Gaussian process models for force fields and wave functions

Aldo Glielmo

Department of Physics

King's College London

This dissertation is submitted for the degree of Doctor of Philosophy

January 2019

To my parents, who provided me with the tools that make me a human being,
and to the memory of Sandro, who showed me his way of using them.

Declaration

This dissertation describes work I have carried out between October 2015 and December 2018 at the department of physics of King's College London, under the supervision of Professor Alessandro De Vita (first supervisor from October 2015 to October 2018), Doctor George Booth (first supervisor from October 2015) and Professor Peter Sollich (second supervisor).

This dissertation contains material appearing in the following articles:

- A. Glielmo, C. Zeni, Á. Fekete and A. De Vita. Building nonparametric n-body force fields using Gaussian process regression. Submitted to K. T. Schütt, S. Chmiela, A. von Lilienfeld, A. Tkatchenko, K. Tsuda, and K. R. Müller, editors, *Machine learning for quantum simulations of molecules and materials* (Springer),
- A. Glielmo, C. Zeni and A. De Vita. Efficient nonparametric n-body force fields from machine learning. *Physical Review B*, 97, 2018,
- A. Glielmo, P. Sollich and A. De Vita. Accurate interatomic force fields via machine learning with covariant kernels. *Physical Review B*, 95, 2017.

In addition to the above, I have contributed to the following publications during the course of my PhD:

- M. Cucuringu, P. Davies, A. Glielmo, H. Tyagi. SPONGE: A generalized eigenproblem for clustering signed networks. Submitted to *International Conference on Artificial Intelligence and Statistics 22*,
- F. Bianchini, A. Glielmo, J. R. Kermode, A. De Vita. Enabling QM-accurate simulation of dislocation motion in gamma-Ni and alpha-Fe using a hybrid multiscale approach. Submitted to *Physical Review Materials*,

- N. Gunkelmann, D. Serero, A. Glielmo, M. Montaine, M. Heckel, T. Pöschel. Stochastic nature of particle collisions and its impact on granular material properties. Submitted to S. Antonyuk, editor, *Particles in contact* (Springer),

- C. Zeni, K. Rossi, A. Glielmo, Á. Fekete, N. Gaston, F. Baletto and A. De Vita. Building machine learning force fields for nanoclusters. *The Journal of Chemical Physics*, 148, 2018,

- M. Heckel, A. Glielmo, N. Gunkelmann, T. Pöschel. Can we obtain the coefficient of restitution from the sound of a bouncing ball?. *Physical Review E*, 93, 2016.

This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements. It has not been submitted in whole or in part for any degree or diploma at this or any other university.

Aldo Glielmo
January 2019

Acknowledgements

...

Summary

Gaussian process models for force fields and wave functions

Aldo Glielmo
King's College London

Algorithms capable of extracting information from data are increasingly finding application in condensed matter physics. ...

Contents

List of Figures

List of symbols and acronyms

Acronyms

DFT	Density functional theory
DFTB	Density functional tight binding
EAM	Embedded atom model
GP	Gaussian process
LJ	Lennard Jones
LOFT	Learn on the fly
MAE	Mean absolute error
MD	Molecular dynamics
MFF	Mapped force field
ML	Machine learning
ML-FF	Machine learning force field
QM	Quantum mechanical
QM/MM	Quantum mechanics/molecular mechanics
VMC	Variational Monte Carlo

Math and Gaussian process conventions

$a \mid \mathbf{a} \mid \mathbf{A} \mid \mathbb{A}$	Scalar, vector, matrix and block matrix
$\mathbf{a}^{\mathrm{T}} \mid \hat{\mathbf{a}}$	The transpose $\mathbf{a}$ and the unit vector in the direction of $\mathbf{a}$
$\sim$	Distributed according to or approximately equal to
$\mathcal{GP}$	Gaussian stochastic process
$\langle\cdot\rangle$	Expectation with respect to a Gaussian stochastic process
$\boldsymbol{\theta}$	Vector of hyperparameters
$\ell \mid \sigma_n^2$	Lengthscale and noise hyperparameter
$\mathcal{D}$	Dataset
N	Number of training points
$\mathcal{O}(\cdot)$	Computational complexity

