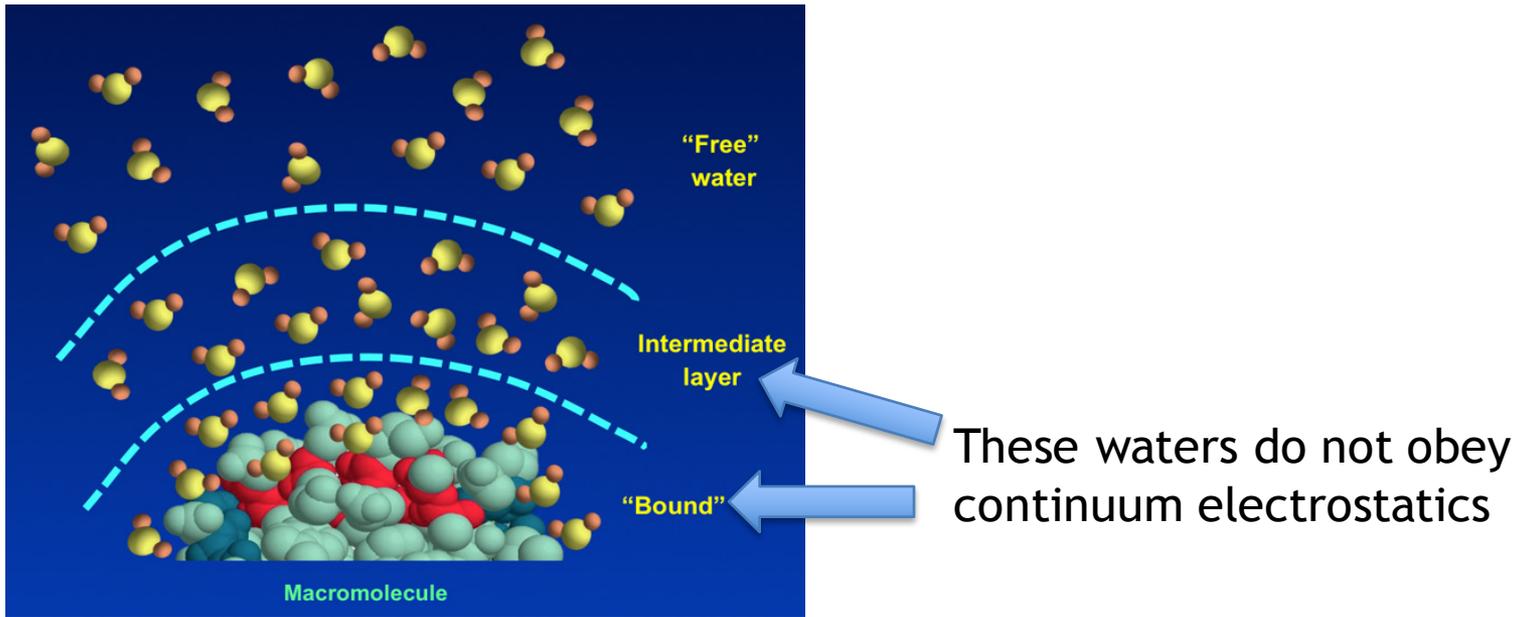
Northeastern University

Reporting work with B. Ren, A. Molavi Tabrizi, A. Rahimi, S. Goossens (NEU)
M. Knepley (Rice), C. Cooper (UTFSM, Chile)

Thanks to funding support from

- NIH R21 (GM102642-01) (PI: J. Bardhan)
- U.S. Dept. of Energy Office of Science, Advanced Scientific Computing Research (DE-AC02-06CH11357)

The multiscale challenge: the hydration shell



- Original motivation: how do we model the deviations from bulk behavior in the first layers of solvent?
- Can such a theory apply to mesoscopic systems e.g. colloids?
- This is true for pure water as well as electrolytes, organic solvents, ionic liquids, ...

The Basic Continuum Solvent Model

1. Inside the protein

$$\nabla^2 \varphi_{\text{protein}}(r) = - \frac{\sum_i q_i \delta(r - r_i)}{\epsilon_{\text{protein}} \epsilon_0}$$

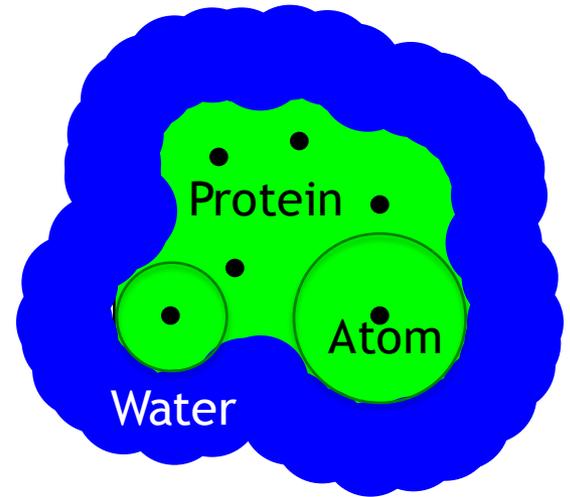
2. Outside (assume the solvent is infinite)*

$$\nabla^2 \varphi_{\text{solvent}}(r) = 0$$

3. Boundary conditions at the interface

$$\varphi_{\text{protein}}(r_{\Gamma}) = \varphi_{\text{solvent}}(r_{\Gamma})$$

$$\epsilon_{\text{protein}} \frac{\partial \varphi_{\text{protein}}}{\partial n} = \epsilon_{\text{solvent}} \frac{\partial \varphi_{\text{solvent}}}{\partial n}$$



*To include solvent ions:

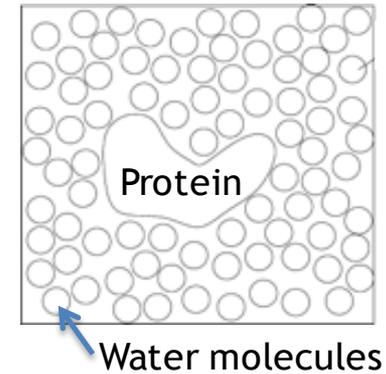
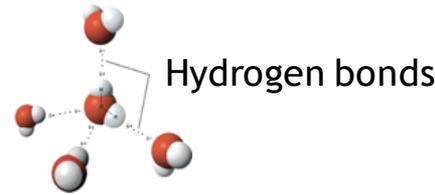
$$\nabla^2 \varphi(r) = \kappa^2 \sinh(\varphi)$$

Outline

- ❑ Not all multiscale models are created equal
- ❑ Hydration-shell Poisson-Boltzmann model
- ❑ HSPB applications and extensions
- ❑ Open questions and possible directions

Beyond the continuum hypothesis: nonlocal models

- On protein length scales, water has finite size
- Water forms semi-structured h-bond networks



- Idea: test *nonlocal continuum models* like in mechanics (gradient theories) and electrodynamics (spatially dispersive media)

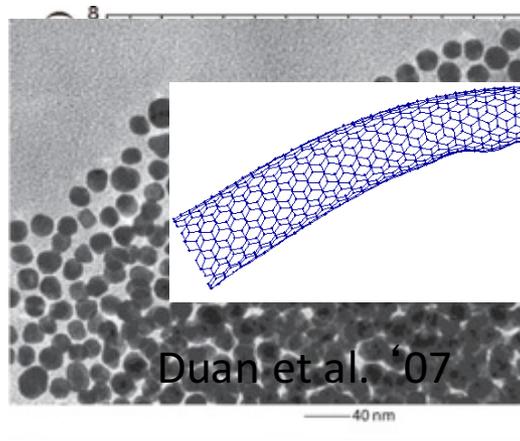
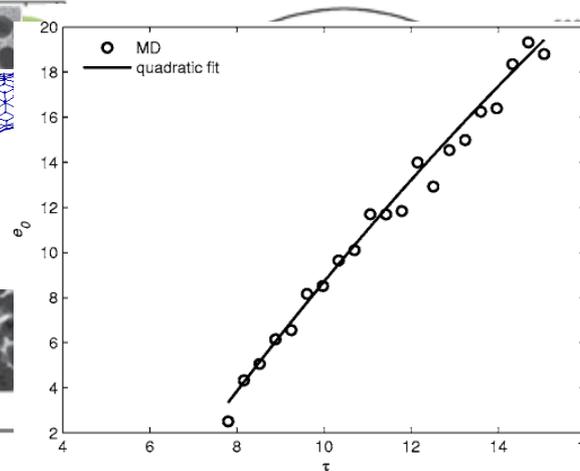
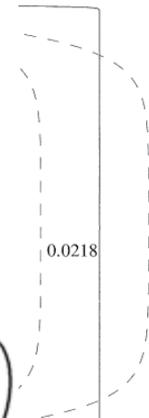


Figure 2. Scanning electron microscope images of DNA-linked gold particle aggregates. The particle diameter is 13 nm, and the DNA linkers are 72 base-pair duplexes.



city



Sharma et al., 2001

Edelman et al. 2001

A simple nonlocal dielectric model

❑ Polarization charge as a function of distance from the ion: not simple

- Short-range: electronic response
- Long-range: bulk behavior

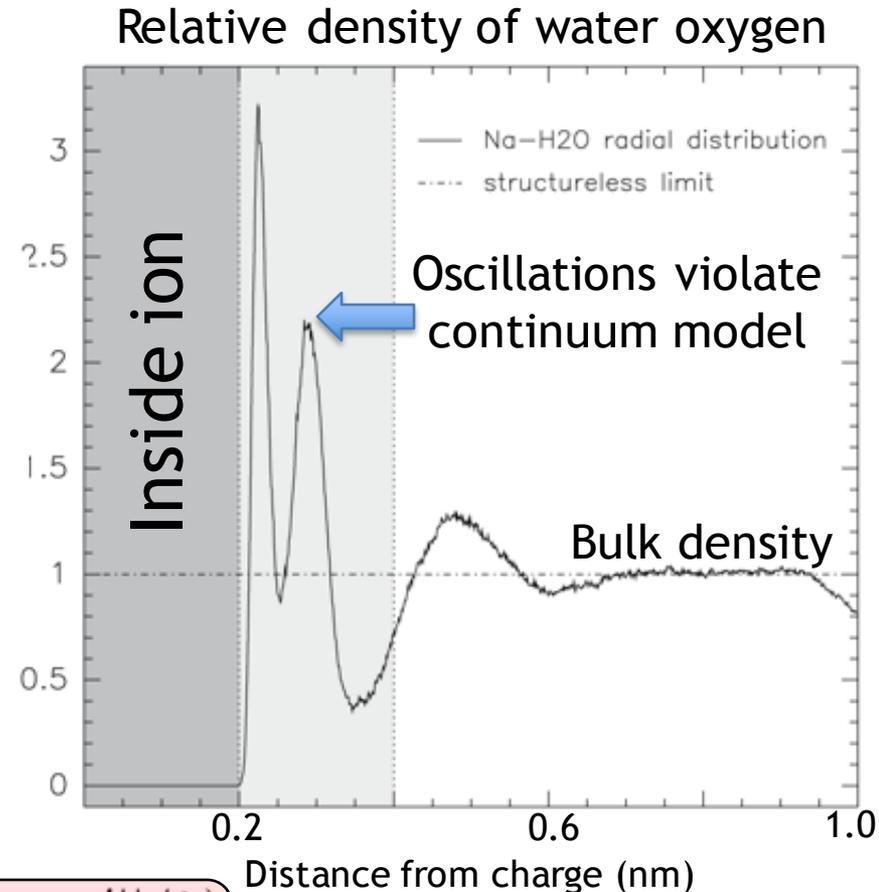
❑ Local: bulk everywhere

❑ Nonlocal: simple function that captures asymptotes

$$\epsilon(\mathbf{r}, \mathbf{r}') = \epsilon_\infty \delta(\mathbf{r} - \mathbf{r}') + \frac{\epsilon_\Sigma - \epsilon_\infty}{\lambda^2} \frac{\exp(-\|\mathbf{r} - \mathbf{r}'\|/\lambda)}{4\pi\|\mathbf{r} - \mathbf{r}'\|}$$

Green's function for

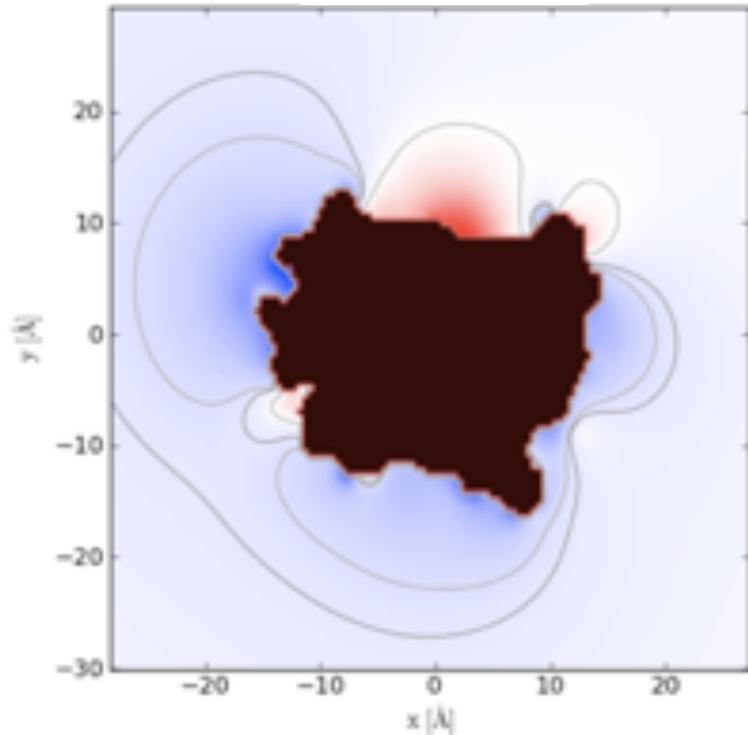
$$\nabla^2 \varphi = \frac{1}{\lambda^2} \varphi$$



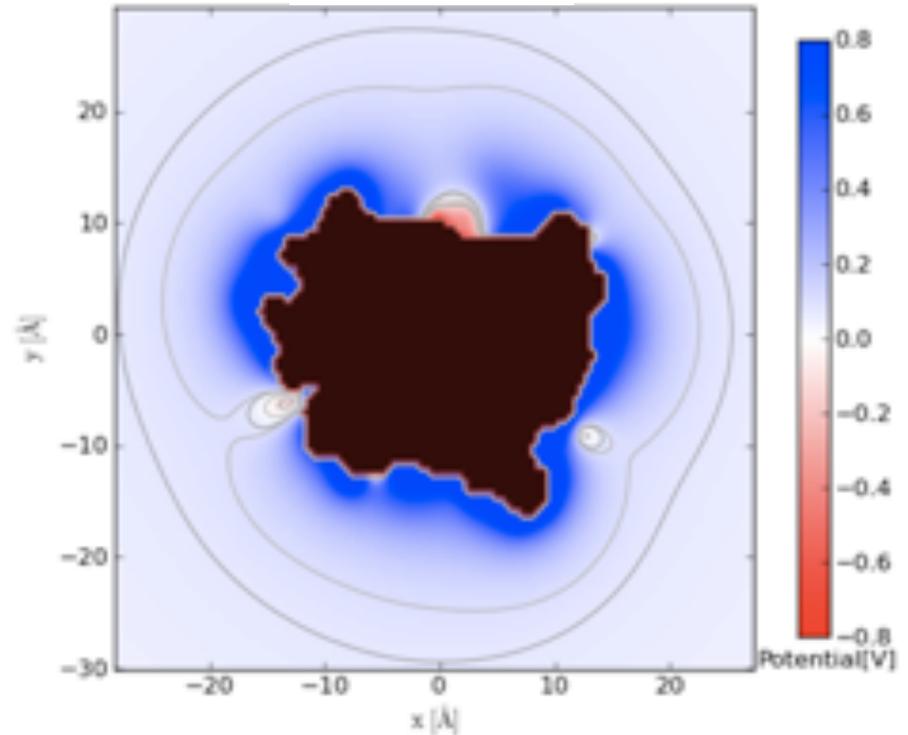
Nonlocality's impact on the water electric potential

