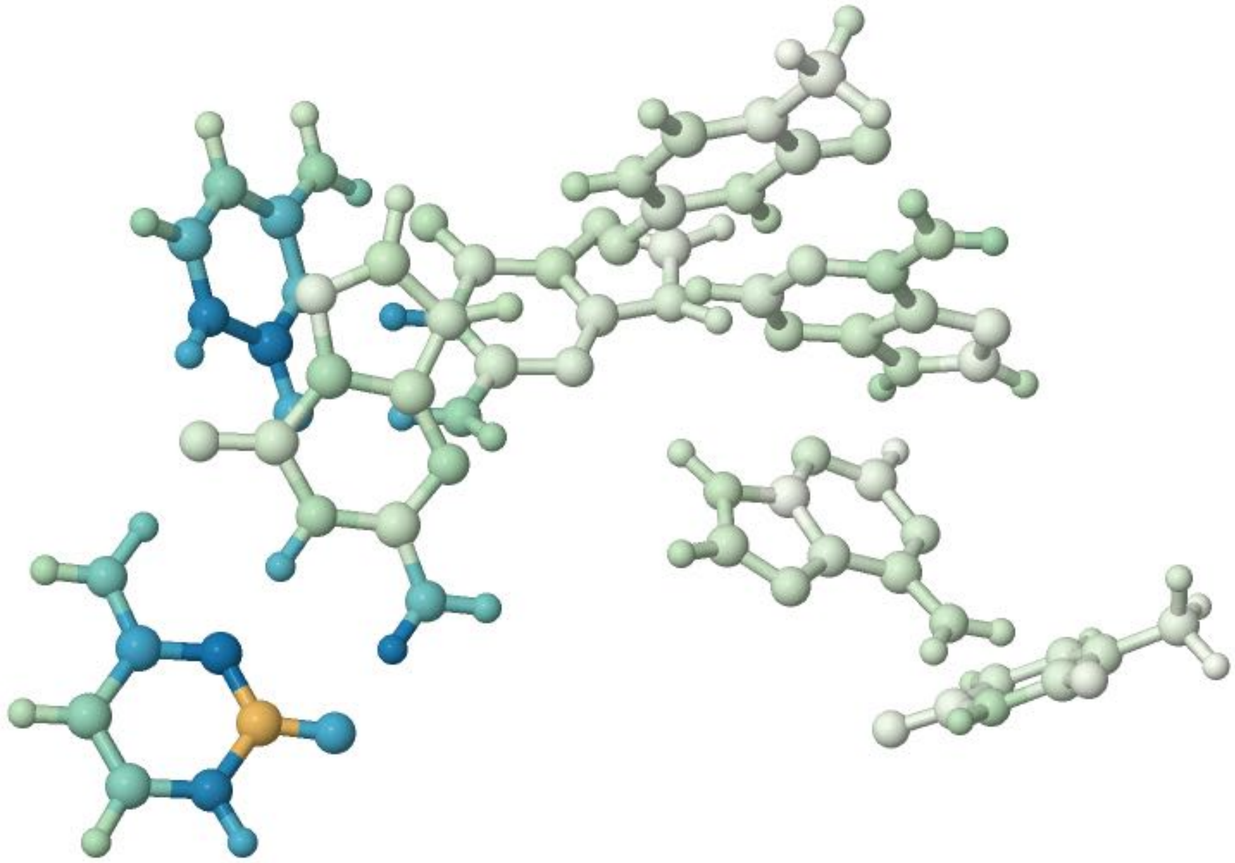
Accurate global machine learning force fields for molecules with hundreds of atoms

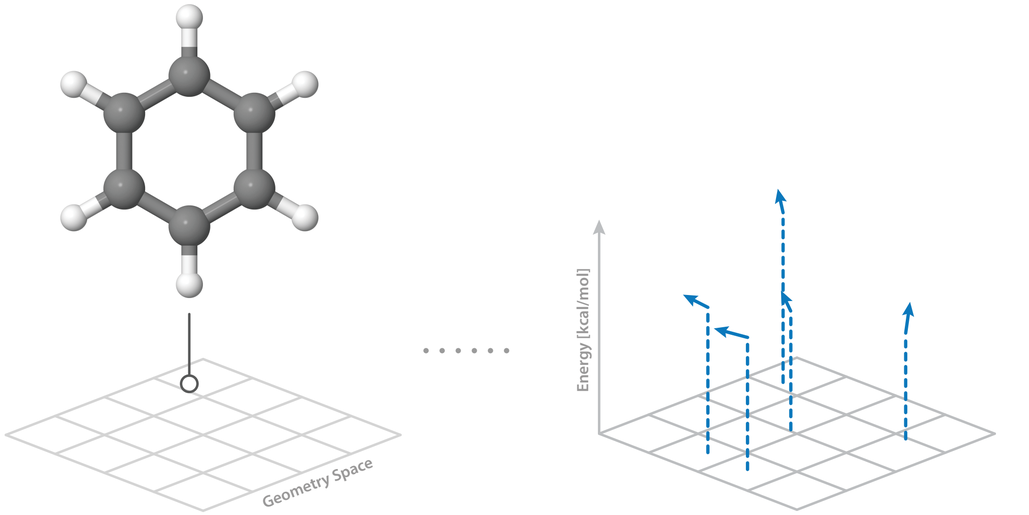
IPAM Workshop: EMS2023



Stefan Chmiela

Force field reconstruction with ML

A unique set of challenges



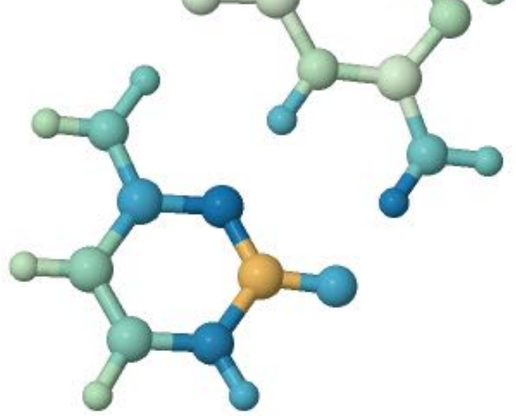
Geometries Energy / atomic forces

- Learning on sets
  (of indistinguishable atoms)
- Invariant / equivariant outputs
- Stringent speed / accuracy requirements

Classical force fields **ML** A*b initio* approximations to Schrödinger's equation

Atomistic modeling with ML

Historical perspective



Numerical analysis

- exploit model structure in training (e.g. correlations, PDE constraints)
- remove uncontrolled approximations

