

1 **A method for redesigning molecular mechanics force field**
2 **parameterization by use of a Bayesian statistical framework**

3 Bryce C. Manubay^{1,*}

4 ¹*University of Colorado - Department of Chemical and Biological Engineering*

5 (Dated: October 6, 2016)

 This document describes a collected set of best practices for computing various physical properties from molecular simulations of liquid mixtures.

Keywords: best practices; molecular dynamics simulation; physical property computation

* bryce.manubay@colorado.edu

Todo list

I. Outline

A. Objectives (0.5pages)

- Molecular dynamics (MD) simulation is fast becoming a more useful tool in many scientific studies.
- However, some limitations remain in the ability of MD force fields to accurately and transferably describe molecular environments.
- Currently, force fields are parameterized with fixed functional forms with, often, poor physical motivation and require the chemical intuition of experts to manually correct parameters, leading to a more suitable product. Additionally, the creation of a transferable method to update existing force fields based on new experimental data is limited due to lack of understanding and lack of consistency in how the original parameterization was done
- A possible solution to these problems is by recasting the force field parameterization process as a bayesian inference problem.
- The objective of this paper is introduce a framework for using high quality experimental data in order to automatically generate families of MD force fields consistent with the data used.
- In this paper I will describe the overall parameterization framework and my roles in the project, first, collecting and organizing large amounts of high quality thermochemical data and, currently, investigating use of the Multistate Bennett Acceptance Ratio (MBAR) as a means to improve throughput by reducing simulation requirements during the parameterization process.

B. Significance (0.5pages)

- A broad variety of research has been greatly impacted by the advent and improvement of MD simulation tools.
 - Observing physical phenomena at a molecular scale (phase changes, ligand docking, etc.)¹
 - Drug discovery and deisgn of new molecules²
- The fundamental part of molecular simulation for describing the energetic interactions of a system is referred to as a force field, hence transferable and quantitatively accurate force fields are imperative for the use of molecular simulation tools to continue to proliferate.

- Transferability of MD force fields and particularly sets of parameters is an extremely popular topic (and current limitation) in the molecular simulation field.³⁻⁶ Transferability encourages use by providing convenience for scientists with wide arrays of research interests and simplifying the mystery that most observe force fields with.
- Inaccurate and poorly parameterized force fields have been shown to grossly misrepresent molecular systems^{7,8}
- A few notable attempts, such as GAAMP and ForceBalance, have been made in recent years towards more automated and systematic force field parameterization methods.⁹⁻¹² Each made important contributions to automated force field parameterization through clever use of objective function optimization, exploiting a variety of fitting data and allowing exploration of functional forms. However none provided the ability for the computer to automatically and systematically explore choices of fitting data, optimization algorithm and functional forms in order to objectively find families of force fields consistent with fitting data and reward those with the least model complexity.

C. Background and related literature (1.5pages ± 0.5pages)

D. Methods (1.5pages ± 0.5pages)

E. Progress (1.5pages ± 0.5pages)

F. Research plan (0.5pages)

-
- [1] G. Jayachandran, V. Vishal, and V. S. Pande, "Using massively parallel simulation and Markovian models to study protein folding: Examining the dynamics of the villin headpiece," *J Chem Phys*, vol. 124, no. 16, p. 164902, Apr. 2006. [Online]. Available: <http://scitation.aip.org/content/aip/journal/jcp/124/16/10.1063/1.2186317>
 - [2] M. De Vivo, M. Masetti, G. Bottegoni, and A. Cavalli, "Role of Molecular Dynamics and Related Methods in Drug Discovery," *J. Med. Chem.*, vol. 59, no. 9, pp. 4035–4061, May 2016. [Online]. Available: <http://dx.doi.org/10.1021/acs.jmedchem.5b01684>
 - [3] N. A. Vellore, J. A. Yancey, G. Collier, R. A. Latour, and S. J. Stuart, "Assessment of the Transferability of a Protein Force Field for the Simulation of Peptide-Surface Interactions," *Langmuir*, vol. 26, no. 10, pp. 7396–7404, May 2010. [Online]. Available: <http://dx.doi.org/10.1021/la904415d>
 - [4] D. A. Puleo and R. Bizios, *Biological Interactions on Materials Surfaces: Understanding and Controlling Protein, Cell, and Tissue Responses*. Springer Science & Business Media, Jun. 2009.
 - [5] F. Sato, S. Hojo, and H. Sun, "On the Transferability of Force Field Parameters With an ab Initio Force Field Developed for Sulfonamides," *J. Phys. Chem. A*, vol. 107, no. 2, pp. 248–257, Jan. 2003. [Online]. Available: <http://dx.doi.org/10.1021/jp026612i>
 - [6] A. Martin-Calvo, J. J. Gutiérrez-Sevillano, J. B. Parra, C. O. Ania, and S. Calero, "Transferable force fields for adsorption of small gases in zeolites," *Phys Chem Chem Phys*, vol. 17, no. 37, pp. 24 048–24 055, Oct. 2015.

- 74 [7] O. F. Lange, D. van der Spoel, and B. L. de Groot, "Scrutinizing Molecular Mechanics Force Fields on
75 the Submicrosecond Timescale with NMR Data," *Biophys J*, vol. 99, no. 2, pp. 647–655, Jul. 2010.
76 [Online]. Available: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2905107/>
- 77 [8] F. Martín-García, E. Papaleo, P. Gomez-Puertas, W. Boomsma, and K. Lindorff-Larsen, "Comparing
78 Molecular Dynamics Force Fields in the Essential Subspace," *PLoS One*, vol. 10, no. 3, Mar. 2015.
79 [Online]. Available: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4374674/>
- 80 [9] L. Huang and B. Roux, "Automated Force Field Parameterization for Nonpolarizable and Polarizable
81 Atomic Models Based on Ab Initio Target Data," *J. Chem. Theory Comput.*, vol. 9, no. 8, pp.
82 3543–3556, Aug. 2013. [Online]. Available: <http://dx.doi.org/10.1021/ct4003477>
- 83 [10] L.-P. Wang, T. J. Martinez, and V. S. Pande, "Building Force Fields: An Automatic, Systematic, and
84 Reproducible Approach," *J. Phys. Chem. Lett.*, vol. 5, no. 11, pp. 1885–1891, Jun. 2014. [Online].
85 Available: <http://dx.doi.org/10.1021/jz500737m>
- 86 [11] L.-P. Wang, T. Head-Gordon, J. W. Ponder, P. Ren, J. D. Chodera, P. K. Eastman, T. J. Martinez, and
87 V. S. Pande, "Systematic Improvement of a Classical Molecular Model of Water," *J. Phys. Chem. B*,
88 vol. 117, no. 34, pp. 9956–9972, Aug. 2013. [Online]. Available: <http://dx.doi.org/10.1021/jp403802c>
- 89 [12] L.-P. Wang, J. Chen, and T. Van Voorhis, "Systematic Parametrization of Polarizable Force Fields
90 from Quantum Chemistry Data," *J. Chem. Theory Comput.*, vol. 9, no. 1, pp. 452–460, Jan. 2013.
91 [Online]. Available: <http://dx.doi.org/10.1021/ct300826t>