- First fast BEM solver for proteins in nonlocal solvent

Local model



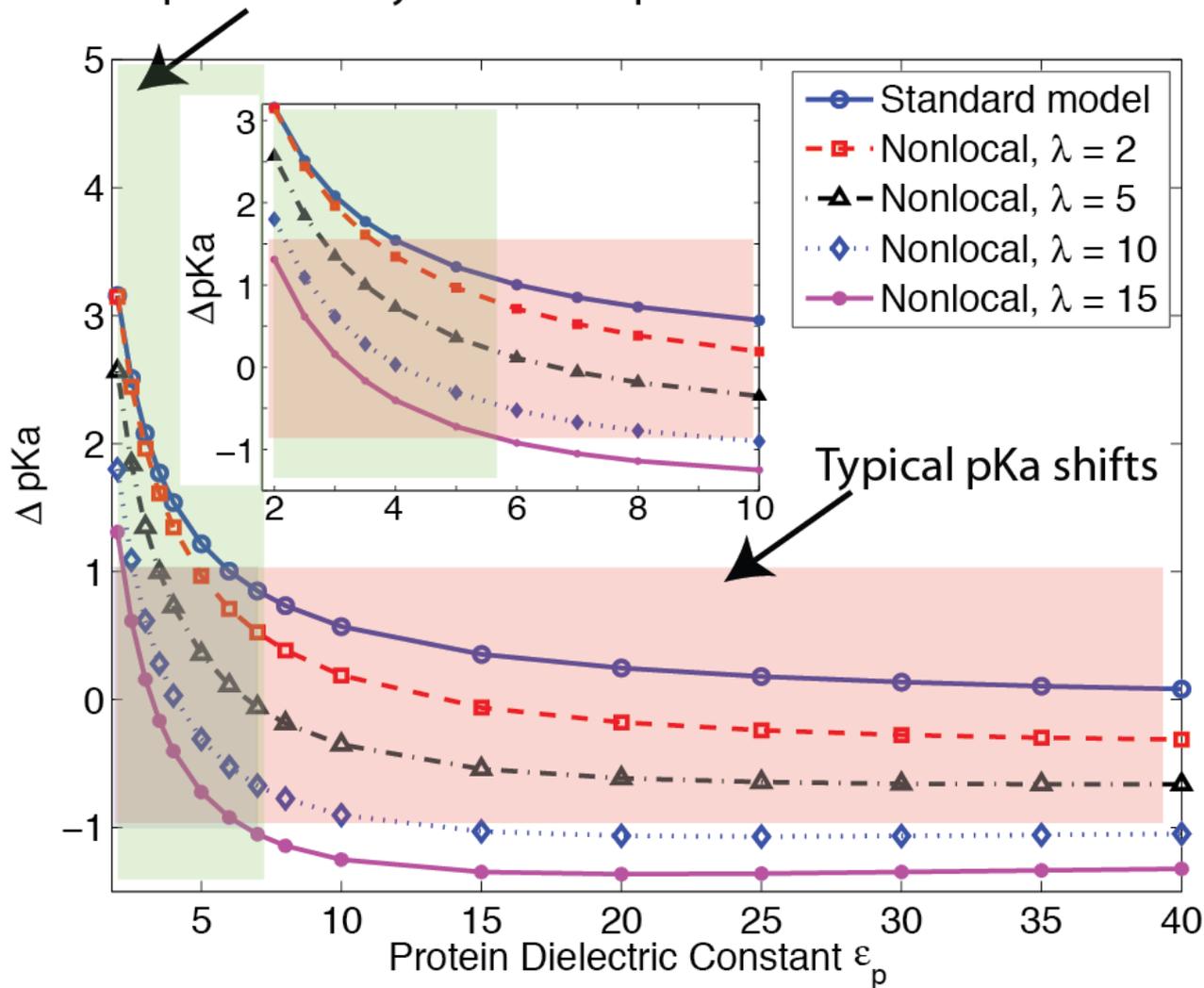
Nonlocal model



Nonlocal Results: New explanation for pKa controversy

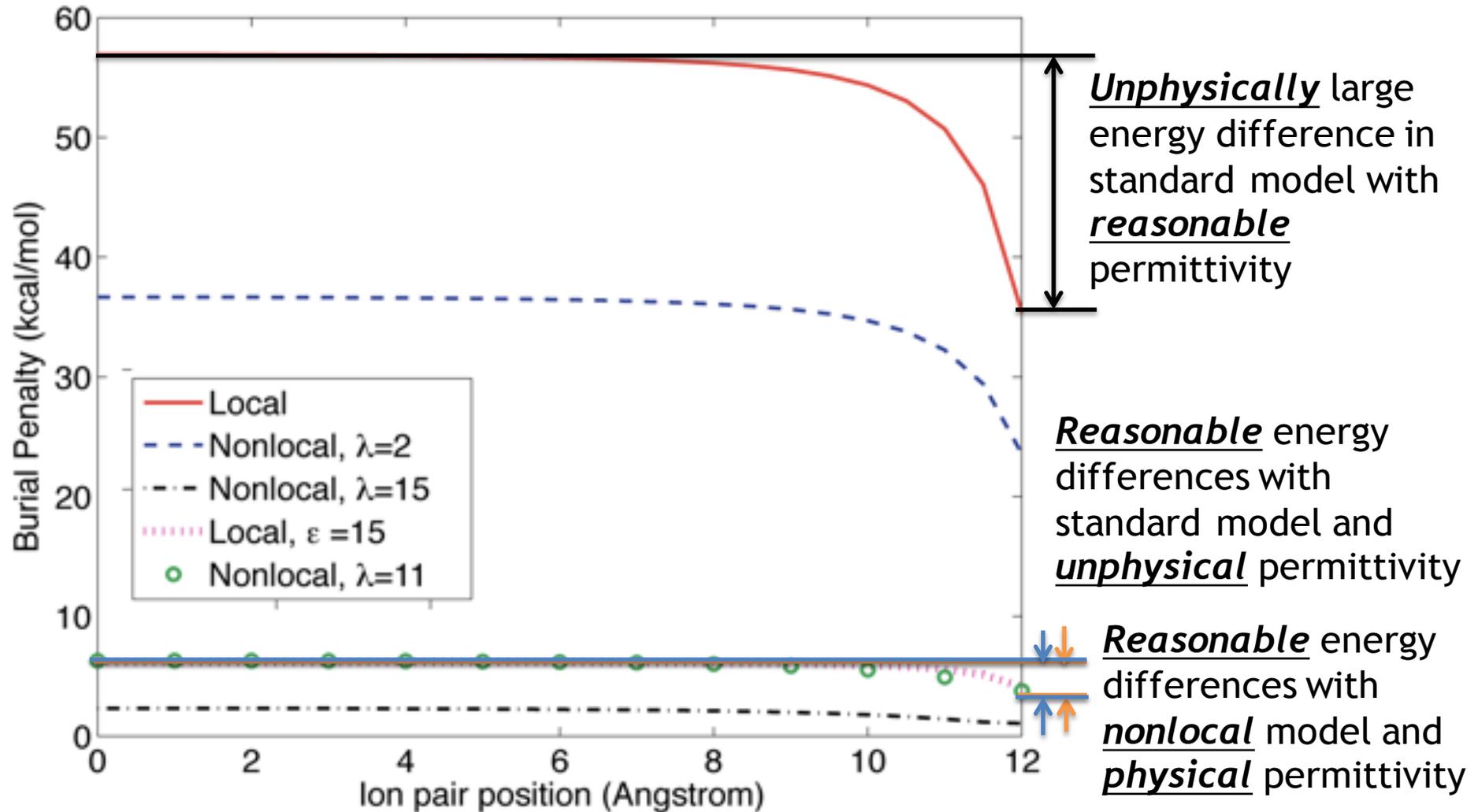
Get realistic answers using experimental dielectric constant

Experimentally measured protein dielectric constants

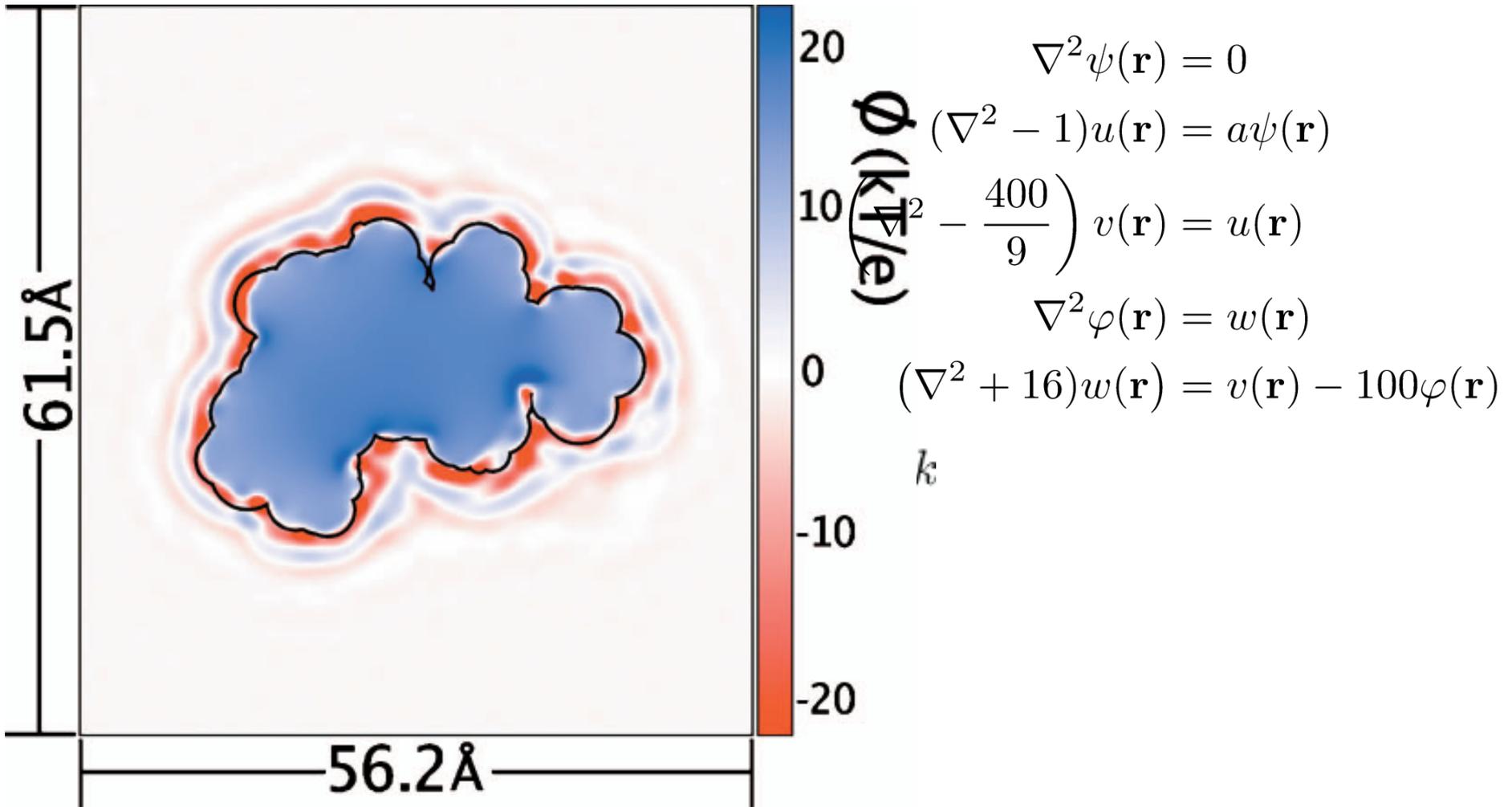


Are charge-burial prediction problems about flexibility or dielectric contrast?

Realistic parameters now give reasonable answers



Adding more realism (the water oscillations)



Nonlocal models from Landau-Ginzburg theory

□ Kornyshev LG theory leads to

$$h = A_1 P_1^2 + A_2 (\nabla P_1)^2 + B_1 P_2^2 + B_2 (\nabla P_2)^2 + C_1 \eta^2 + C_2 (\nabla \eta)^2 + \gamma P_2 \nabla \eta - (P_1 + P_2) D.$$

□ Medvedev added another coupling

$$\begin{aligned} F[\mathbf{P}_i(\mathbf{r}), \delta \mathbf{S}(\mathbf{r})] = & \frac{1}{2} \sum_{i=0}^2 \int \chi_{0i}^- \mathbf{1}(\mathbf{r} - \mathbf{r}') \mathbf{P}_i(\mathbf{r}) \mathbf{P}_i(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \\ & - \int \mathbf{P}(\mathbf{r}) \mathbf{E}_{\text{ext}}(\mathbf{r}) d\mathbf{r} + \frac{\mathbf{b}}{2} \int [\delta \mathbf{S}^2(\mathbf{r}) \\ & + l^2 (\nabla \delta S(\mathbf{r}))^2] d\mathbf{r} + \int \gamma(\mathbf{r} - \mathbf{r}') \mathbf{P}_2(\mathbf{r}) \nabla \delta \mathbf{S}(\mathbf{r}) d\mathbf{r}. \end{aligned}$$

First look: Poisson solvent with charge oscillations

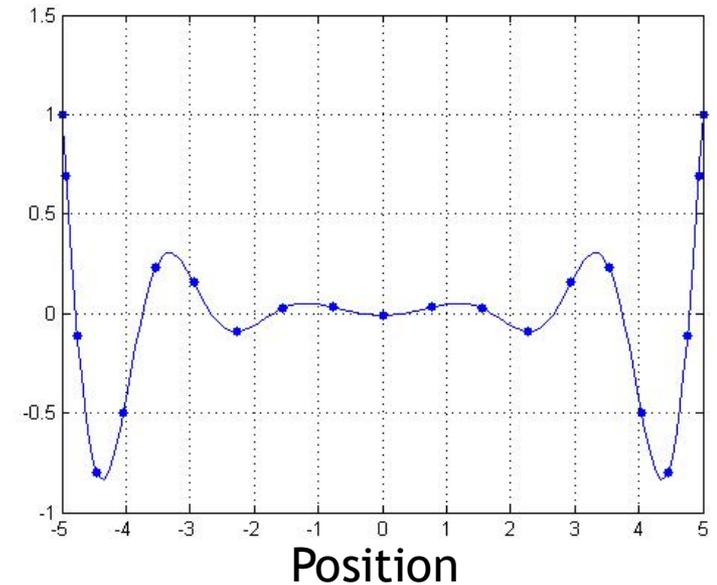
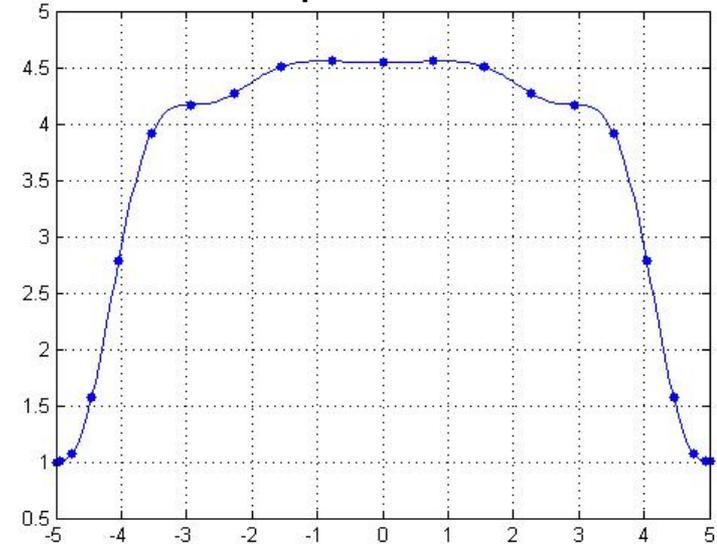
Electrostatic potential for both:
Dirichlet boundary conditions

Displacement potential:

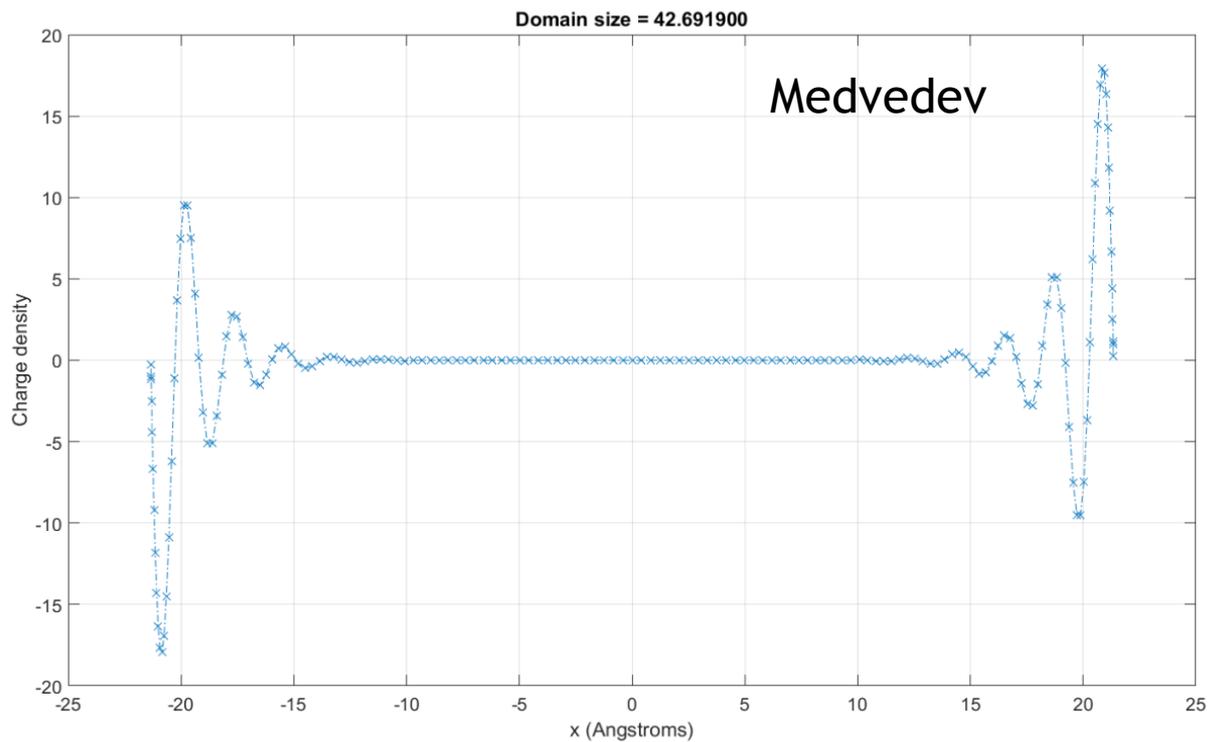
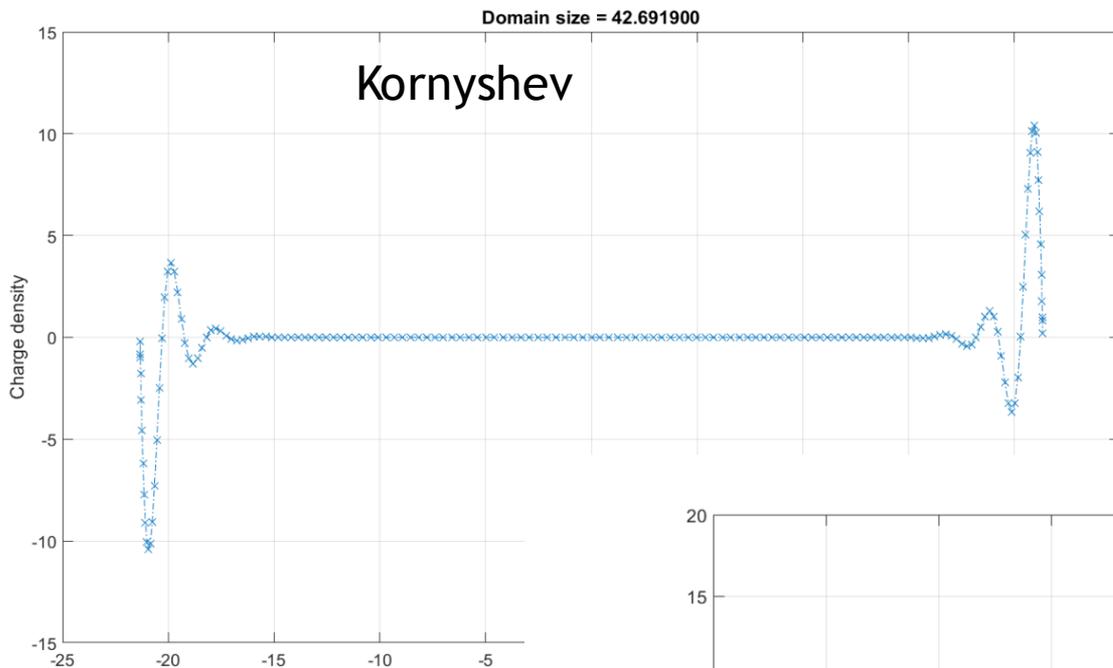
Lorentz nonlocal model
boundary conditions