Feature engineering

- descriptor + linear models

Graph neural networks

- representation learning

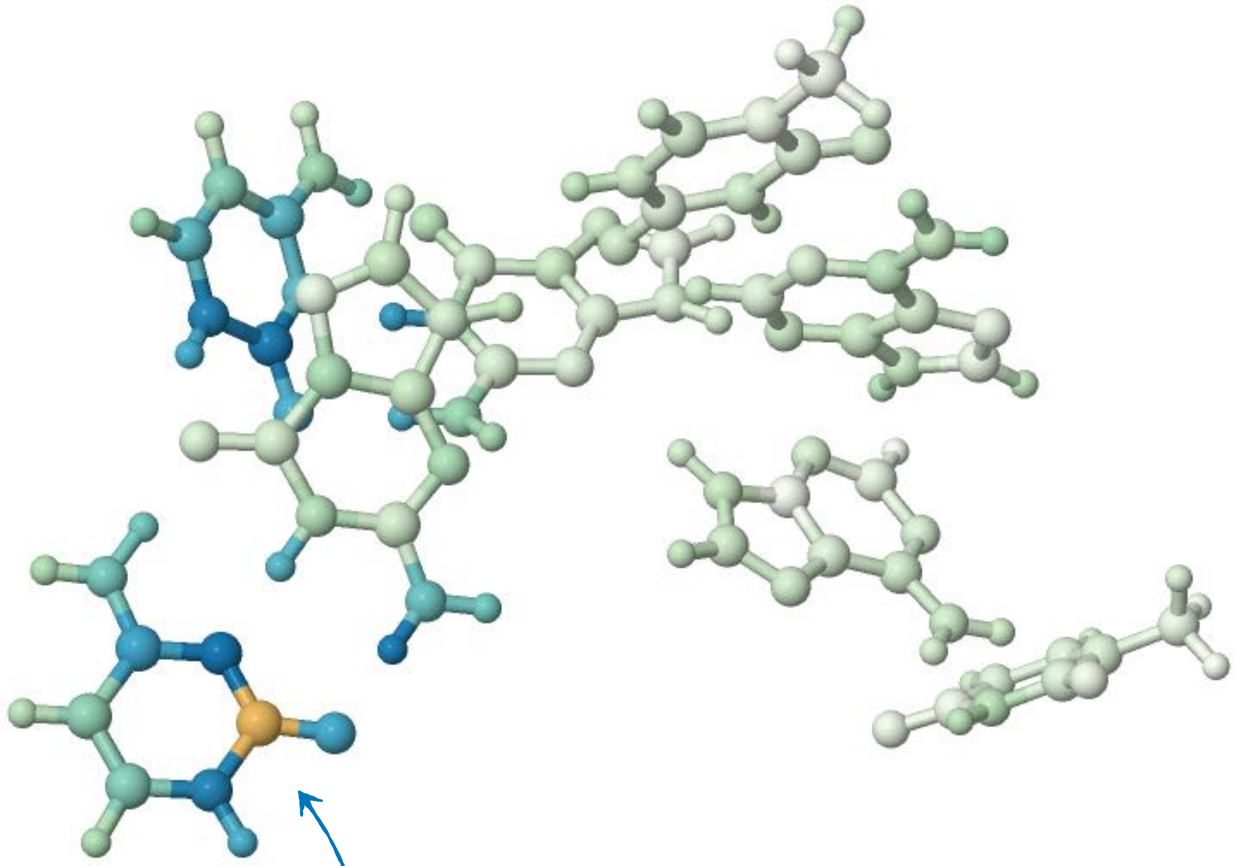
Architecture engineering

- sophisticated filter parameterizations (e.g. YLM)
- "geometric deep learning"

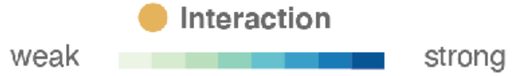
Behler 2011, Bartók et al. 2013, Hirn et al. 2013, Cohen-Welling 2016, Gilmer et al. 2017, Duvenaud et al. 2015, Schütt et al. 2017, Thomas-Smidt et al. 2018, Weiler et al. 2018, Anderson et al. 2019, Unke et al. 2021, Frank et al. 2021

Atomistic modeling with ML

Efficient modeling of global force fields

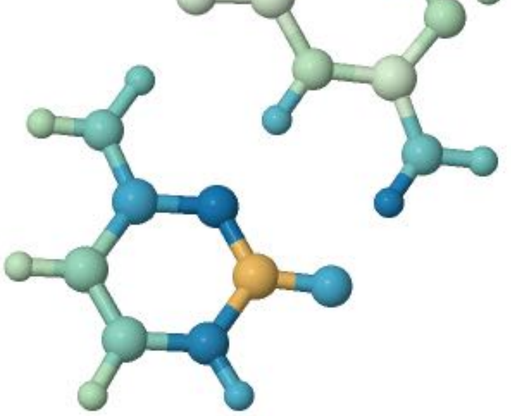


All atoms interact in
quantum many-body systems.



Global atomic interactions

An emerging field without universal solution



TensorMol

Hybrid nearsighted NN potential with screened long-range electrostatic and van der Waals physics

Yao, K., Herr, J. E., Toth, D. W., Mckintyre, R. & Parkhill, J. The TensorMol-0.1 model chemistry: a neural network augmented with long-range physics. Chem. Sci. 9, 2261 (2018).

IPML

Physics-based mechanistic descriptions combined with environment-dependent ML corrections

Bereau, T., DiStasio Jr, R. A., Tkatchenko, A. & Von Lilienfeld, O. A. Non-covalent interactions across organic and biological subsets of chemical space: physics-based potentials parametrized from machine learning. J. Chem. Phys. 148, 241706 (2018).

Long-distance equivariant (LODE) representation + SOAP

Atom-density potential folds global structural and compositional information information into a local representation

Grisafi, A. & Ceriotti, M. Incorporating long-range physics in atomic-scale machine learning. J. Chem. Phys. 151, 204105 (2019).

Fourth-generation Behler-Parinello neural network (4G-BPNN)

ML-FF with non-local charge transfer correction using independent ML model

Ko, T. W., Finkler, J. A., Goedecker, S., & Behler, J. (2021). A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer. Nat. Commun., 12(1), 1-11.

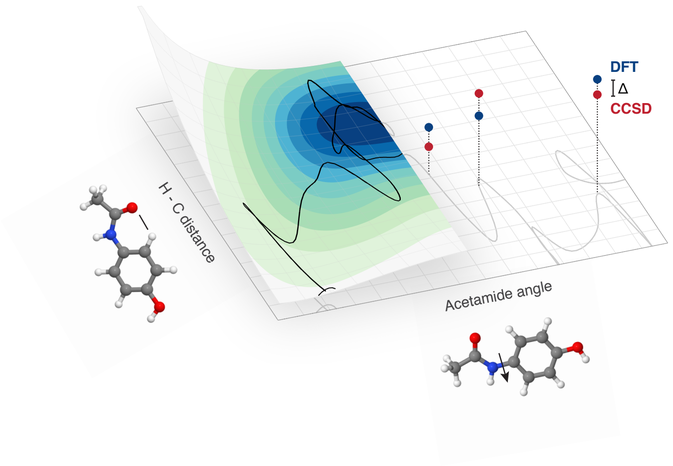
SpookyNet

ML force field augmented with physically motivated corrections for long-ranged electrostatic and dispersion interactions

Unke, O. T., Chmiela, S., Gastegger, M., Schütt, K. T., Sauceda, H. E., & Müller, K. R. (2021). SpookyNet: Learning force fields with electronic degrees of freedom and nonlocal effects. Nat. Commun., 12(1), 1-14.

Structure and properties in MLFFs

Physical symmetries, data correlations, modeling choices



Symmetries /
conservation laws

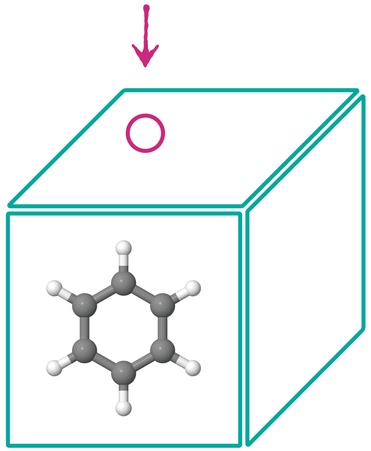
Physical

Frank et al. 2021, Schmitz et al. 2022, Chmiela et al. 2023, Blücher et al. 2023

Structure and properties in MLFFs

Physical symmetries, data correlations, modeling choices

Physically feasible subspace is much smaller.



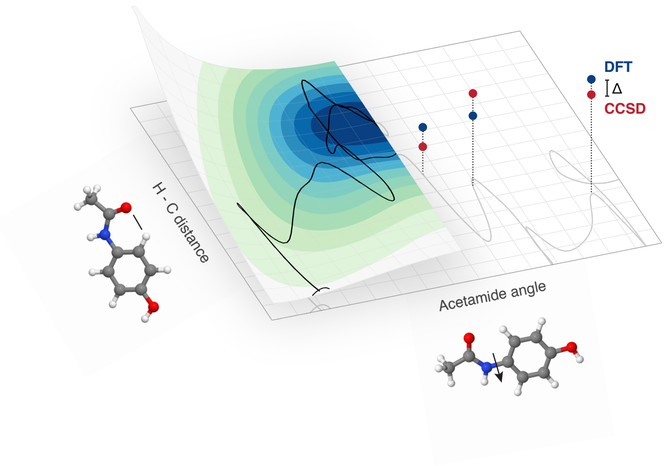
e.g. Noether’s theorem,
geometric priors etc.