Force field learning

ρ	Local atomic configuration
M	Number of atoms within the cutoff radius of an atomic configuration ρ
$\varepsilon^r \mid \varepsilon(\rho)$	Reference and latent local energy function
$\hat{\varepsilon}(\rho) \mid \hat{\sigma}^2(\rho)$	Mean and variance of the local energy predictive distribution
$k(\rho, \rho')$	Local energy kernel evaluated for configurations ρ and ρ'
$k_n \mid k_n^s \mid k_n^{\neg u}$	Local energy n-body kernel, symmetric n-body kernel and symmetric n-body kernel that is non unique
$k_{MB} \mid k_{MB}^s \mid k_{MB}^{ds}$	Local energy many-body kernels, many-body kernel symmetric over the full orthogonal group and many-body kernel symmetric over a discrete group
$\mathbf{f}^r \mid \mathbf{f}(\rho)$	Reference and latent force on the central atom of a configuration
$\hat{\mathbf{f}}(\rho) \mid \hat{\mathbf{\Sigma}}(\rho)$	Mean and covariance of the force predictive distribution
$\mathbf{K}(\rho, \rho')$	Matrix-valued force kernel evaluated for two configurations
$\mathbf{K}_n^G \mid \mathbf{K}_{MB}^G$	Matrix-valued n-body and many-body force kernels that are covariant over the group G
$\mathcal{P} \mid \mathbf{P}$	Abstract operator and matrix representation of a permutation of atoms of the same chemical species
$\mathcal{R} \mid \mathbf{R}$	Abstract operator and matrix representation of a rotation
$\mathcal{F} \mid \mathbf{F}$	Abstract operator and matrix representation of a reflection
$\mathcal{Q} \mid \mathbf{Q}$	Abstract operator and matrix representation of an element of the orthogonal group (either a rotation or a reflection)
$C_n \mid D_n$	Cyclic and dihedral group of order n

Wave function learning

$\mathbf{x}$	An electronic configuration
$\psi^r \mid \psi(\mathbf{x})$	The reference and the latent wave function
$\psi_S(\mathbf{x})$	Ground state wave function of a quadratic Hamiltonian
$\lambda(\mathbf{x})$	The natural logarithm of the absolute value of the wave function $\psi(\mathbf{x})$ or of the ratio $\psi(\mathbf{x})/\psi_S(\mathbf{x})$
$k(\mathbf{x}, \mathbf{x}')$	Kernel between two electronic configurations $\mathbf{x}$ and $\mathbf{x}'$
$k \mid k_n^d \mid k_c$	Plaquette kernel of size n, distance dependent kernel of size n and complete kernel

Chapter 1

Introduction

DATA, and the fast development of efficient machine learning algorithms capable of exploiting them, are changing the paradigm under which great computational challenges are tackled. Up until ten or fifteen years ago, algorithms were designed with the idea of performing a sequence of instructions, well determined by a programmer, in order to solve a given problem. In more recent times instead, this "paradigm" is shifting towards the design of algorithms that automatically *learn* from data the particular sequence of instructions best suited to solve the problem at hand.

This change can be perhaps best understood by means of an emblematic comparison: that between the software used within the IBM "Deep Blue" computer in 1997 to beat the chess world champion Garri Kasparov [1], and that used twenty years later, in 2016, by Google's "AlphaGo" to beat Lee Sedol, an international champion in the game of Go [2]. While the first algorithm was based on a human-programmed smart search of all possible moves aimed at finding the one yielding to the largest advantage [3–5], within AlphaGo the function yielding to the best move at any given point of the game had been previously learned by the algorithm by both analysing big amounts of previously played matches and by playing against itself [6]. [1]

A similar change is also taking place in the domain of computational condensed matter physics research, where the exploitation of data is recently taking a prominent role.

[1] Interestingly, a modified version of AlphaGo later repeatedly defeated a modified (and more powerful) version of the Deep Blue chess algorithm [7].

Appendices

A.1 On the derivation of the GP predictive distribution

This appendix gives a sketch of the procedure by which Eq. (**??**) is obtained, which substantially relies on the properties of multivariate Gaussian distributions. For full details on this one can consult the excellent Refs. [8] and [9].

Bibliography

[1] Harding, L. & Barden, L. Deep Blue win a giant step for computerkind. *The Guardian* (1997). URL `https://www.theguardian.com/theguardian/2011/may/12/deep-blue-beats-kasparov-1997`.

[2] Borowiec, S. Google's AlphaGo AI defeats human in first game of Go contest. *The Guardian* (2016). URL `https://www.theguardian.com/technology/2016/mar/09/google-deepmind-alphago-ai-defeats-human-lee-sedol-first-game-go-contest`.

[3] Russell, S. & Norvig, P. *Artificial Intelligence: A Modern Approach* (Pearson, 1995).

[4] Hsu, F. H. IBM's deep blue chess grandmaster chips. *IEEE Micro* **19**, 70–81 (1999).

[5] Hsu, F. H. *Behind Deep Blue: Building the computer that defeated the world chess champion* (Princeton University Press, 2004).

[6] Silver, D. *et al.* Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484–489 (2016).

[7] Silver, D. *et al.* A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* **362**, 1140–1144 (2018).

[8] Bishop, C. M. *Pattern Recognition and Machine Learning* (Springer, 2006).

[9] Rasmussen, C. E. & Williams, C. K. I. *Gaussian Processes for Machine Learning* (The MIT Press, 2006).