ad hoc boundary conditions

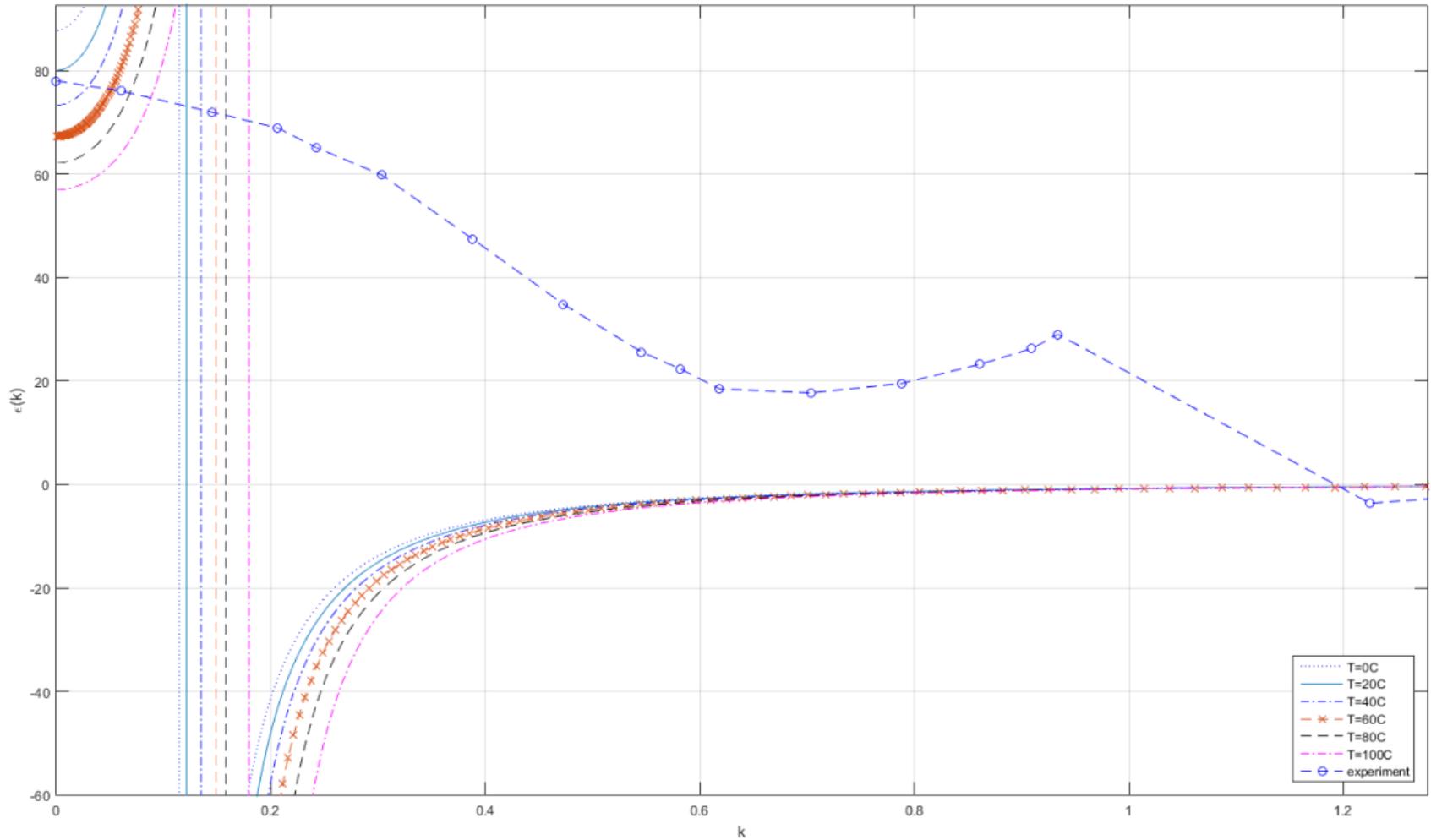
Electrostatic potential



Difference between Kornyshev+Medvedev



Temperature effects? Not so good.

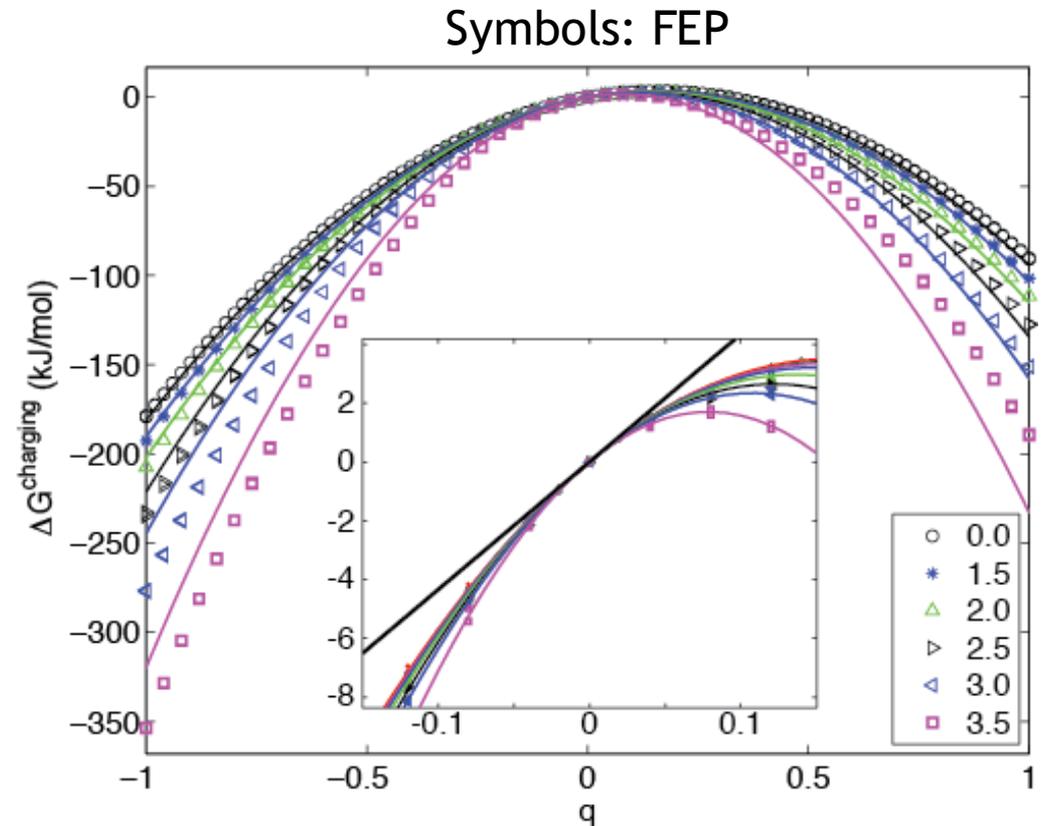
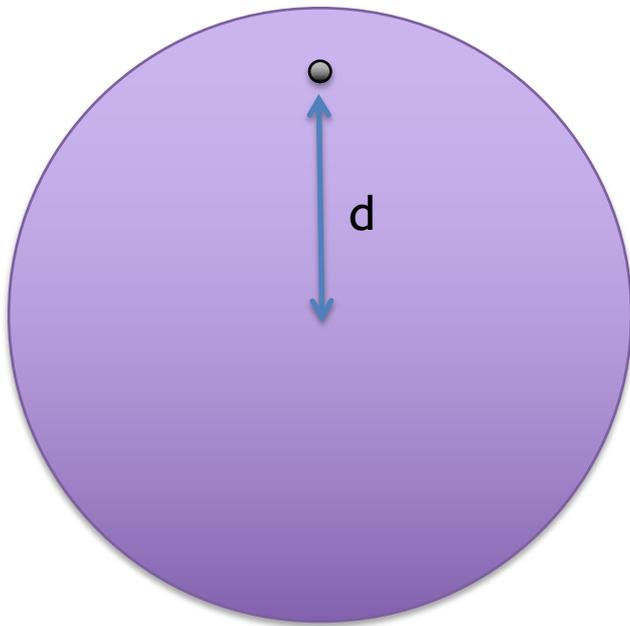


Parameterizing using explicit-solvent MD

□ Charging free-energy perturbation (FEP) calculations

Nonlocal models use correlation lengths from 2 to 40 Angstroms.

- Small values (1-4 Å) are appropriate for matching explicit solvent FEP
- **BUT charge-sign asymmetry dominates for surface charges**



Known theory failure: hydration asymmetry

FEBRUARY, 1939

JOURNAL OF CHEMICAL PHYSICS

VOLUME 7

The Free Energy of Hydration of Gaseous Ions, and the Absolute Potential of the Normal Calomel Electrode

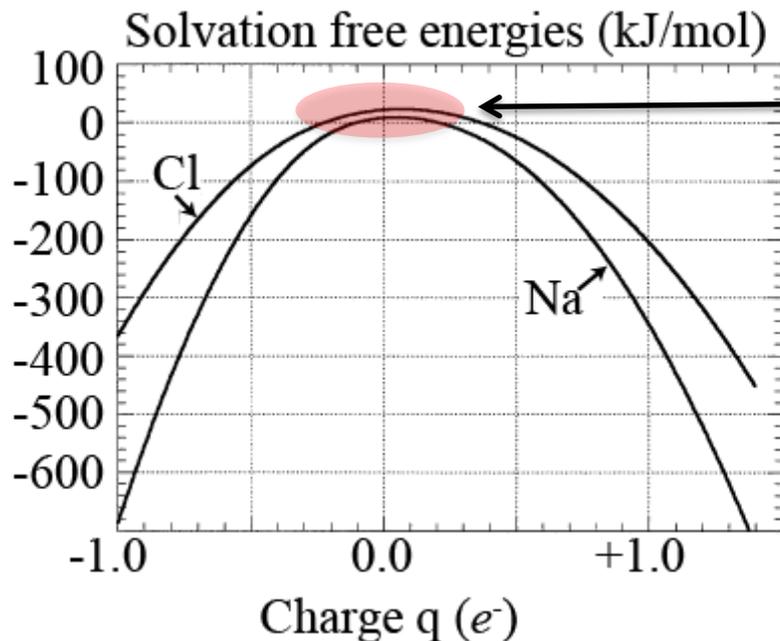
WENDELL M. LATIMER, KENNETH S. PITZER AND CYRIL M. SLANSKY

Department of Chemistry, University of California, Berkeley, California

(Received December 7, 1938)

Easy solution for ions:
Adjust radii

The free energies of hydration of the alkali and halide ions are found to agree reasonably well with the simple expression of Born ($-\Delta F = (1 - 1/D)Ne^2/2r_e$) for solution of charged spheres in a dielectric medium, provided the crystal radii are suitably modified so as to correspond to the radii of the cavities in the dielectric medium. The results show that the dielectric constant of water remains large even in the intense field next to the ion. The entropies of hydration are also found to be consistent with these radii. Because of the simplicity of this calculation, the resulting free energies of solution of individual ions are considered to be *a priori* the most probable and are used to calculate a value of -0.50 volt for the absolute potential of the calomel half-cell.



But: the slope is not zero at $q=0$

Lynden-Bell et al. '01

Outline

- ❑ Not all multiscale models are created equal
- ❑ Hydration-shell Poisson-Boltzmann model
- ❑ HSPB applications and extensions
- ❑ Open questions and possible directions

Continuum model's failure is simple for ions..

FEBRUARY, 1939

JOURNAL OF CHEMICAL PHYSICS

VOLUME 7

The Free Energy of Hydration of Gaseous Ions, and the Absolute Potential of the Normal Calomel Electrode

WENDELL M. LATIMER, KENNETH S. PITZER AND CYRIL M. SLANSKY

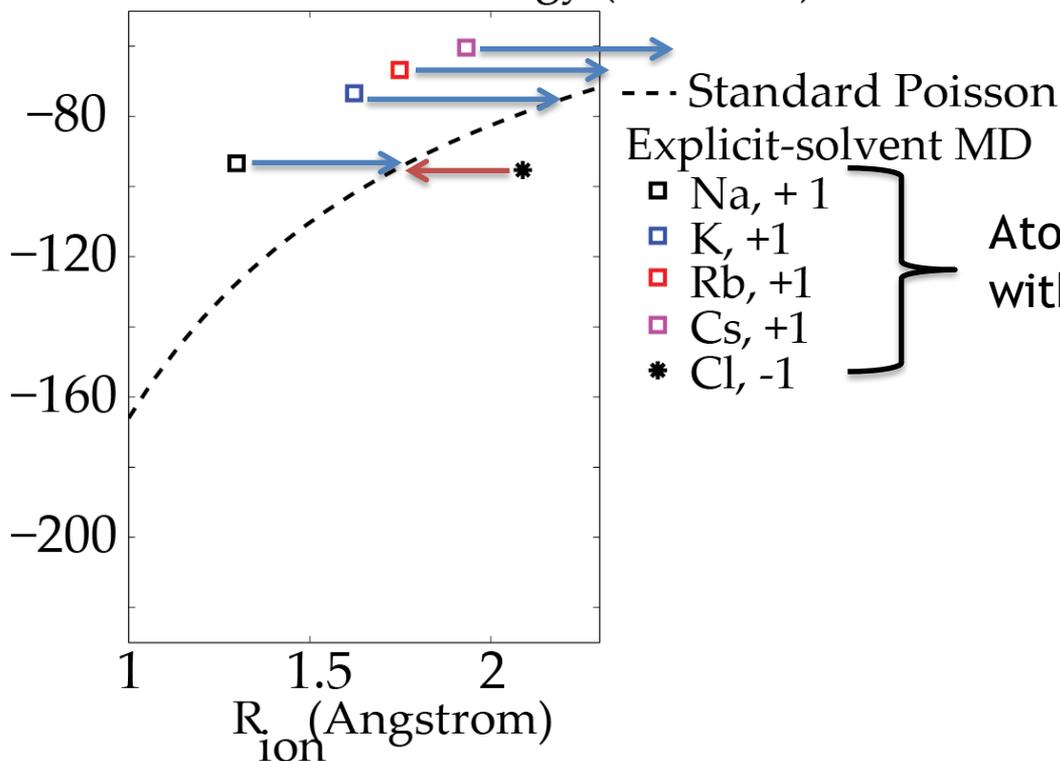
Department of Chemistry, University of California, Berkeley, California

(Received December 7, 1938)

Easy solution for spheres:
Adjust their radii!