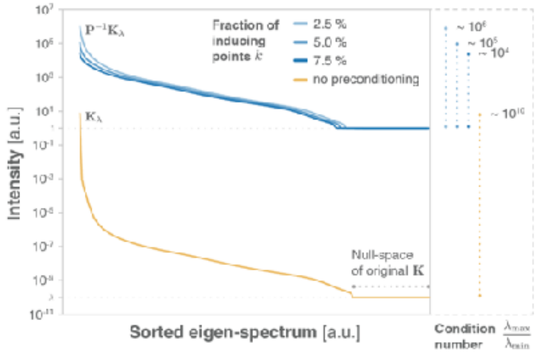
Physical

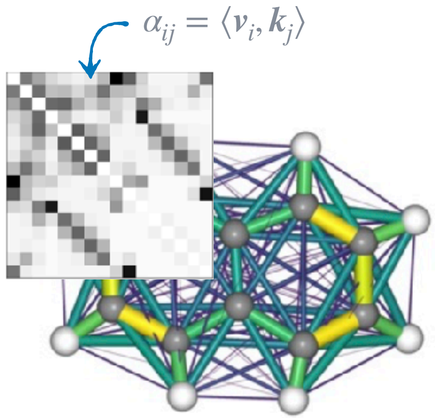
Frank et al. 2021, Schmitz et al. 2022, Chmiela et al. 2023, Blücher et al. 2023

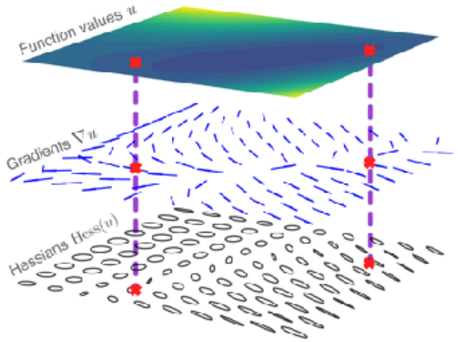
Structure and properties in MLFFs

Physical symmetries, data correlations, modeling choices









Symmetries / conservation laws

Correlation patterns in the reference data

Data driven interactions

Efficient training of constrained models

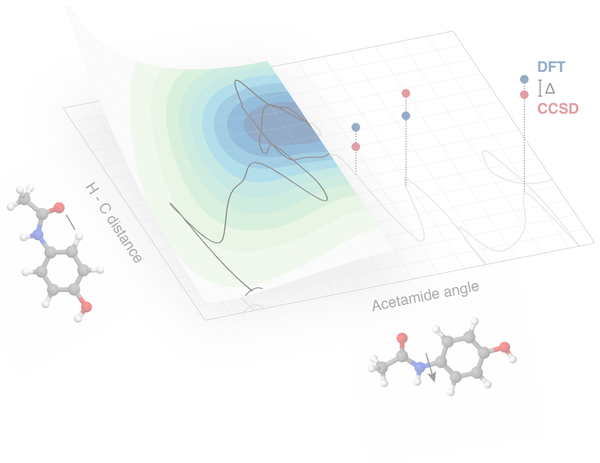
Physical ····· Statistical (ML) ····· Architectural (ML)

Frank et al. 2021, Schmitz et al. 2022, Chmiela et al. 2023, Blücher et al. 2023

Structure and properties in MLFFs

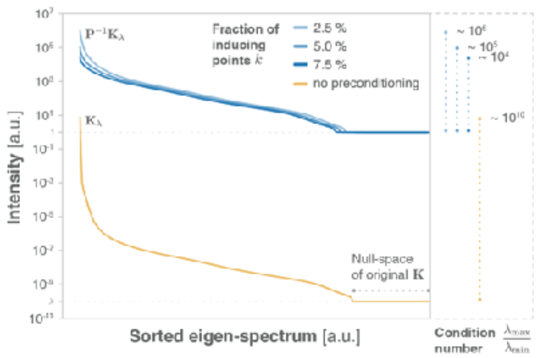
Physical symmetries, data correlations, modeling choices

A

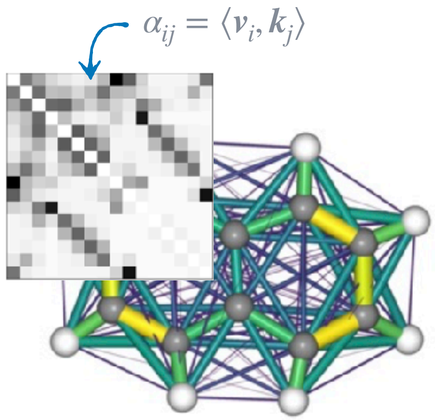


Symmetries / conservation laws

C

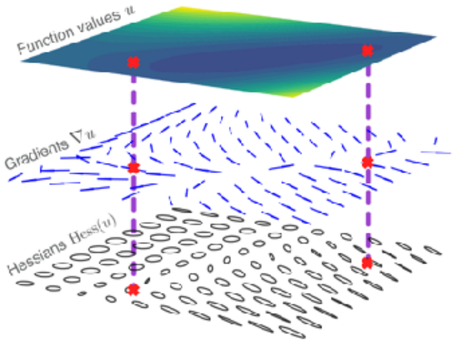


Correlation patterns in the reference data



Data driven interactions

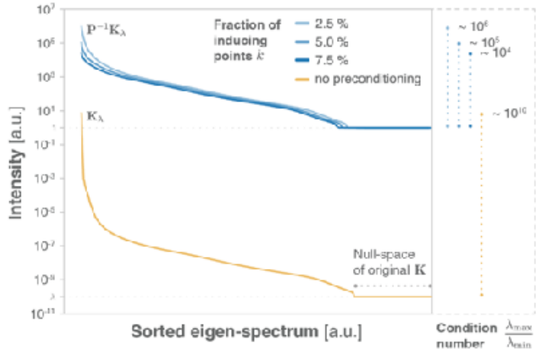
B



Efficient training of constrained models

Physical ····· Statistical (ML) ····· Architectural (ML)

Frank et al. 2021, Schmitz et al. 2022, Chmiela et al. 2023, Blücher et al. 2023

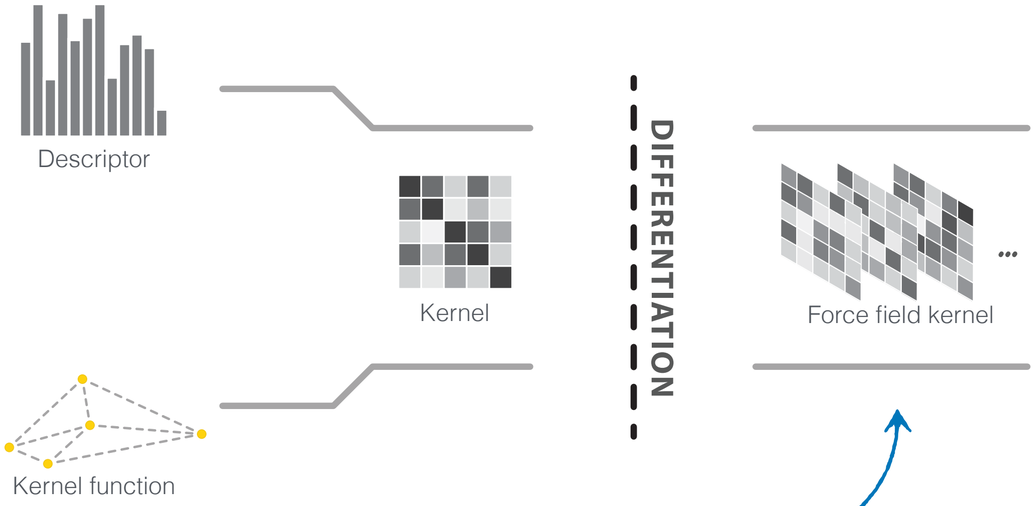


Correlation patterns in the reference data

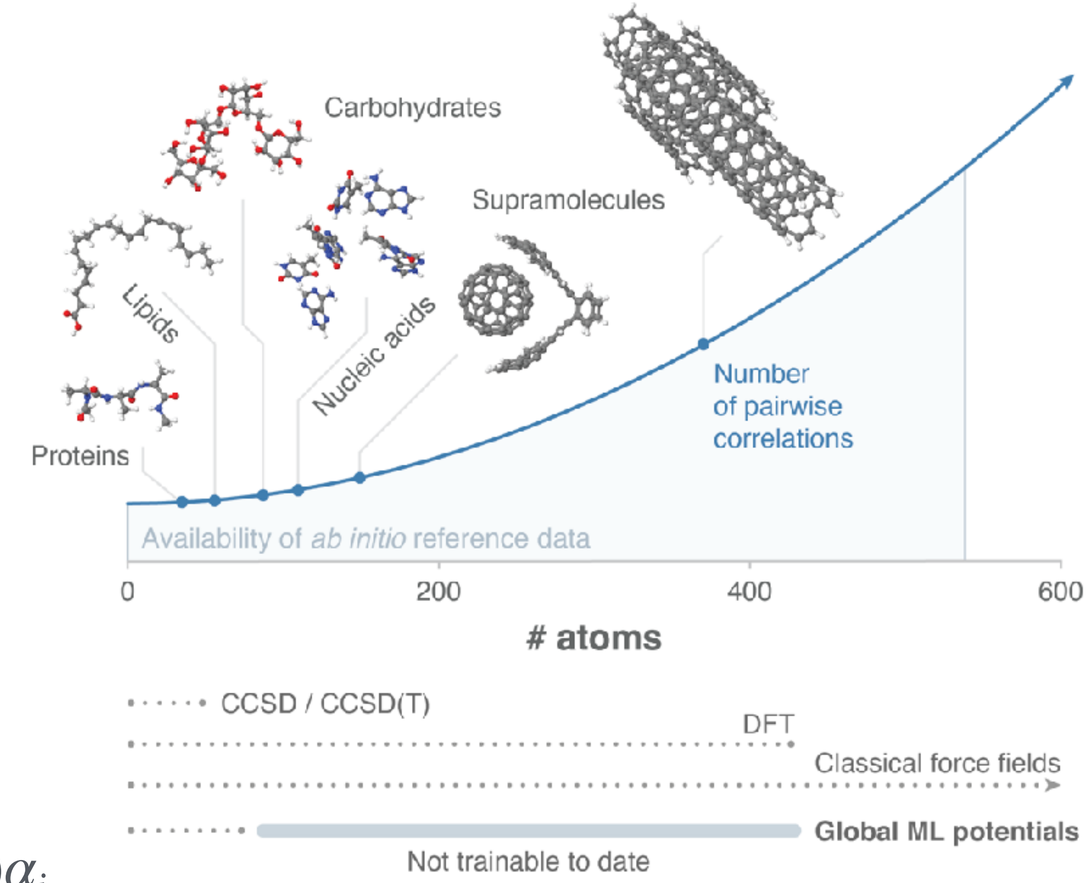
Force field reconstruction with ML

Representing global interactions

All atoms interact:
quadratic complexity (minimum)



Linear model
Pairwise atom-correlations: $\hat{f}(\mathbf{x}) = \mathbf{K}\alpha = \sum_{i}^{M} k(\mathbf{x}, \mathbf{x}_i)\alpha_i$



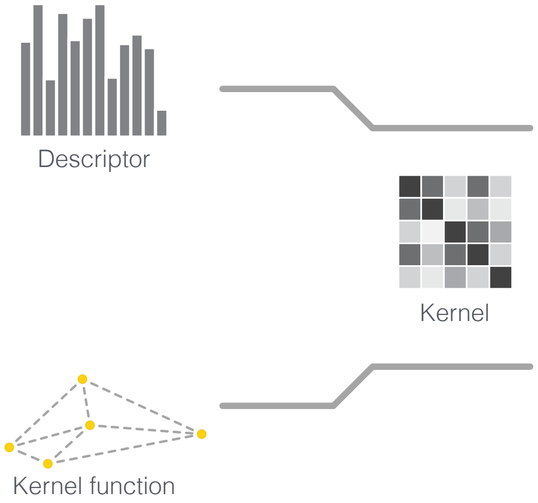
Reference data,
but no models

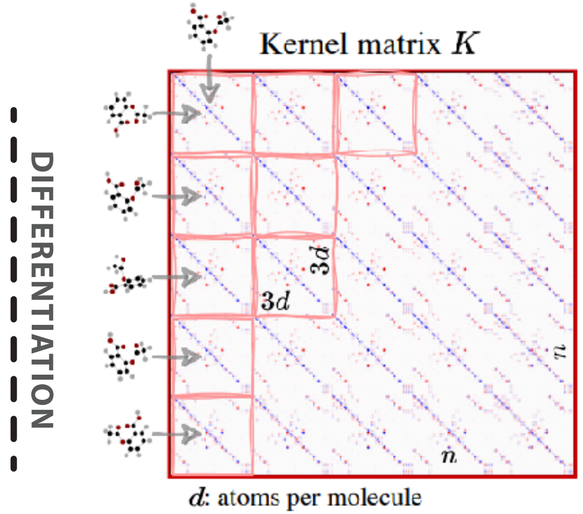
Chmiela et al. 2023, Blücher et al. 2023

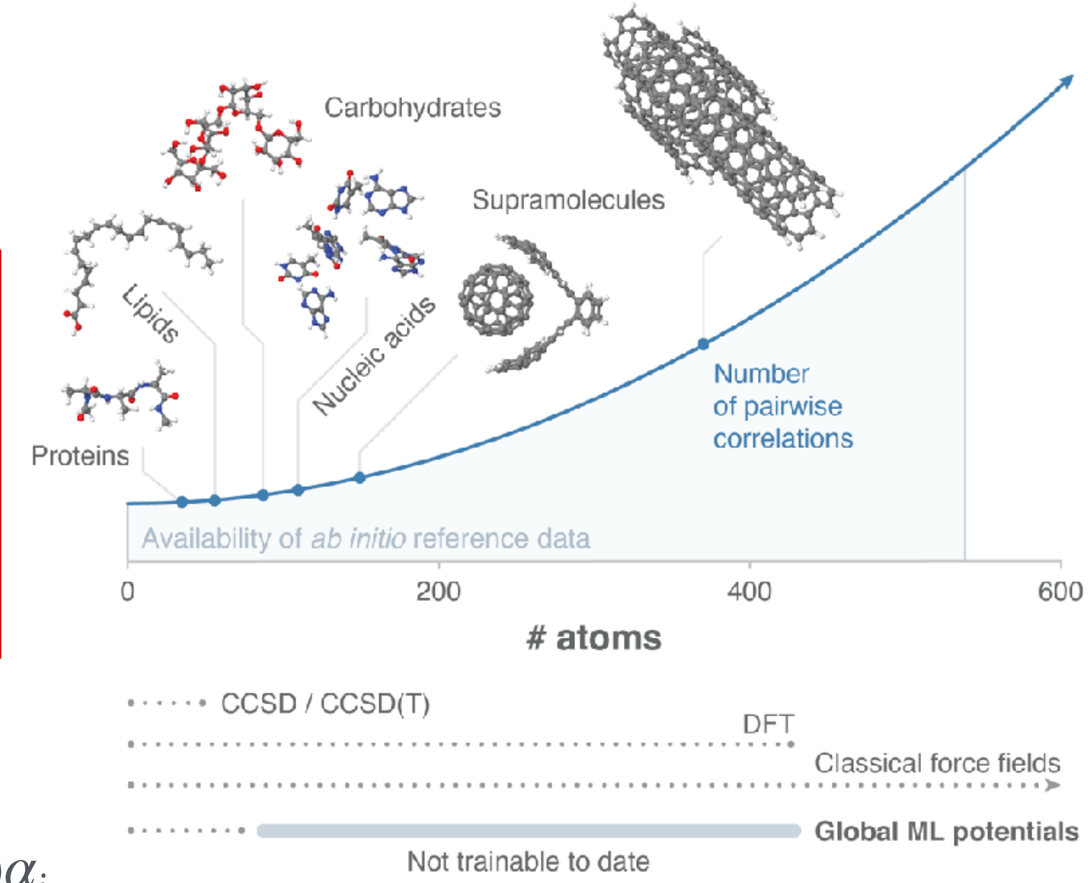
Force field reconstruction with ML

Representing global interactions

All atoms interact:
quadratic complexity (minimum)