The free energies of hydration of the alkali and halide ions are found to agree reasonably well with the simple expression of Born ($-\Delta F = (1 - 1/D)Ne^2/2r_e$) for solution of charged spheres in a dielectric medium, provided the crystal radii are suitably modified so as to correspond to the radii of the cavities in the dielectric medium. The results show that the dielectric constant of water remains large even in the intense field next to the ion. The entropies of hydration are also found to be consistent with these radii. Because of the simplicity of this calculation, the resulting

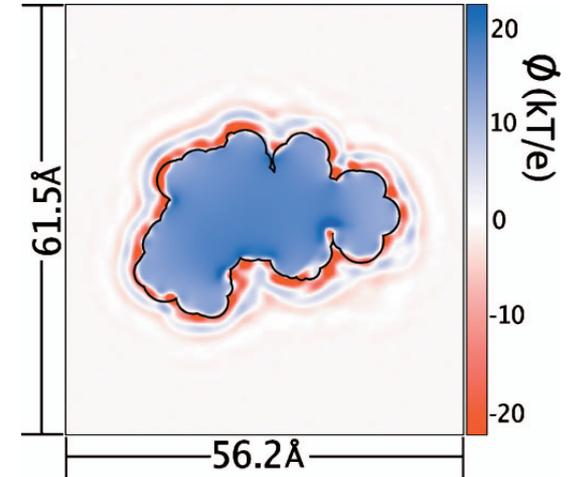
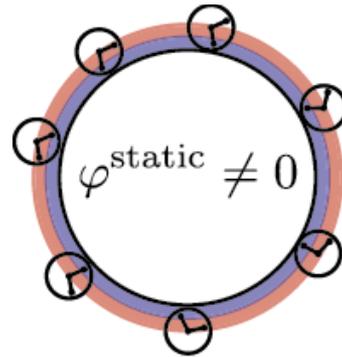
Electrostatic free energy (kcal/mol)



Atomistic simulations as “computational microscope”

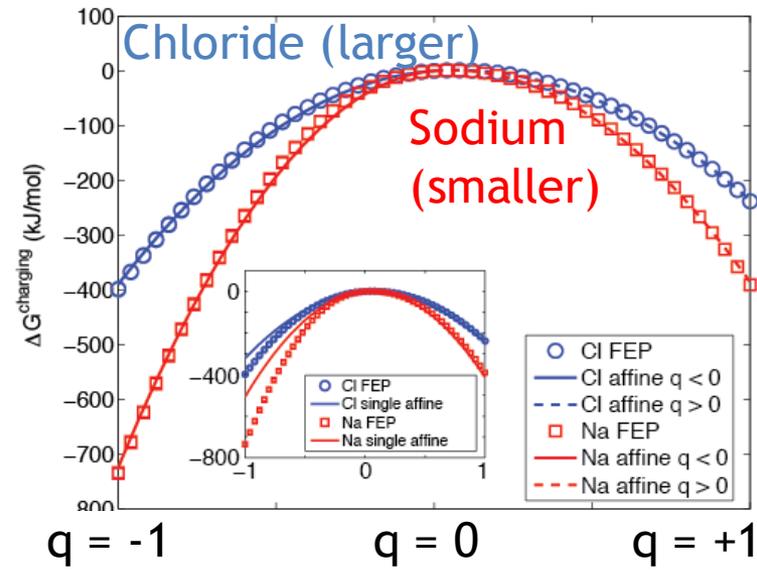
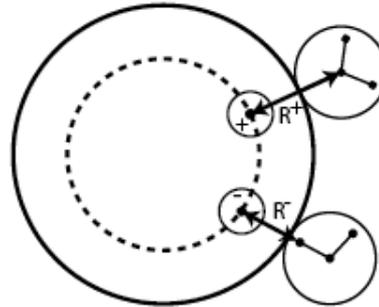
1. Surface potential:

Liquid-vapor interface potential exists *even in the absence of solute charge*



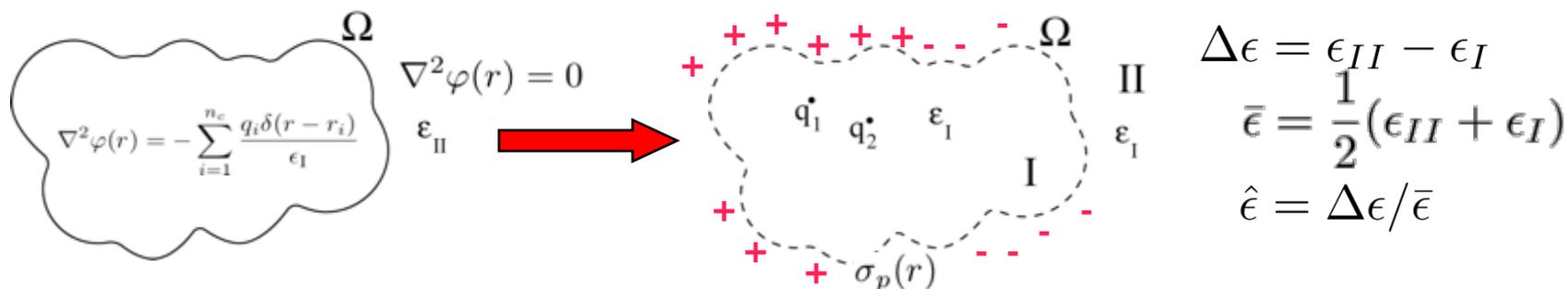
2. Hydrogen-oxygen size difference:

Protein surface charges see different “closest approaches” depending on their sign



Bardhan, Jungwirth, Makowski '12

The basic boundary-integral equation model



$$\sigma_p(\mathbf{s}) + \frac{\Delta\epsilon(\mathbf{s})}{4\pi\bar{\epsilon}(\mathbf{s})} \mathbf{n}(\mathbf{s}) \cdot \int_{\Omega} \frac{\mathbf{s} - \mathbf{s}'}{|\mathbf{s} - \mathbf{s}'|^3} \sigma_p(\mathbf{s}') d\mathbf{s}' = -\frac{\Delta\epsilon(\mathbf{s})}{4\pi\bar{\epsilon}(\mathbf{s})} \mathbf{n}(\mathbf{s}) \cdot \sum_k \frac{q_k}{\epsilon(\mathbf{r}_k)} \frac{\mathbf{s} - \mathbf{r}_k}{|\mathbf{s} - \mathbf{r}_k|^3}$$

$$(\mathcal{I} - \hat{\epsilon}\mathcal{K}) \sigma = \hat{\epsilon}G_n q$$

$$\varphi^{\text{REAC}} = \int_S \frac{1}{\|r - r'\|} \sigma(r') dA'$$

$$E^{\text{REAC}} = \frac{1}{2} q^T C A^{-1} B q$$

Tomasi, 1981
Shaw, 1985
Zauhar, 1988

Key modeling step: Think of a Born ion. Stop taking the boundary condition as a given...

GIVEN

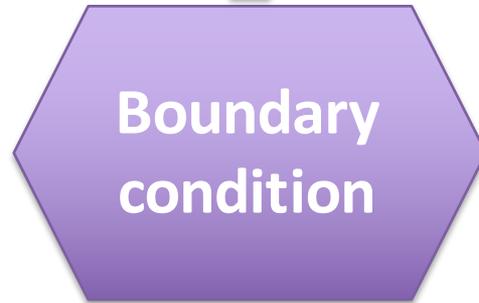


From MD or
experiment

From macroscopic
dielectric theory

$$\Delta G^{es} = \frac{1}{2} q_{\text{ion}} R_{\text{ion}} \sigma$$

GIVEN



$$\sigma^{\text{SMBC}} = \frac{E^{\text{Coul}}}{1 + \frac{\epsilon_{\text{III}}}{\epsilon_{\text{II}} - \epsilon_{\text{III}}}}$$

DERIVED



$$\varphi_{\text{center}} = \int \frac{\sigma(r')}{4\pi|r-r'|} dA'$$

$$\varphi_{\text{center}} = R_{\text{ion}} \sigma$$

Note: written here for a spherical ion (the boundary integral operator simplifies due to symmetry)

...and ask, if the physically-based radius is the given, what is the boundary condition?

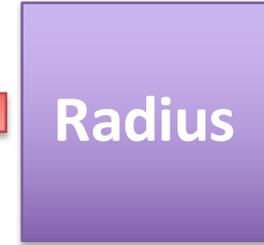
GIVEN

GIVEN

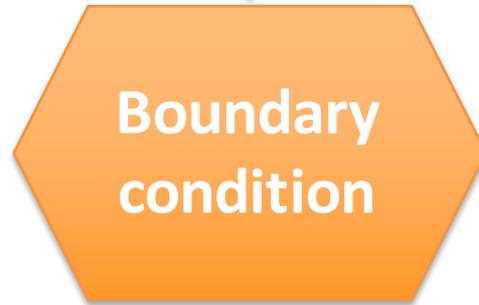


From MD or experiment

$$\Delta G^{es} = \frac{1}{2} q_{\text{ion}} R_{\text{ion}} \sigma$$



DERIVED



$$\sigma = g(E^{\text{Coul}}) \quad \text{More general expression}$$

$$\varphi_{\text{center}} = \int \frac{\sigma(r')}{4\pi|r-r'|} dA'$$
$$\varphi_{\text{center}} = R_{\text{ion}} \sigma$$

Why boundary-integral equation modeling?

BIE for standard Maxwell BC

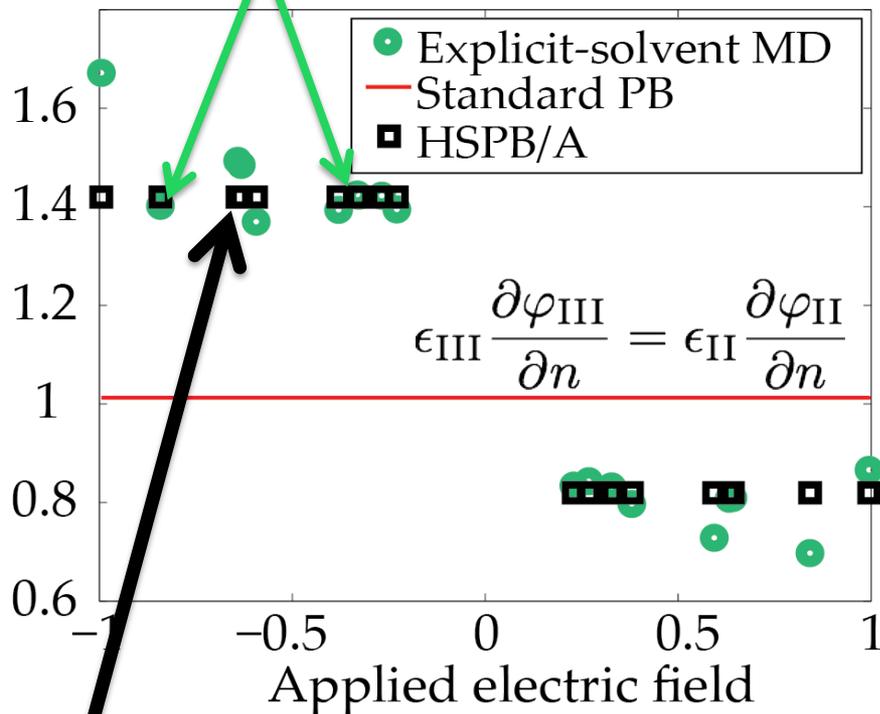
$$\left(I + \frac{\epsilon_{\text{III}} - \epsilon_{\text{II}}}{\epsilon_{\text{II}}} \left(-\frac{1}{2}I + K \right) \right) \sigma = E_n$$

BIE for proposed nonlinear BC

$$\left(I + \frac{\epsilon_{\text{III}} - \epsilon_{\text{II}}}{\epsilon_{\text{II}}} \left(-\frac{1}{2}I + K \right) + h(E_n) \right) \sigma = E_n$$

$$h(E_n) = \alpha \tanh(\beta E_n - \gamma) + \mu$$

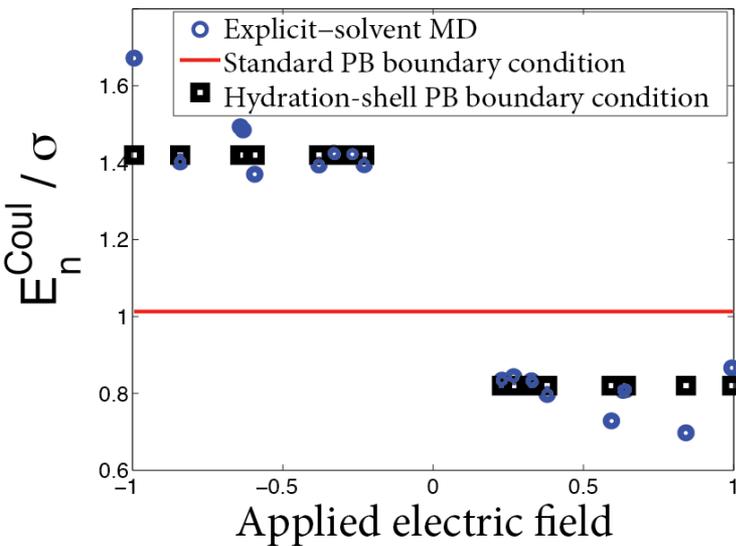
Atomistic calc's



$$\left(\frac{\epsilon_{\text{II}}}{\epsilon_{\text{II}} - \epsilon_{\text{III}}} - \alpha \tanh(\beta E_n - \gamma) + \mu \right) \frac{\partial \varphi_{\text{I}}}{\partial n} = \left(\frac{\epsilon_{\text{III}}}{\epsilon_{\text{II}} - \epsilon_{\text{III}}} - \alpha \tanh(\beta E_n - \gamma) + \mu \right) \frac{\partial \varphi_{\text{I}}}{\partial n}$$

Small wonder that the PDE modelers didn't come to this on their own!

Modifying the boundary condition



$$\epsilon_{in} \frac{d\phi_{in}}{dn} = \epsilon_{out} \frac{d\phi_{out}}{dn}$$

Add nonlinear correction

$$(\epsilon_{in} - \Delta\epsilon h(E_n)) \frac{d\phi_{in}}{dn} = (\epsilon_{out} - \Delta\epsilon h(E_n)) \frac{d\phi_{out}}{dn}$$

$$\Delta\epsilon = \epsilon_{out} - \epsilon_{in}$$

$$h(E_n) = \alpha \tanh(\beta E_n - \gamma) - \alpha \tanh(-\gamma)$$

From more than 50 parameters to just 4!

□ The NLBC model has only 4 fitting parameters:

1. α : magnitude of the asymmetry
2. β : width of the asymmetry transition
3. γ : water's "intrinsic" orientational preference
4. ξ : uniform scaling factor applied to all MD radii

□ Contrast to standard symmetric models:

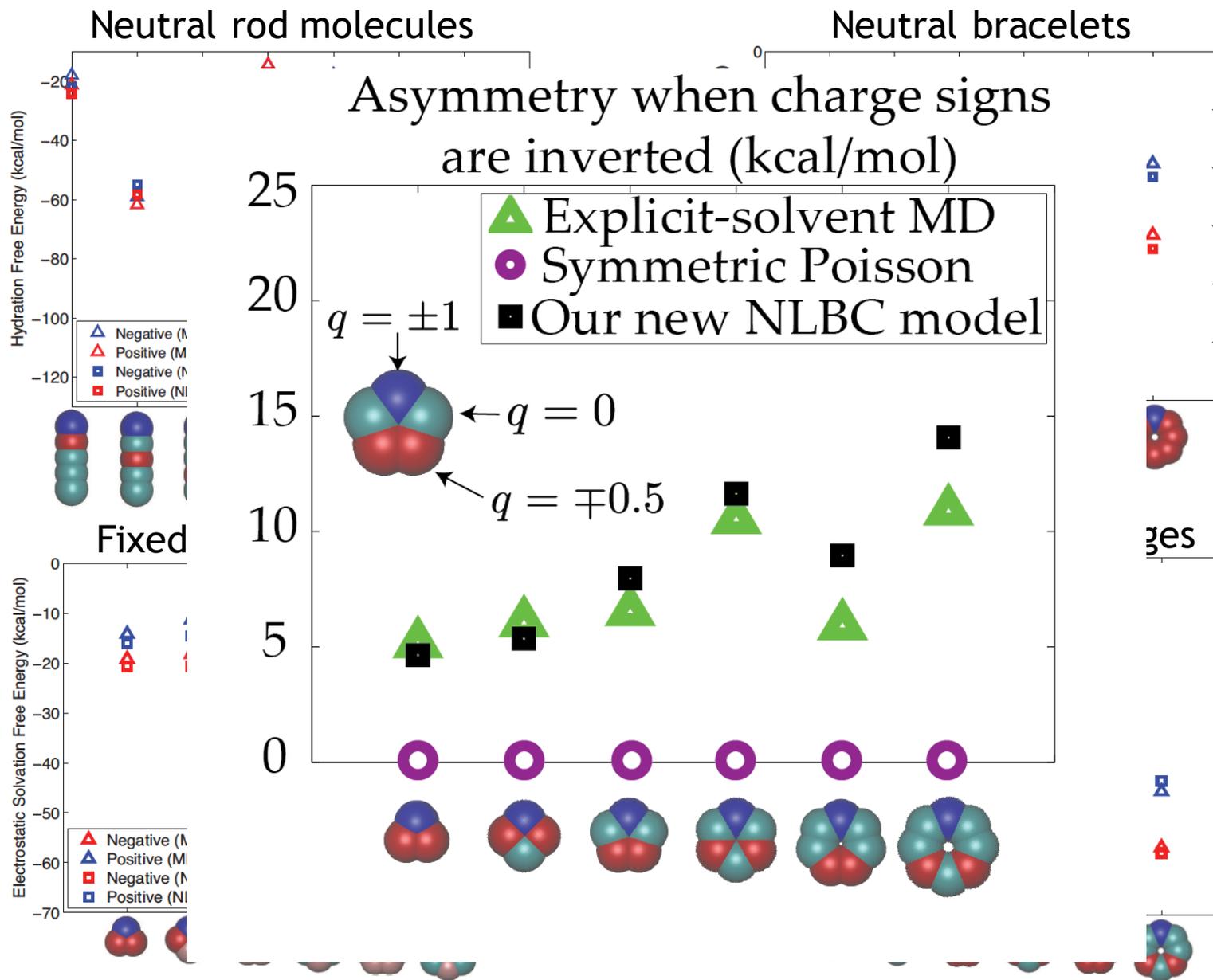
□ Our 4 parameters were fit against 52 unphysically hard test problems covering asymmetric solvation