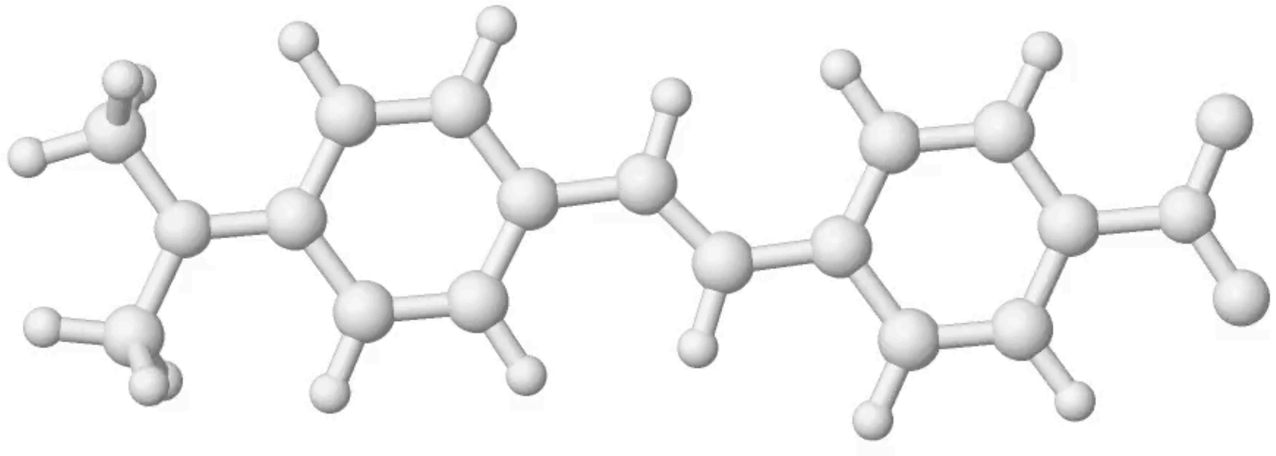


Linear model
Pairwise atom-correlations: $\hat{f}(\mathbf{x}) = \mathbf{K}\alpha = \sum_{i}^{M} k(\mathbf{x}, \mathbf{x}_i)\alpha_i$

Chmiela et al. 2023, Blücher et al. 2023

Force field reconstruction with ML

Atoms interact non-locally



“Donor-Bridge-Acceptor”

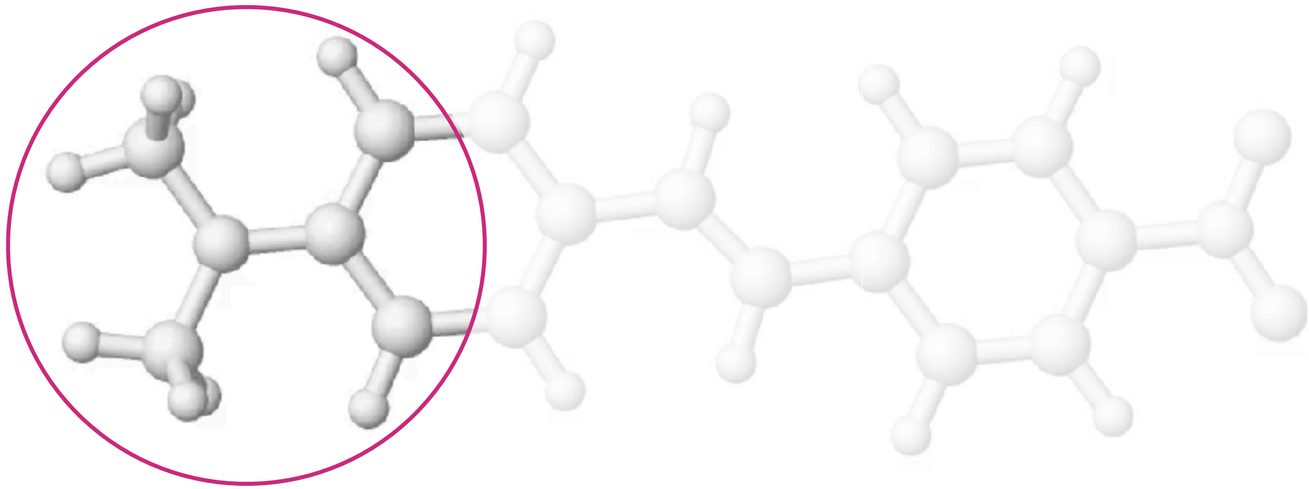
(E)-N,N-dimethyl-4-(4-nitrostyryl)aniline

All atoms need to interact to make accurate predictions.

Chmiela et al. 2023

Force field reconstruction with ML

Atoms interact non-locally



Limit interaction length

“Donor-Bridge-Acceptor”

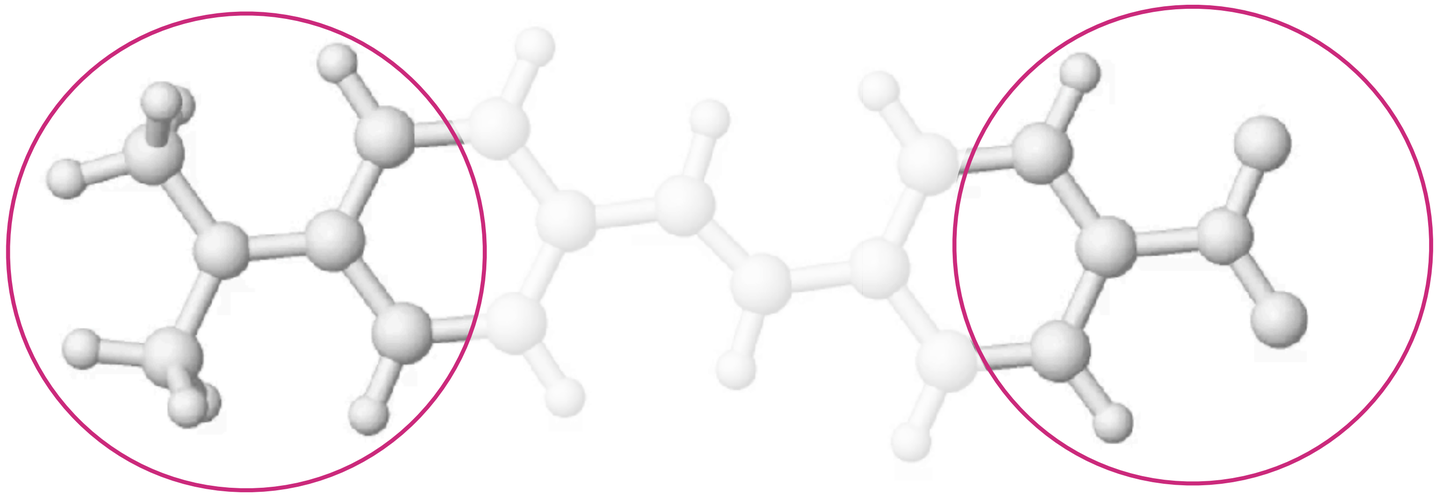
(E)-N,N-dimethyl-4-(4-nitrostyryl)aniline

All atoms need to interact to make accurate predictions.