TABLE 2: Atomic Born Radii Derived from Solvent Electrostatic Charge Distribution Tested with Free Energy Perturbation Methods in an Explicit Solvent^a

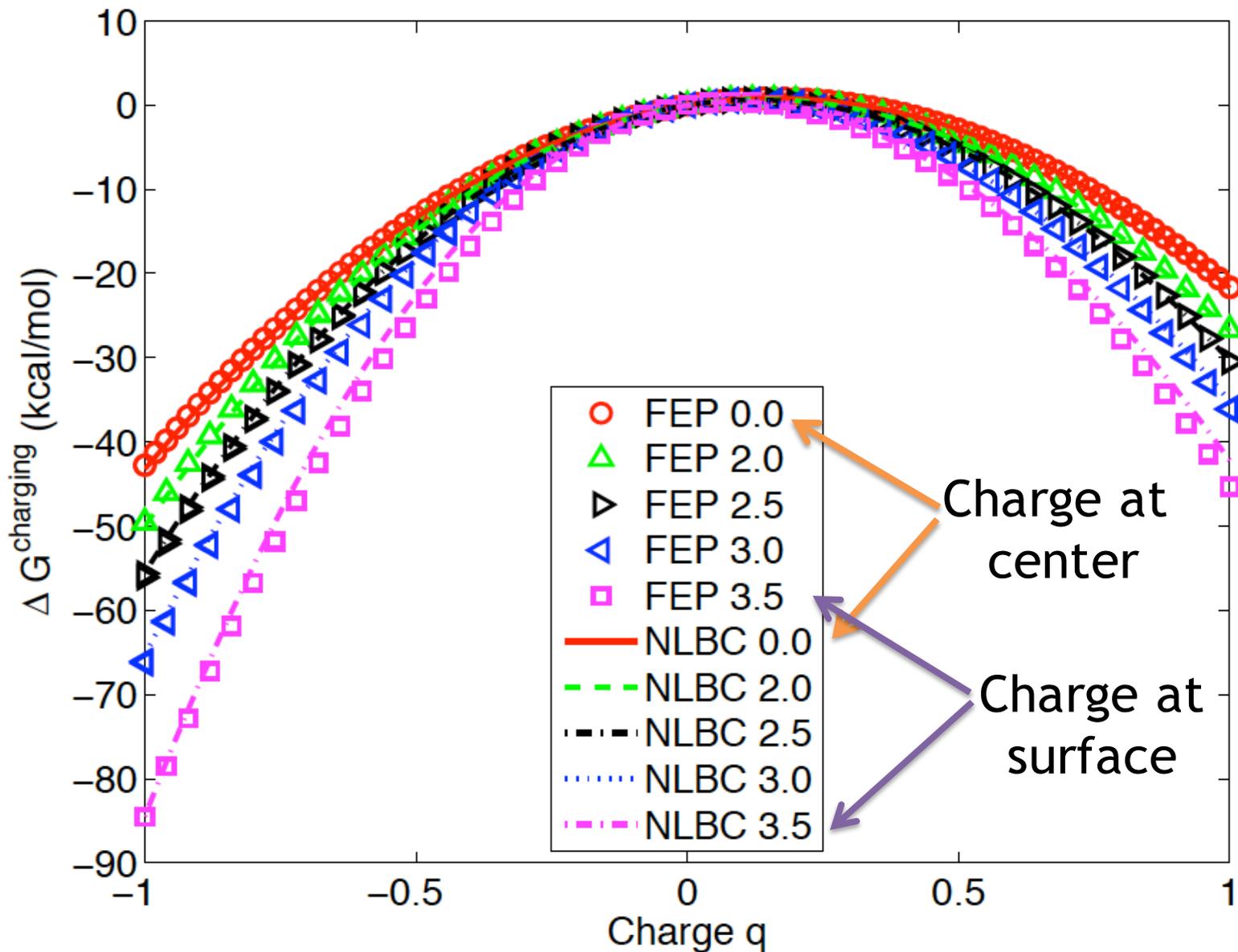
atom	radius (Å)
Backbone	
C	2.04 carbonyl C, peptide backbone
O	1.52 carbonyl oxygen
	2.23 peptide nitrogen
CA	2.86 all CA except Gly
CA	2.38 Gly only
Hydrogens	
H*	0.00 all hydrogens
Side Chains	
CB	2.67 all residues
CG*	2.46 Val, Ile, Arg, Lys, Met, Phe, Thr, Trp, Gln, Glu
CD*	2.44 Ile, Leu, Arg, Lys
CD, CG	1.98 Asp, Glu, Asn, Gln
CB, CG, CD	1.98 Pro only
CE*, CD*, CZ,	2.00 Tyr, Phe rings
CE*, CD*, CZ*, CH2	1.78 Tip ring only
CE	2.10 Met only
CZ, CE	2.80 Arg, Lys
OE*, OD*	1.42 Glu, Asp, Asn, Gln
OG*	1.64 Ser, Thr
OH	1.85 Tyr only
NH*, NE, NZ	2.13 Arg, Lys
NE2, ND2	2.15 Gln, Asn
NE2, ND1	2.31 His only
NE1	2.40 Trp
S*	2.00 Met, Cys

^a Patches N-term and C-term CAT, CAY: 2.06 Å. CY: 2.04 Å. OY: 1.52 Å. NT: 2.23 Å. * refers to a wild card character.

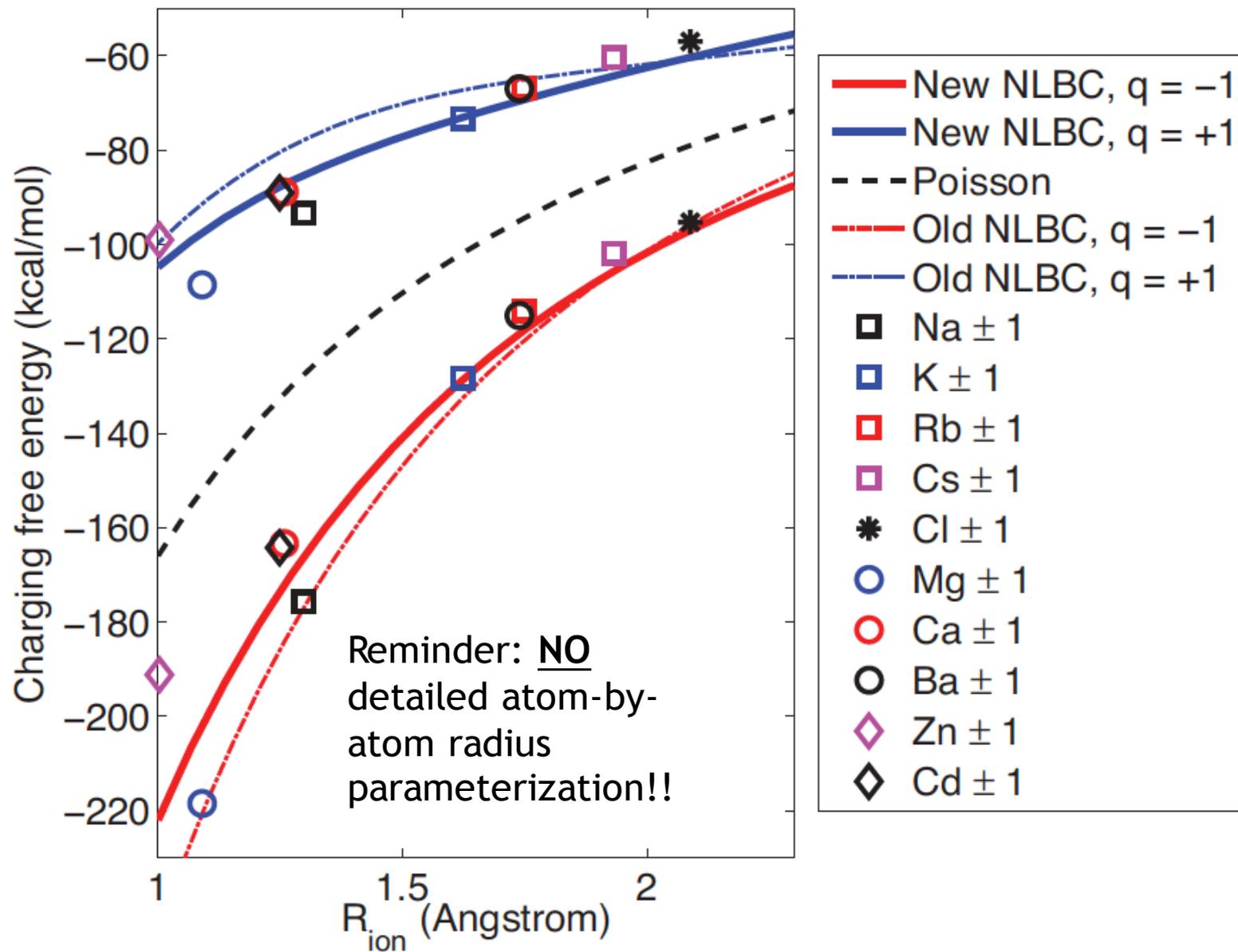
NLBC model is accurate for many hard problems



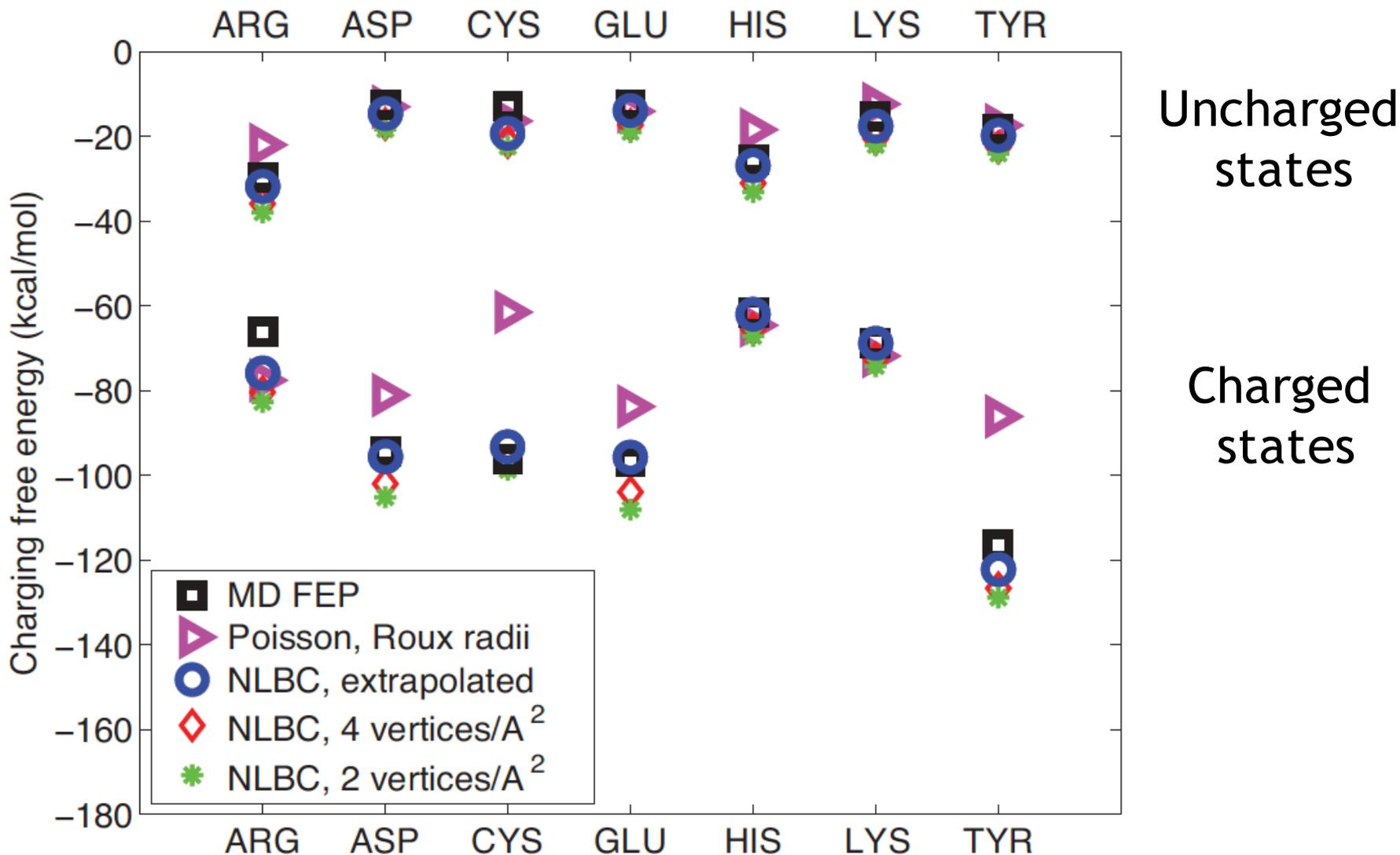
Accurate Energies in Sphere



Accurate Ion Energies



Returning to the pKa problem



Outline

- ❑ Not all multiscale models are created equal
- ❑ Hydration-shell Poisson-Boltzmann model
- ❑ HSPB applications and extensions
- ❑ Open questions and possible directions

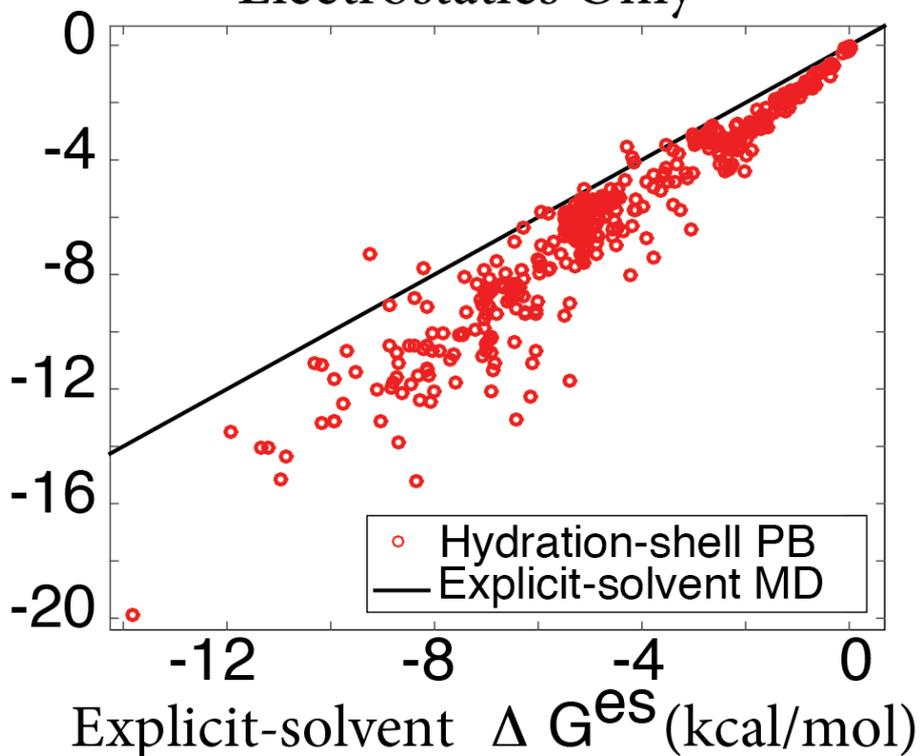
Applications of original HSPB

1. Parameterization approaches
2. Test set of 500 small molecules (experiment + MD)
3. Single-atom charging free energies in amino acids

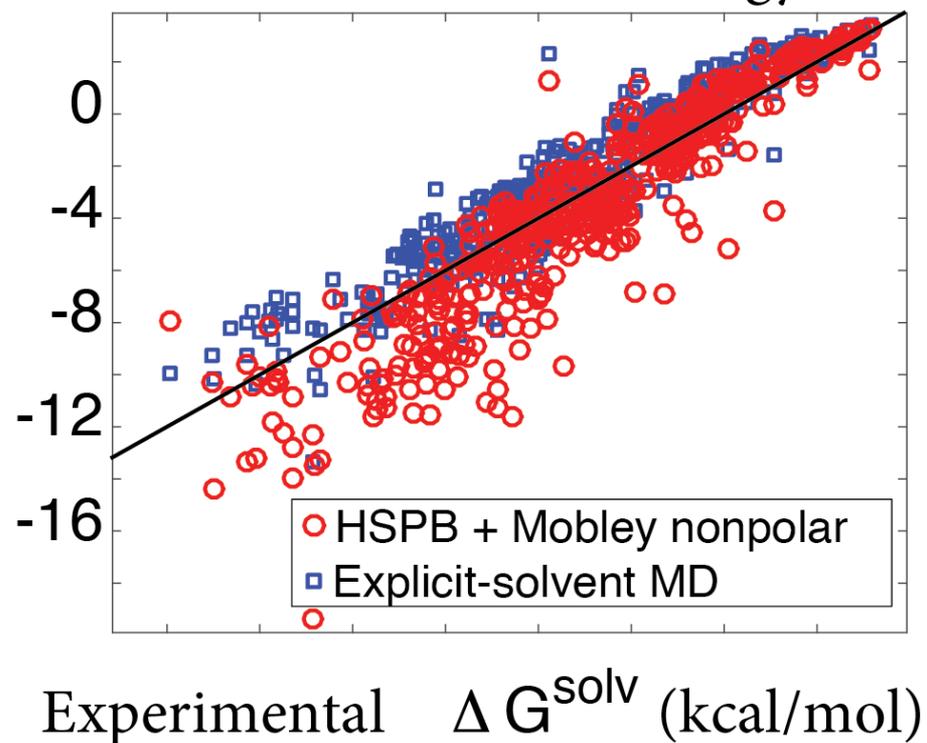
Test set of 500 small molecules

Parameterized asymmetric HSPB using 6 monovalent ions and 6 amino-acid side chain analogues

Electrostatics Only



Total Solvation Free Energy

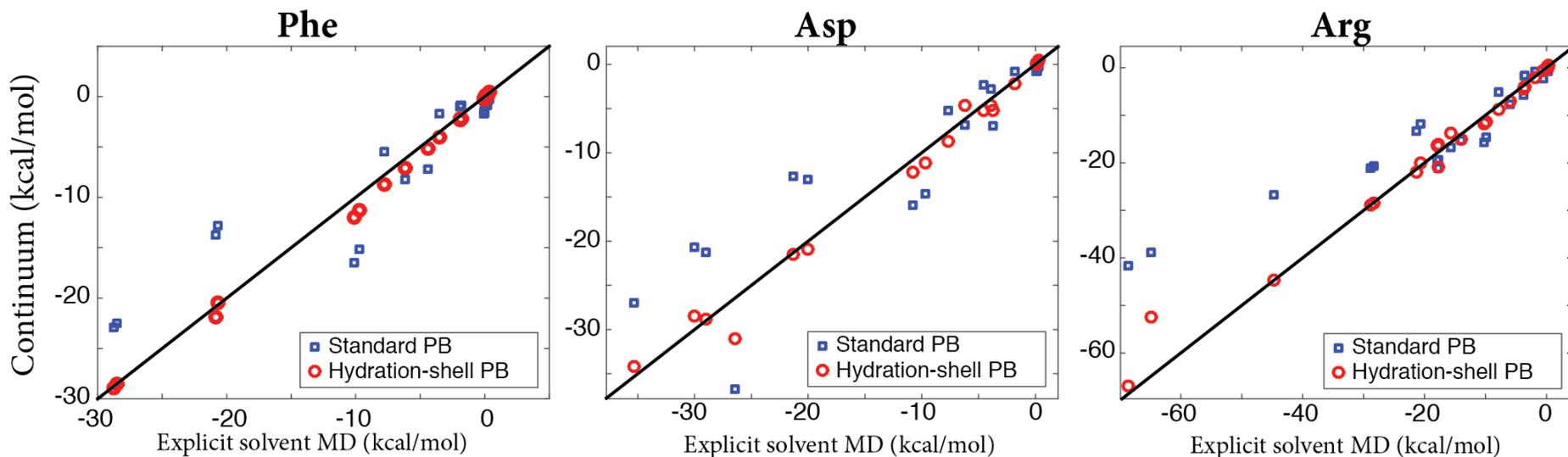


Reminder: No atomic radii have been optimized

Note: Dominant errors are associated with oxygens. Further investigation underway!

Single-atom charging free energies

- A more detailed window into the reaction potential operator



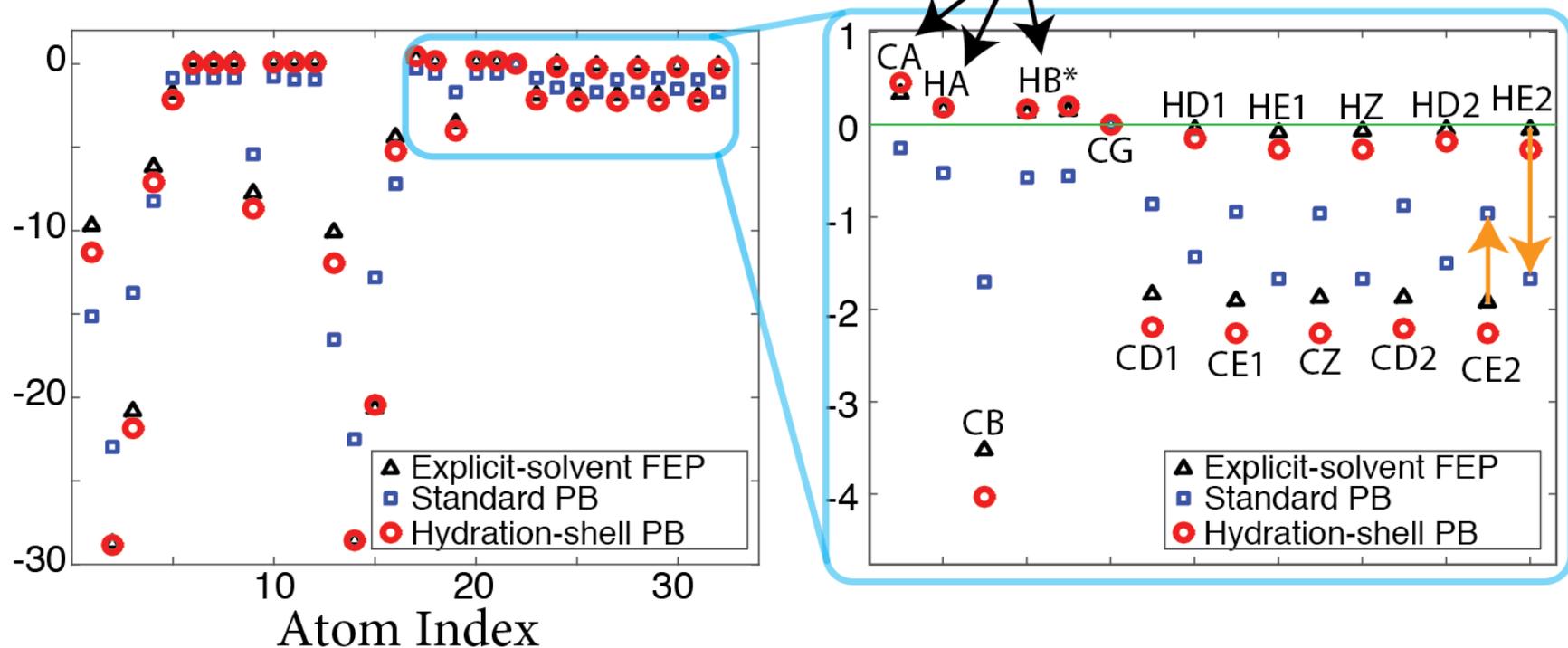
Standard PB: using Roux radii

HSPB: No radii fitted

The standard PB theory obtains correct total solvation energies through compensating errors!

More on single-atom charging free energies

- Looking specifically at phenylalanine



Extension: the Mean Spherical Approximation (MSA) defines a different HSPB

Looking at the MSA expression for Born ion solvation free energy

	ΔG^{Born}	Electric flux boundary condition
Standard	$-\frac{N_a}{4\pi\epsilon_0} \cdot \frac{1}{2} \left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_{out}} \right) \frac{q^2}{R}$	$\epsilon_{in} \frac{d\phi_{in}}{dn} = \epsilon_{out} \frac{d\phi_{out}}{dn}$
MSA	$-\frac{N_a}{4\pi\epsilon_0} \cdot \frac{1}{2} \left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_{out}} \right) \frac{q^2}{R + \delta_s}$	$(\epsilon_{in} - \Delta\epsilon h(E_n)) \frac{d\phi_{in}}{dn} = (\epsilon_{out} - \Delta\epsilon h(E_n)) \frac{d\phi_{out}}{dn}$

Radius perturbation from MSA:
 $\delta_s = \delta_s(\epsilon_{out}(T), r_s)$

Solvent size parameter, also from MSA: for example, for water $r_s = 1.42 \text{ \AA}$

$\Delta\epsilon = \epsilon_{out} - \epsilon_{in}$

MSA hydration-shell correction
 $h(E_n) = \delta_s \sqrt{|E_n|}$

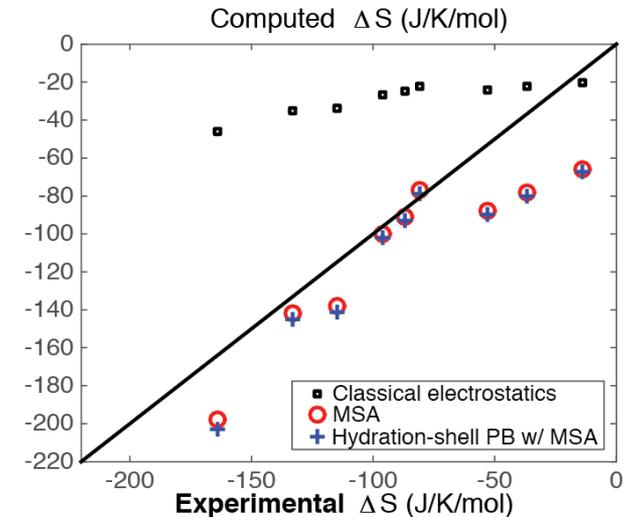
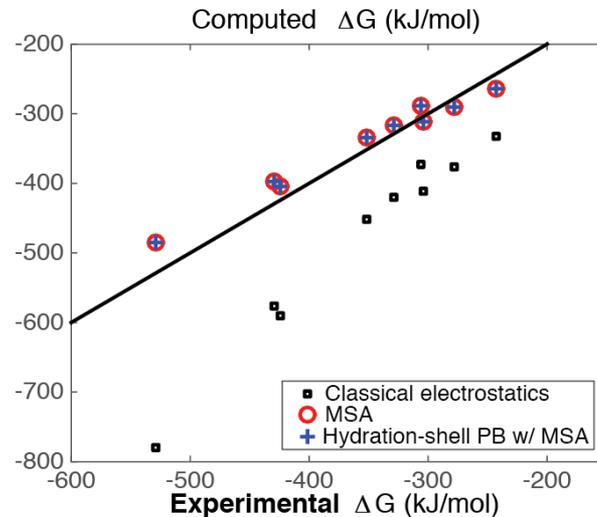
No free parameters

HSPB+MSA=

Poisson-based solvation thermodynamics

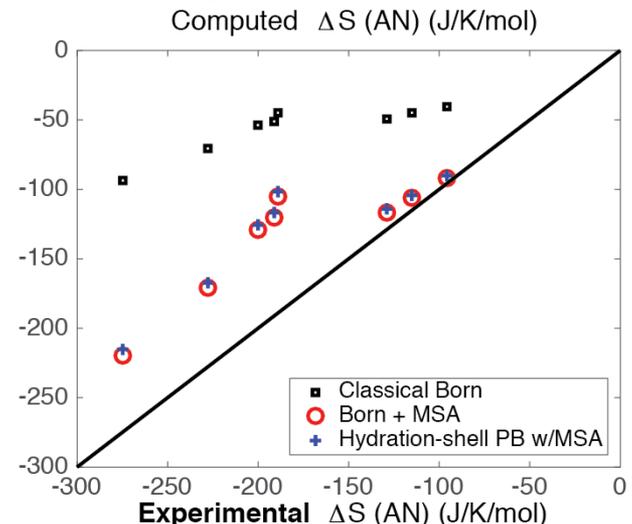
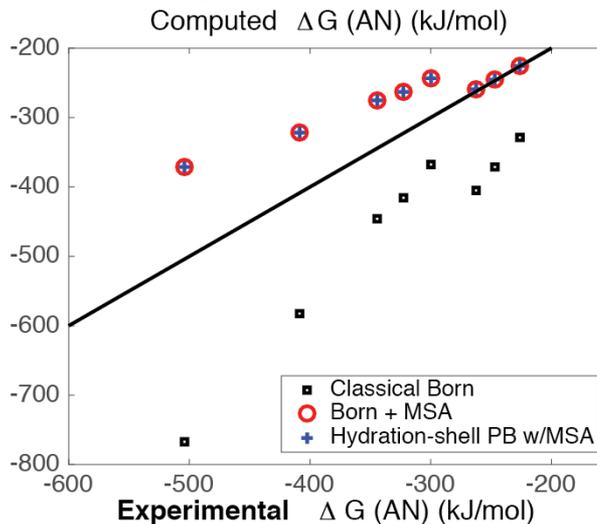
□ And not just in protic solvents

Water at 298 K,
using Shannon-Prewitt
radii



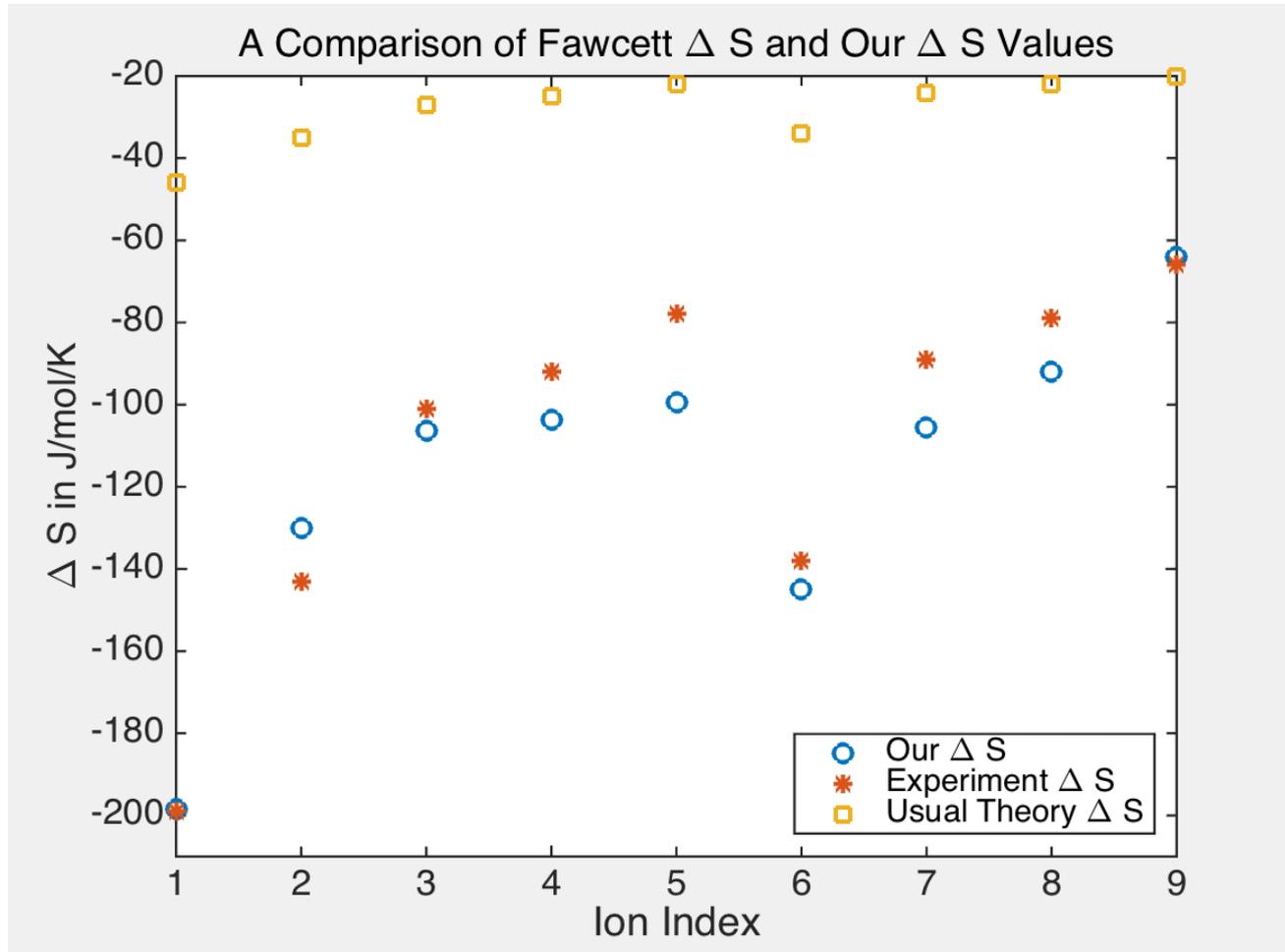
Why HSPB imp:
Acetonitrile at 298 K,
using Shannon-Prewitt
radii

$$\Delta G^{classical}(T) = -\frac{z^2}{4}$$
$$\Delta G^{MSA}(T) = -\frac{z^2}{4}$$



Temperature-dependent asymmetric HSPB

□ Build on success of temp-dependent HSPB+MSA

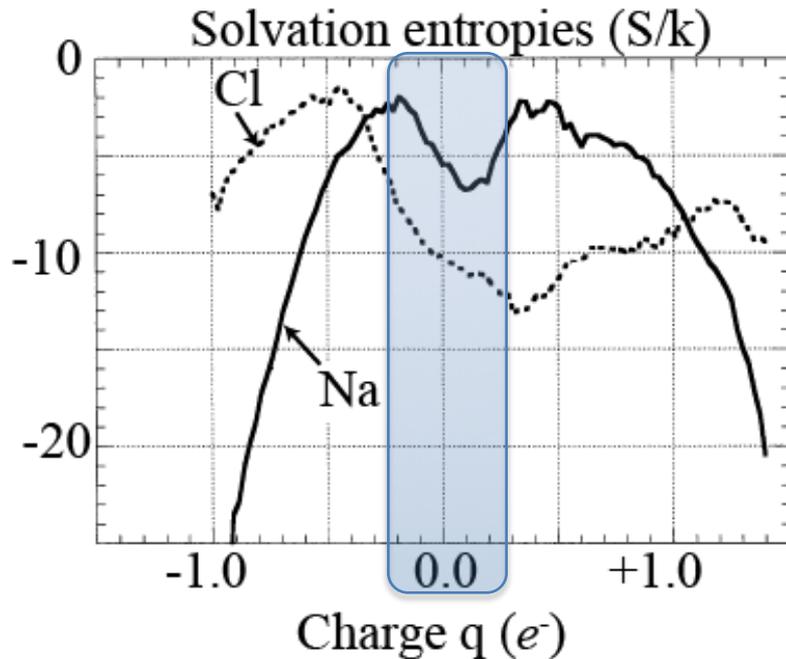


Outline

- ❑ Not all multiscale models are created equal
- ❑ Hydration-shell Poisson-Boltzmann model
- ❑ HSPB applications and extensions
- ❑ Open questions and possible directions

What does the new boundary condition miss?