Chmiela et al. 2023

Force field reconstruction with ML

Atoms interact non-locally



Limit interaction length

“Donor-Bridge-Acceptor”

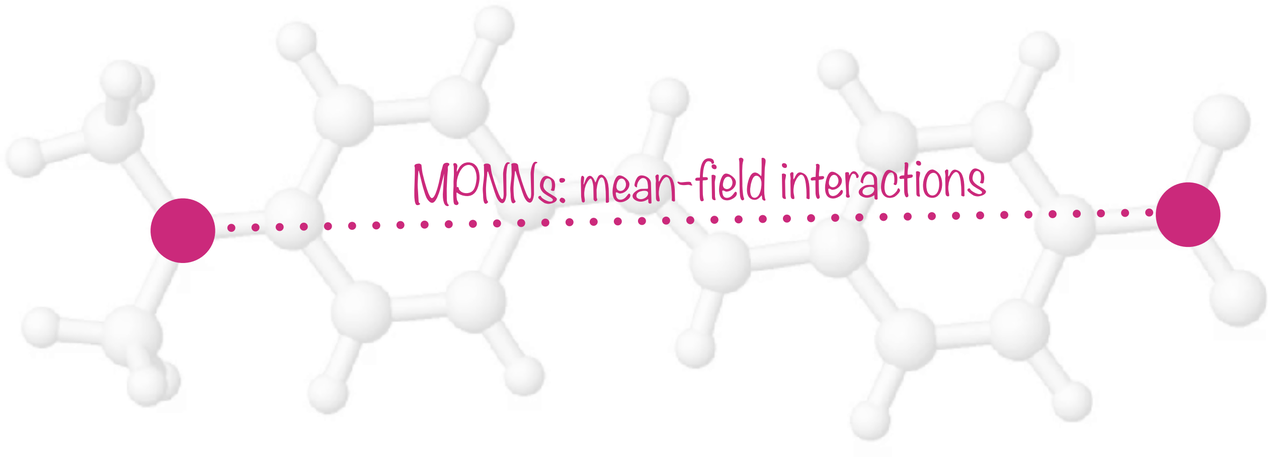
(E)-N,N-dimethyl-4-(4-nitrostyryl)aniline

All atoms need to interact to make accurate predictions.

Chmiela et al. 2023

Force field reconstruction with ML

Atoms interact non-locally



Localization can remove important interaction-scales!

“Donor-Bridge-Acceptor”

(E)-N,N-dimethyl-4-(4-nitrostyryl)aniline

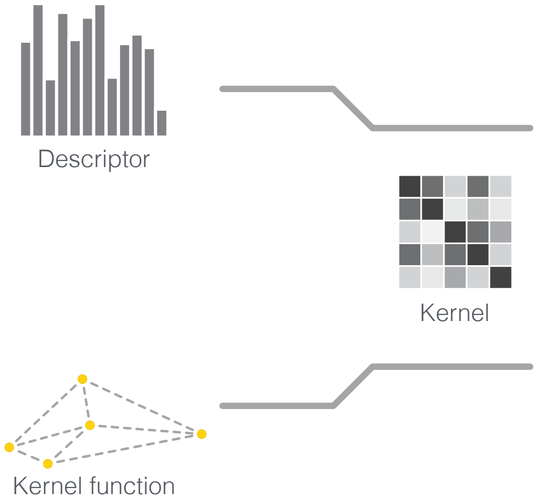
All atoms need to interact to make accurate predictions.

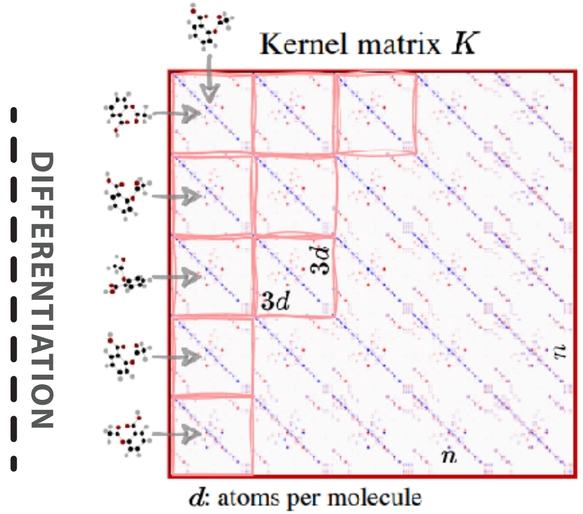
Chmiela et al. 2023

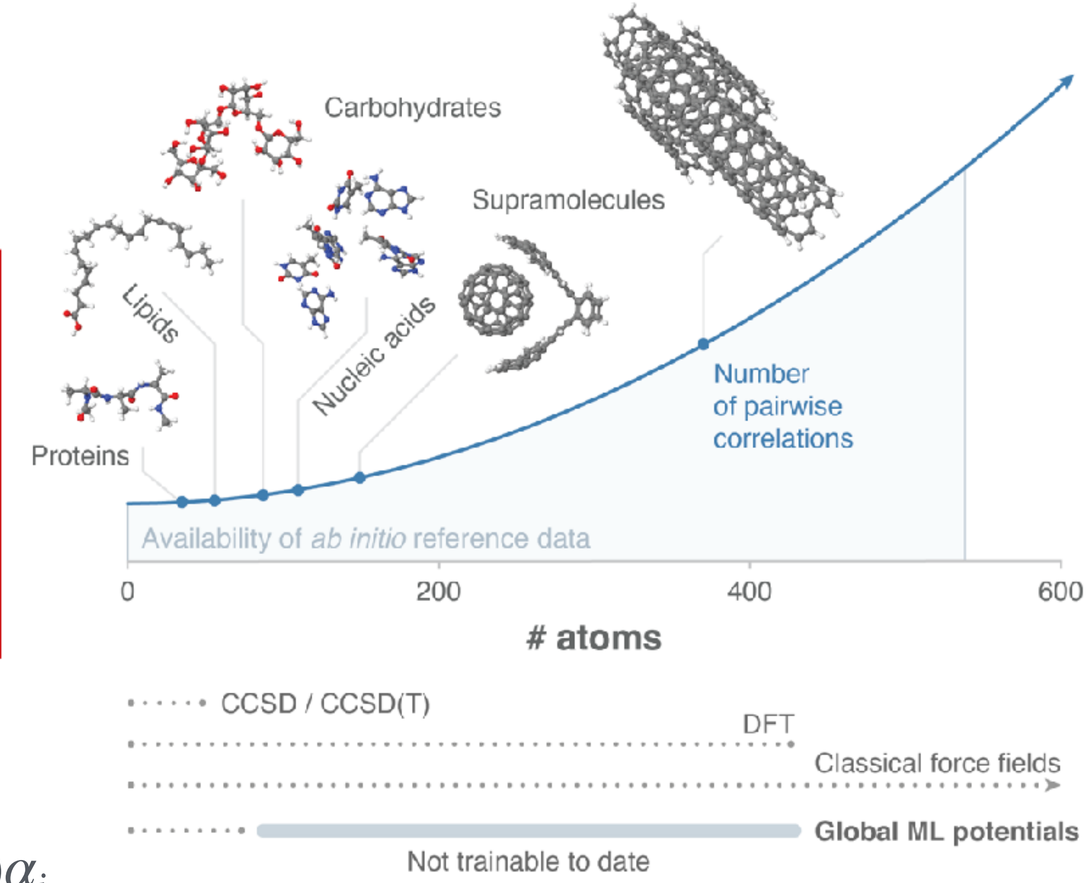
Force field reconstruction with ML

The most immediate application of ML in QM

All atoms interact:
quadratic memory complexity (minimum)







Linear model
Pairwise atom-correlations: $\hat{f}(\mathbf{x}) = \mathbf{K}\alpha = \sum_{i}^{M} k(\mathbf{x}, \mathbf{x}_i)\alpha_i$