- ❑ Still a single scale theory—no charge oscillations
- ❑ Can't model actual dielectric saturation
- ❑ How should the NLBC results be analyzed in terms of solution thermodynamics?



Lynden-Bell et al. 2001

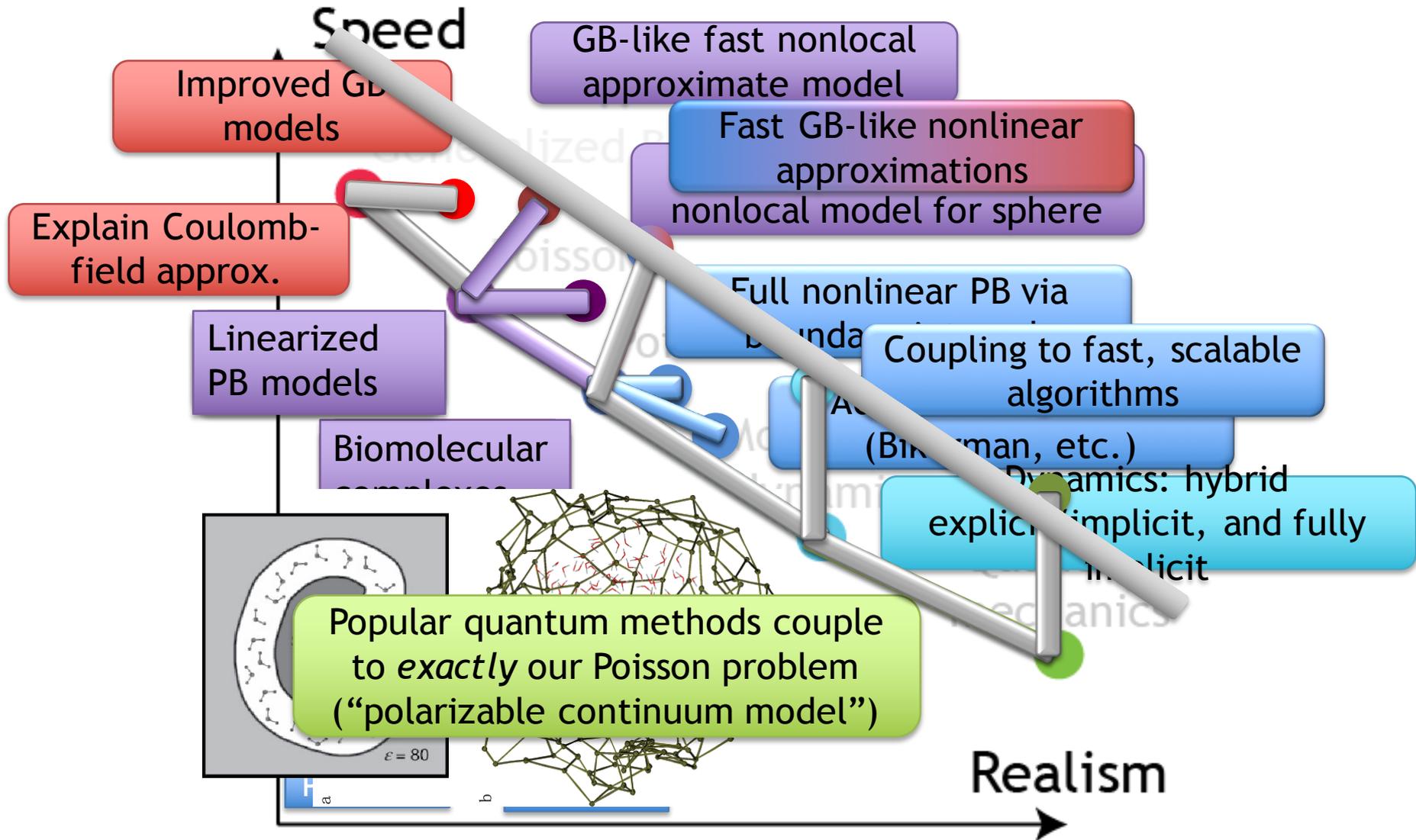
Does the small q region inform the width of the nonlinear transition region?

Other future directions

- Obtain HSPB from volumetric models e.g. RISM?
- Incorporate into electronic structure methods (PCM)
- Adding a nonpolar term (SPT? Hummer's info. theory?)
- Calibrate temperature and pressure dependence capability
- Solution specific electrolyte BC
- Implicit-solvent molecular dynamics via BEM-based GB
- Extensions for more sophisticated nonlinear BC functions

Future directions

Advancing a sustainable modeling framework



A fast, rigorous Generalized Born

Popular heuristic model: Generalized Born

$$\hat{L}_{ii}^{GB} = q_i C (I - \mathbf{0})^{-1} B q_i$$

\hat{L}_{ij}^{GB} via interpolation

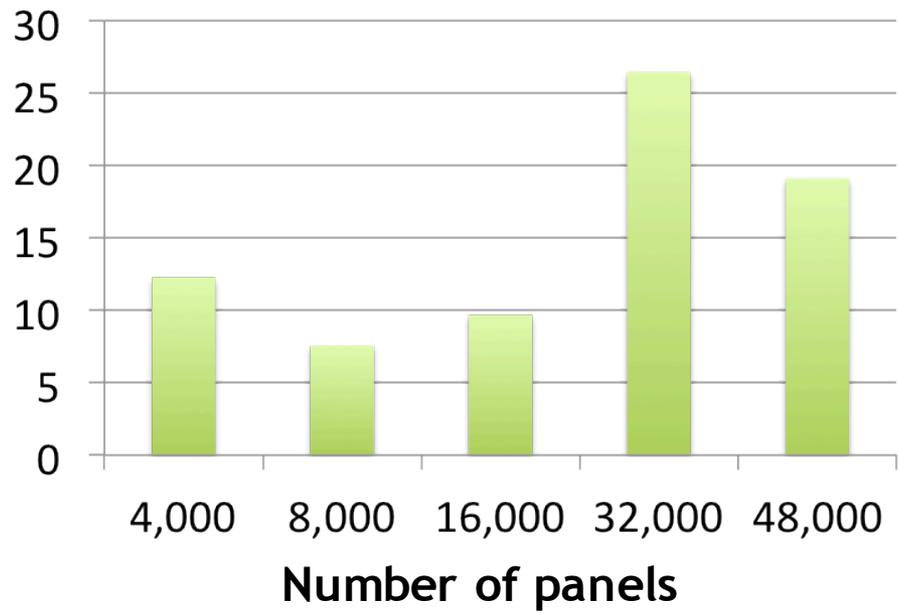


New discretization techniques reduce the size of B and $C \rightarrow$ further speedup

$$\hat{L} = C (I - \hat{U} \hat{\Sigma} \hat{V}^T)^{-1} B$$

Approximate inverse

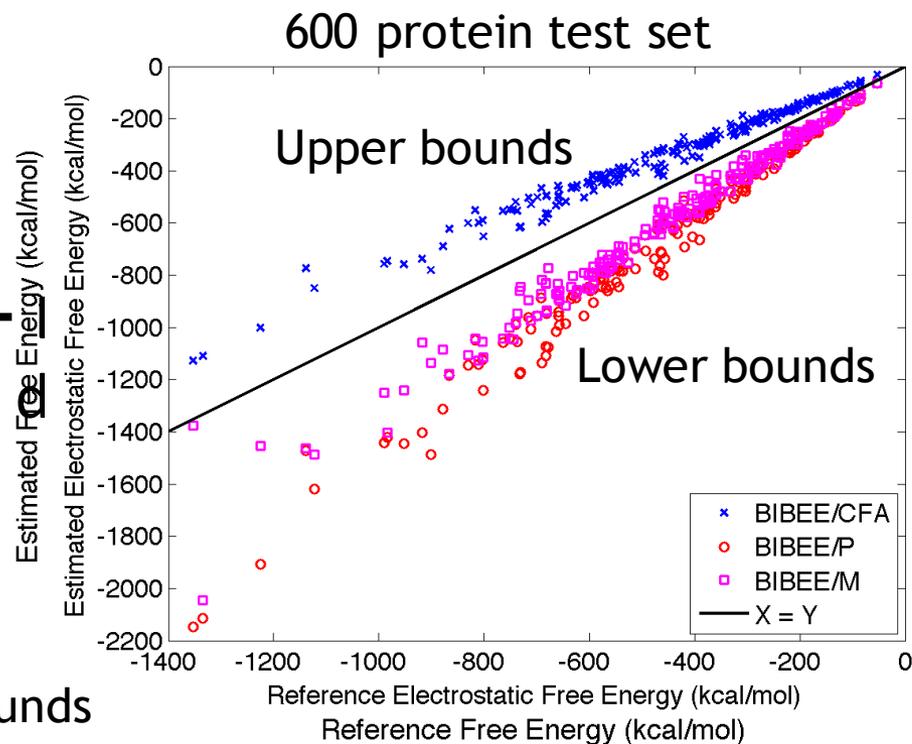
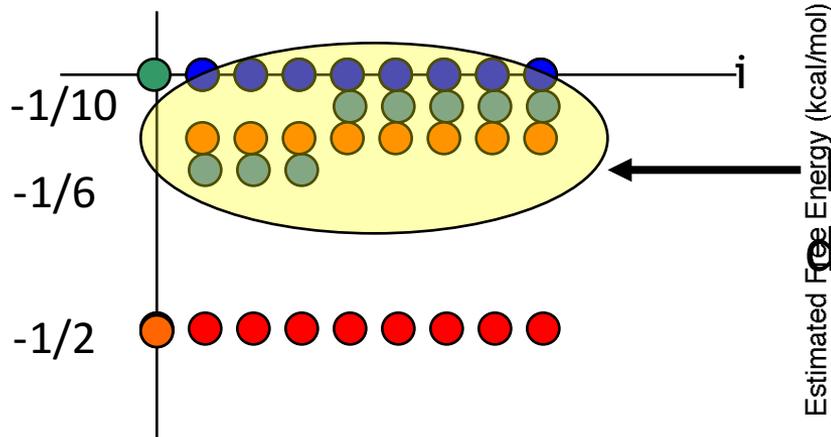
Speedup (approx)



Nystrom methods offer another order of magnitude (Knepley+Bardhan, 2015)

Multiscale approximation methods

Result: A flexible fast approximation scheme



We have proven

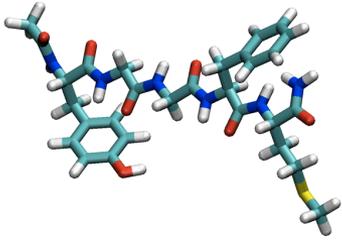
1. How to obtain upper and lower bounds (using the fact that the operator is quasi-Hermitian)
2. That the model is a deformation of the boundary condition
3. Eigenfunctions are exact in separable geometries

Mean absolute error: 4% !

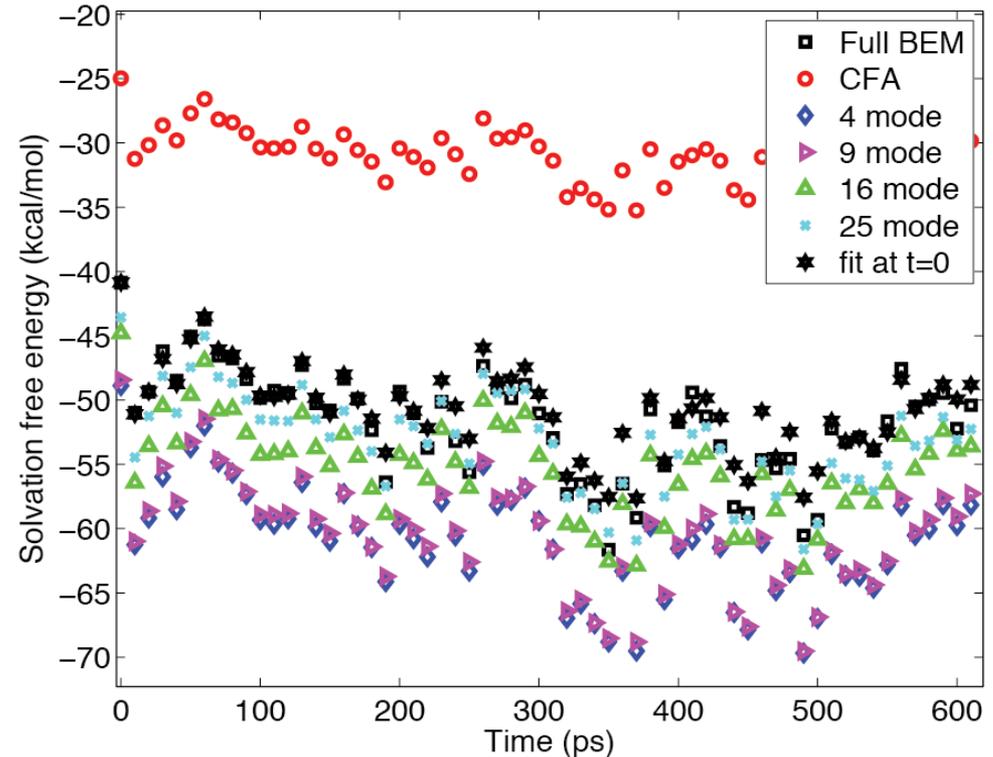
Multiscale approximation methods

Result: High accuracy under geometry variations

- Example: sampling protein conformations from MD



Example: neuropeptide met-enkephalin

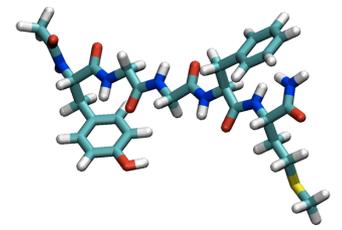


- Key feature:

- Advance scale of simulation while preserving our ability to add detail to the forward model.