Chmiela et al. 2023, Blücher et al. 2023

Correlation matrices

Numerical challenges in training

Complexity:
Memory: $\mathcal{O}(N^2)$
Time: $\mathcal{O}(N^3)$

Involves matrix diagonalization: $\alpha = (\mathbf{X}\mathbf{X}^\top + \lambda\mathbb{I})^{-1}\mathbf{y} = \left(\mathbf{L}\mathbf{L}^\top\right)^{-1}\mathbf{y}$ ← Cholesky decomposition

$\mathbf{X}$: data points
$\mathbf{K} = \mathbf{X}\mathbf{X}^\top$: symmetric, PSD

Alternative approach: iterative optimization using **matrix-vector products**

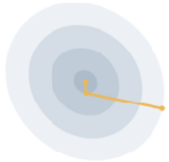
- Gradient descent: $\nabla_{\mathbf{w}} L(\hat{f}(\mathbf{x}), \mathbf{y}) = (\mathbf{X}\mathbf{X}^\top + \lambda\mathbb{I})\alpha^{t-1} - \mathbf{y}$

$$\alpha^t = \alpha^{t-1} - \gamma\left[(\mathbf{X}\mathbf{X}^\top + \lambda\mathbb{I})\alpha^{t-1} - \mathbf{y}\right]$$

γ: learning rate

Search direction

- Krylov subspace solver (**conjugate gradients**):



Dynamic learning rate: $\gamma_t = \dfrac{\mathbf{p}_t^\top \mathbf{y}}{\mathbf{p}_t^\top \mathbf{K}_\lambda \mathbf{p}_t}$

$$\mathbf{p}_t = \mathbf{r}_t + \frac{\mathbf{r}_t^\top \mathbf{r}_t}{\mathbf{r}_{t-1}^\top \mathbf{r}_{t-1}}\mathbf{p}_{t-1}$$

$\mathbf{p}_t$: search direction

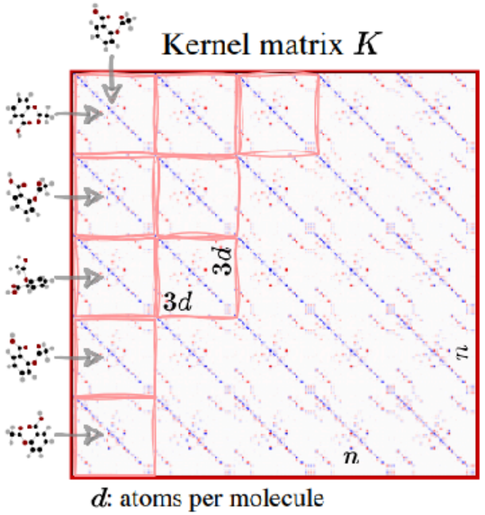
Conjugate optimization steps: $\mathbf{p}_t = \mathbf{r}_t - \sum_{i<t} \dfrac{\mathbf{p}_i^\top \mathbf{K}_\lambda \mathbf{r}_t}{\mathbf{p}_i^\top \mathbf{K}_\lambda \mathbf{p}_i}\mathbf{p}_i$

Only one of these is kept in memory at a time.

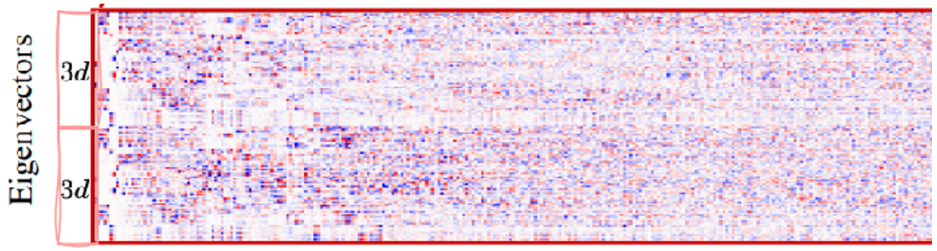
Rank-1 decomposition of $\mathbf{K} = \sum_t^N \mathbf{p}_t \mathbf{p}_t^\top$ ← N steps with perfect arithmetic!

Correlation matrices

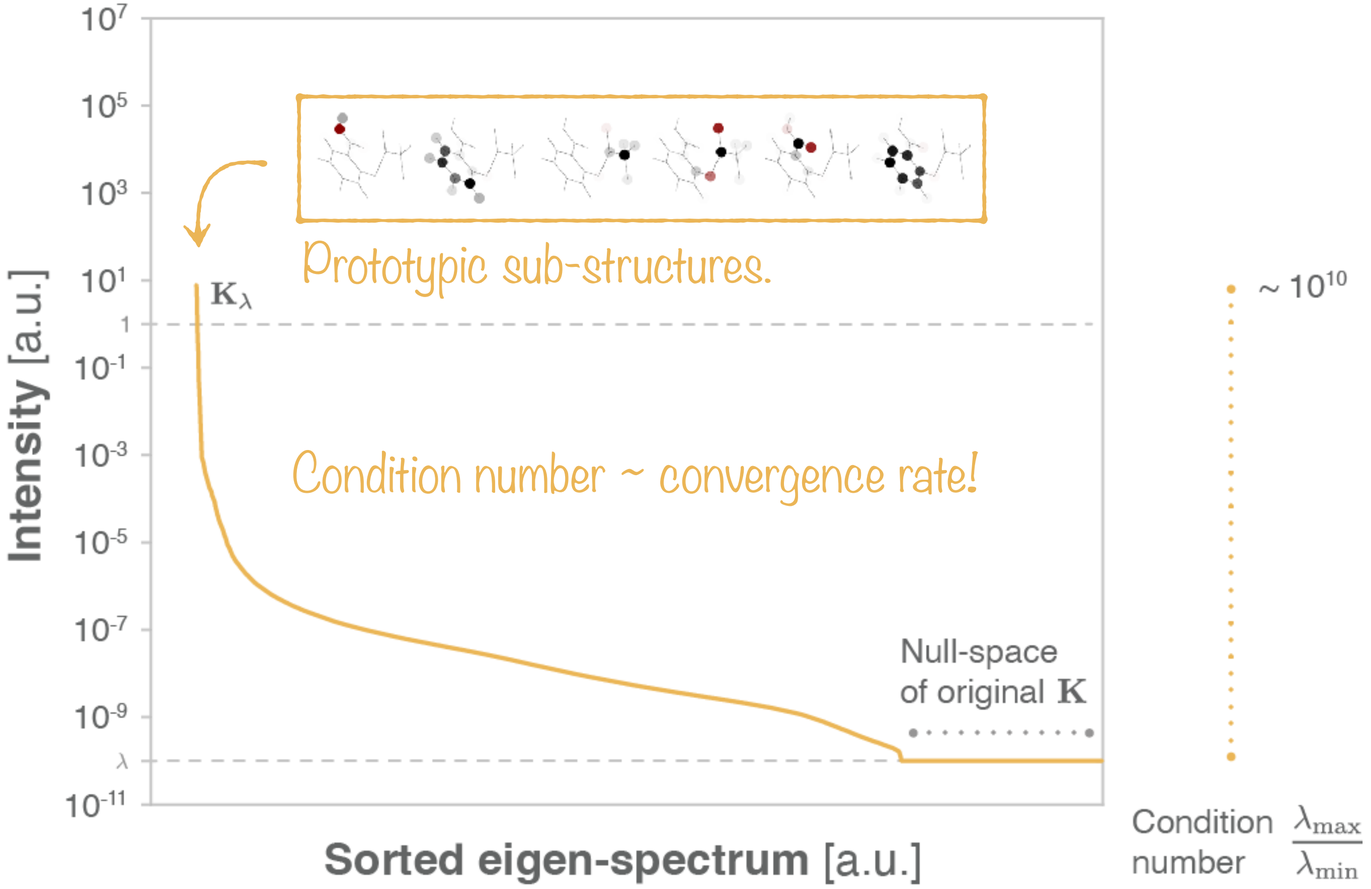
Strongly correlated reference geometries