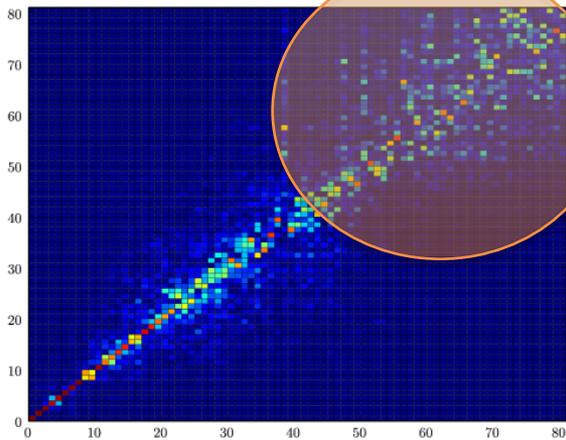
Multiscale approximation methods

Details: PDE-regularizer vs. ad hoc models

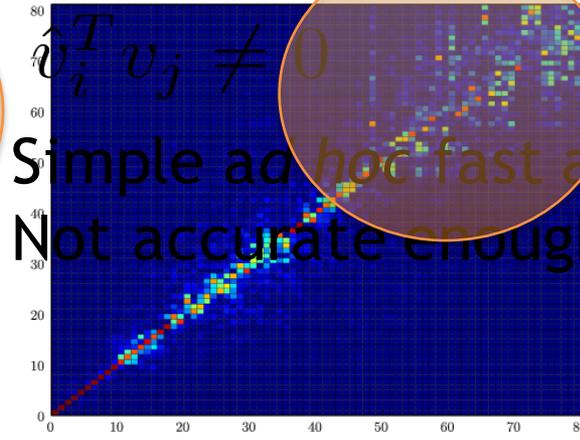


Comparing actual eigenvectors to approximate ones

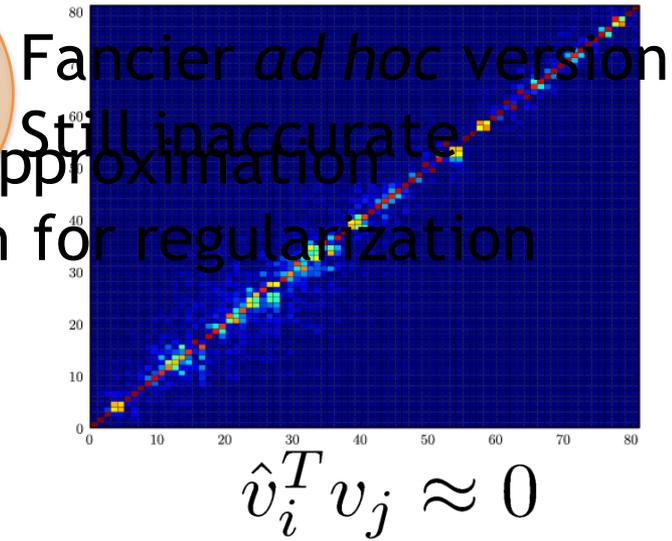
SGB/CFA



GBMV



My approach

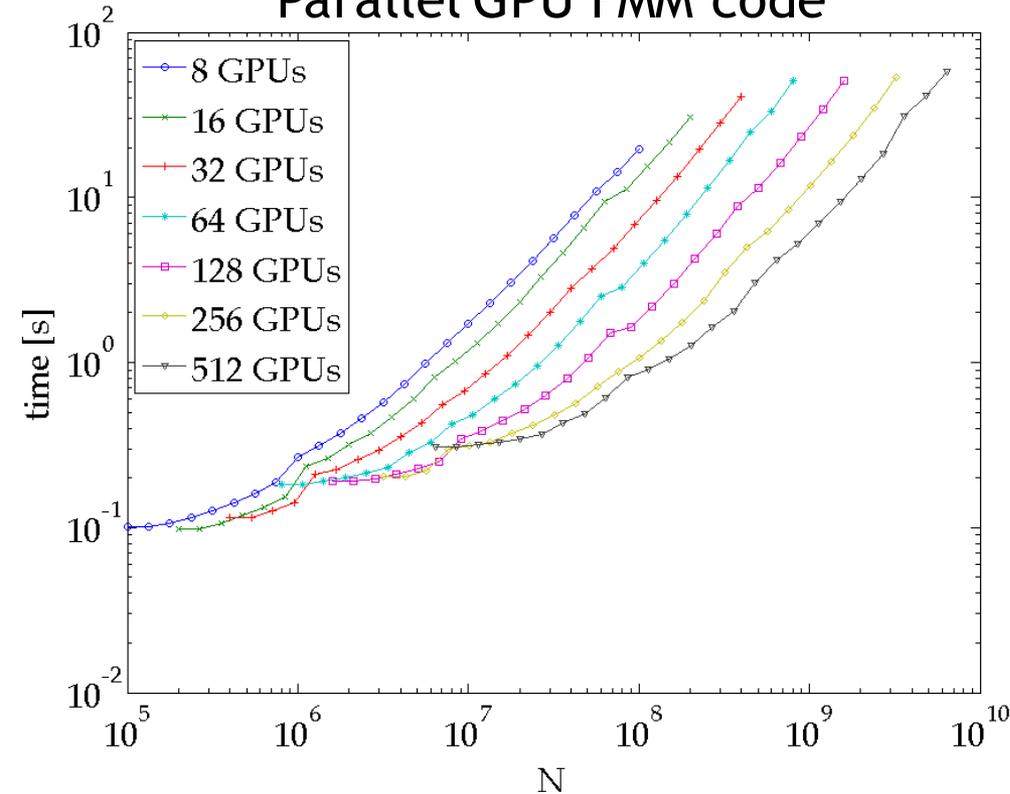


Bardhan, 2008; Bardhan et a. 2009, 2011

Massive parallelism for free

Uses existing scalable algorithmic primitives

Parallel GPU FMM code



760 node GPU cluster



Picture courtesy T. Hamada

❑ Other BIBEE implementations have used

- FFTSVD (Altman and Bardhan et al., 2006, 2009): OpenMP
- Tree codes (Cooper, Barba, et al., in prep.): GPU

PetFMM code of Yokota, Cruz, Barba, Knepley, Hamada

What can be done on a GPU-based workstation?



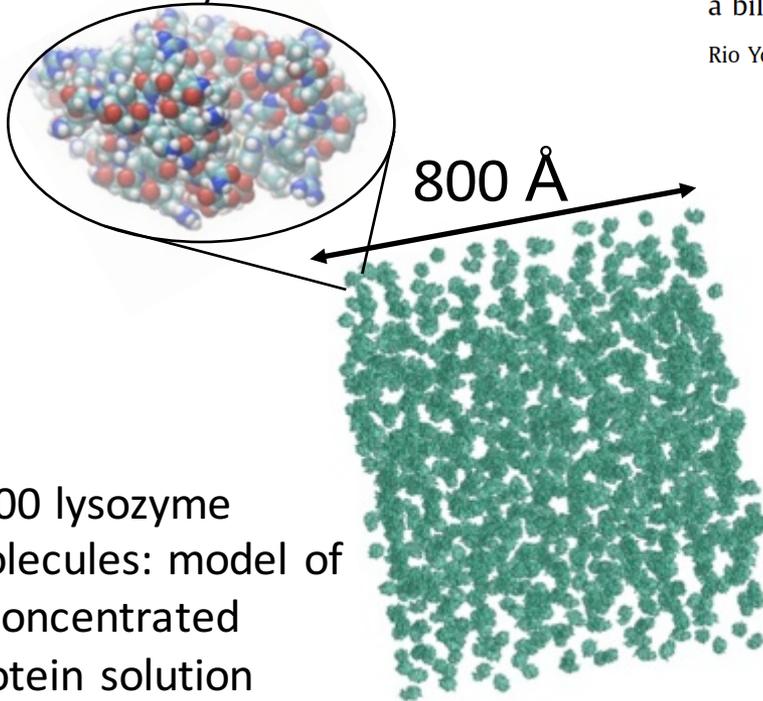
Contents lists available at ScienceDirect

Computer Physics Communications

www.elsevier.com/locate/cpc

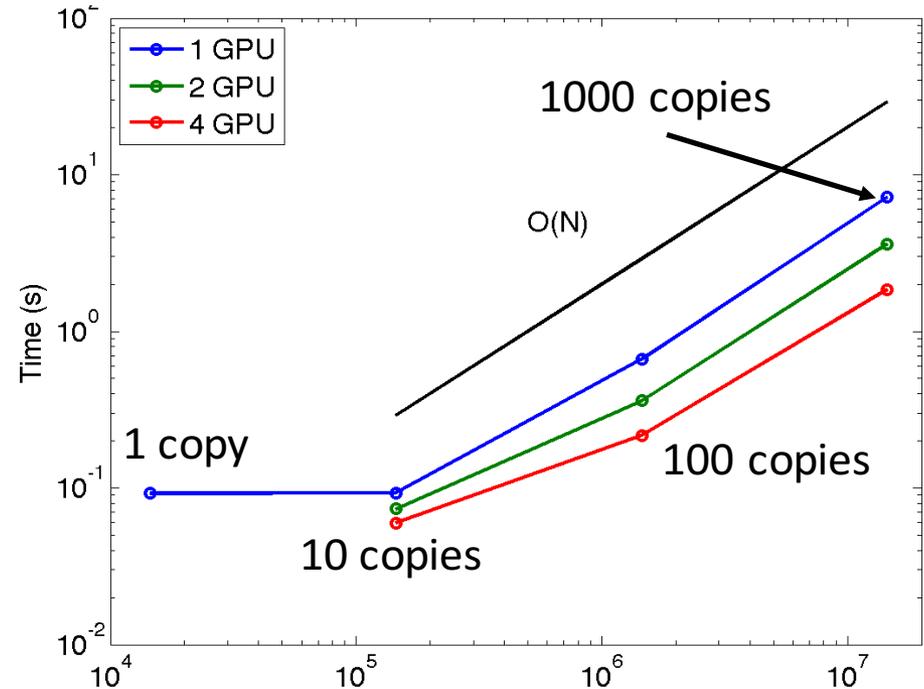


Lysozyme: ~2K atoms,
~15K boundary elements



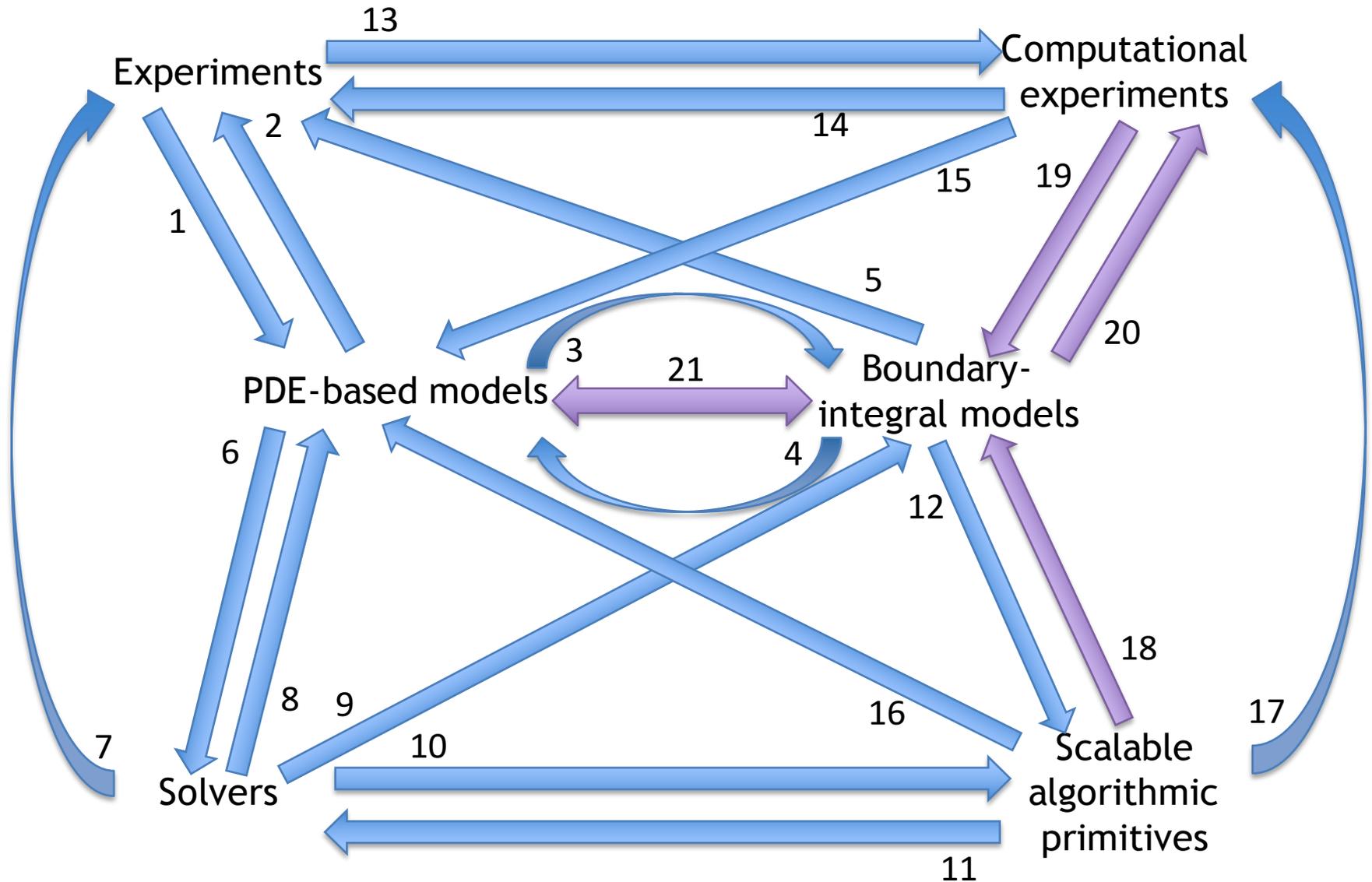
Biomolecular electrostatics using a fast multipole BEM on up to 512 GPUs and a billion unknowns

Rio Yokota^a, Jaydeep P. Bardhan^b, Matthew G. Knepley^c, L.A. Barba^{a,*}, Tsuyoshi Hamada^d



□ Applications in colloid and interface science, phase behavior in crowded solutions

Boundary integrals: Now more than ever



Key collaborators



Matt Knepley,
Rice University



Lee Makowski
Northeastern



Mihai Anitescu
Argonne and U. Chicago



Michael Altman
Merck



Pavel Jungwirth,
Czech Acad. Arts Sci.



Mala Radhakrishnan
Wellesley



Andreas Hildebrandt
U. Mainz

Acknowledgments

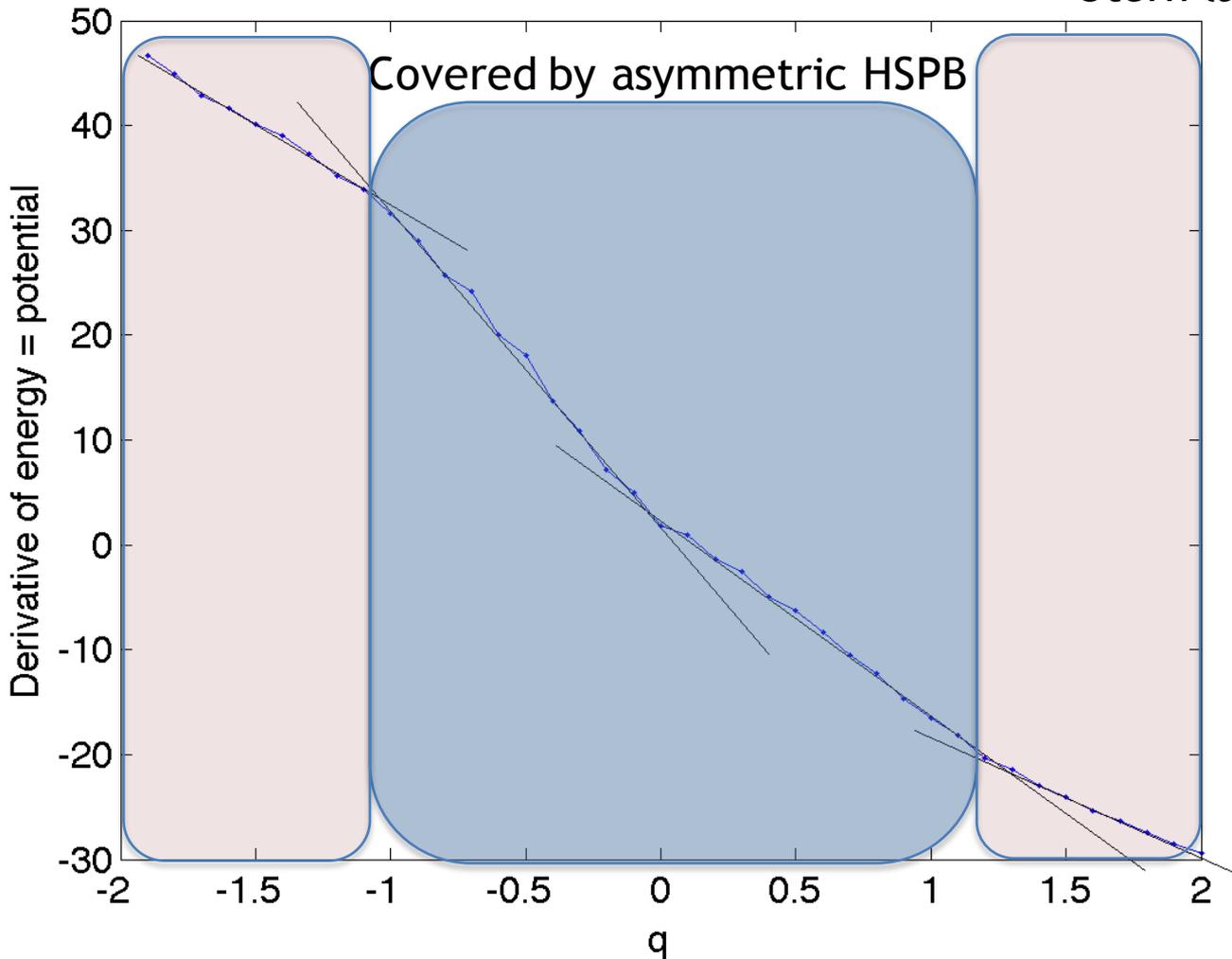
□ Thanks to funding support from

- NIH R21 (GM102642-01) (PI: J. Bardhan)
- U.S. Dept. of Energy Office of Science, Advanced Scientific Computing Research (DE-AC02-06CH11357, M. Knepley)

A totally insane idea: Piecewise-linear response

□ Charging Na between -2 and +2

Additional NLBC at
Stern layer



Bardhan+Jungwirth, unpublished