Spectrum



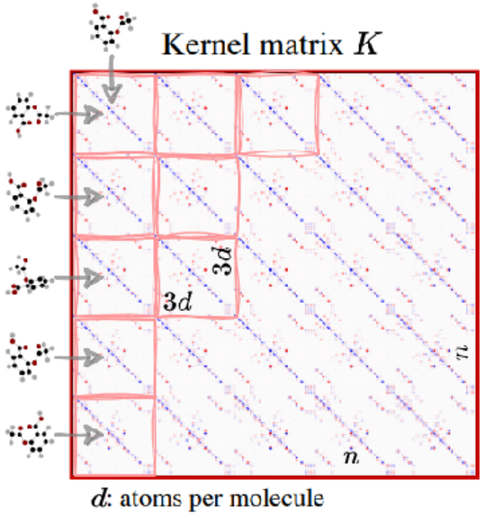
Dominant eigenvectors



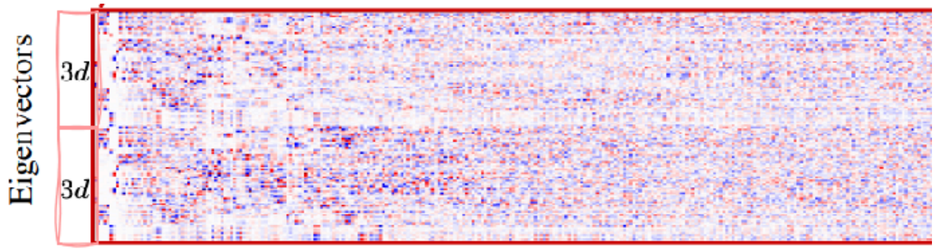
Chmiela et al. 2023, Blücher et al. 2023

Correlation matrices

Strongly correlated reference geometries

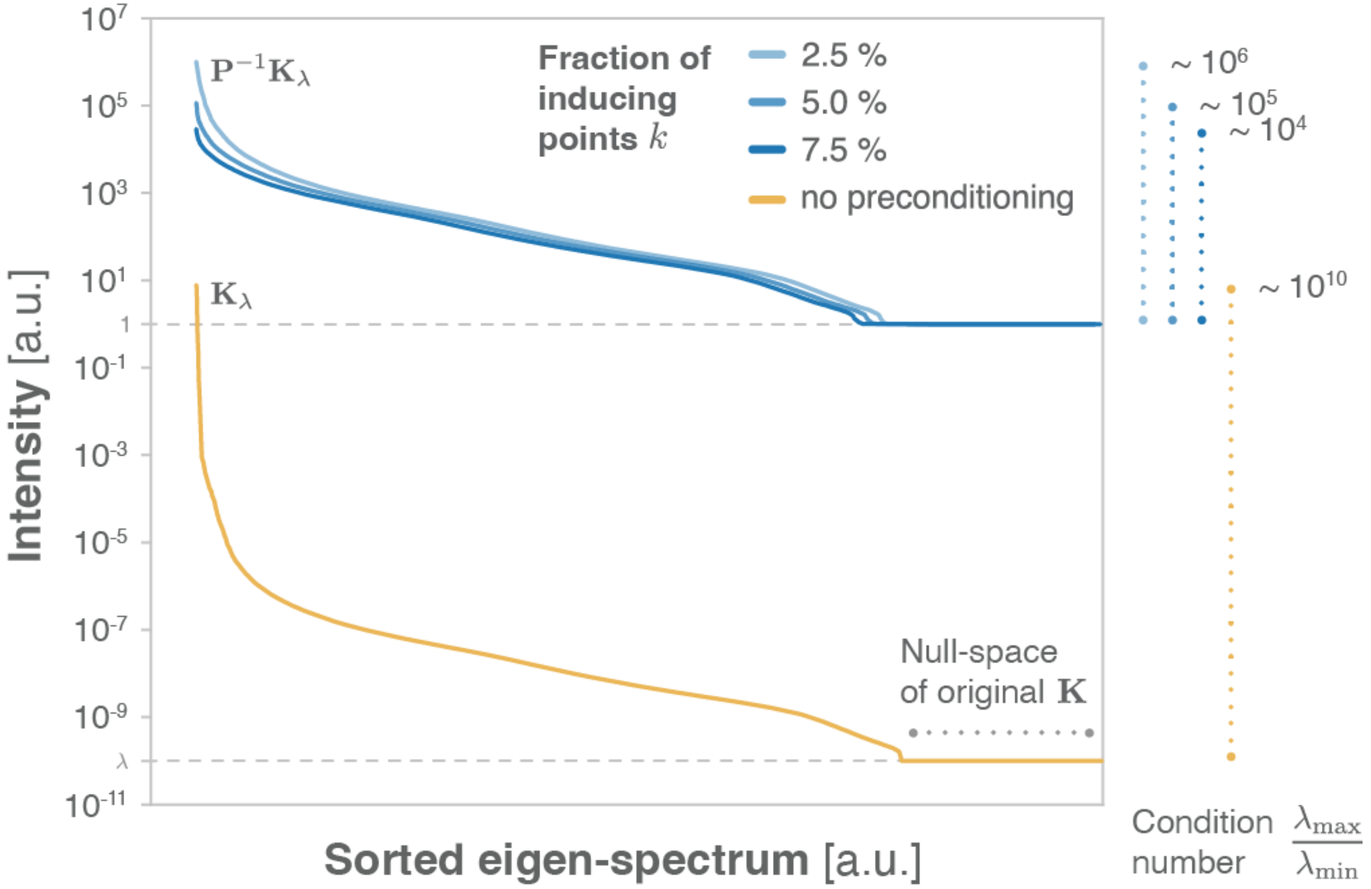


Spectrum



Dominant eigenvectors

Preconditioning captures dominant spectral components



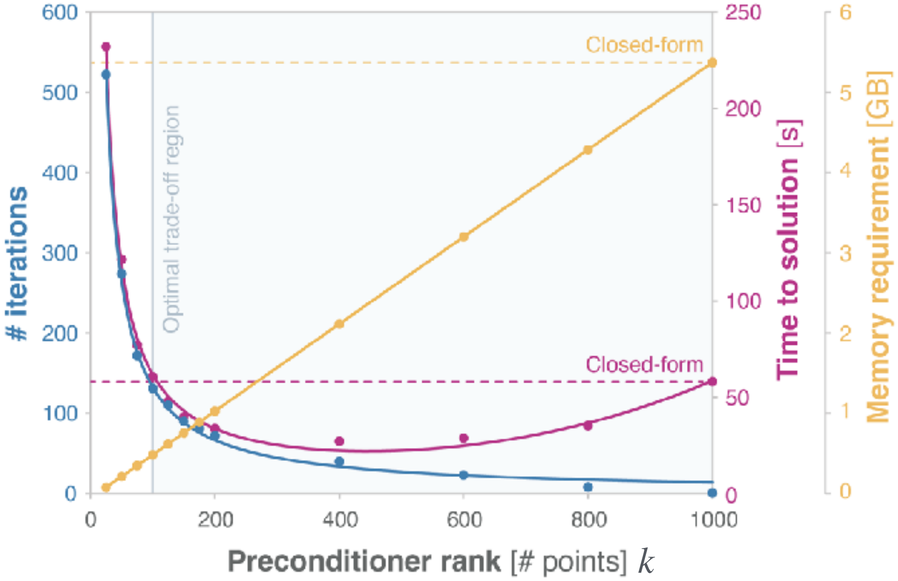
Chmiela et al. 2023, Blücher et al. 2023

Effective preconditioners for correlation matrices

Strongly correlated reference geometries

General idea: $\mathbf{K} = \underbrace{\sum_{t=1}^{k} \mathbf{u}_t \mathbf{u}_t^\top}_{\text{Precond.}} + \underbrace{\sum_{t=k+1}^{n} \mathbf{p}_t \mathbf{p}_t^\top}_{\text{CG decomp.}}$

$\bar{\mathbf{K}} = \mathbf{P}^{-1}\mathbf{K}$ of $\bar{\mathbf{K}}$



Cost ↓

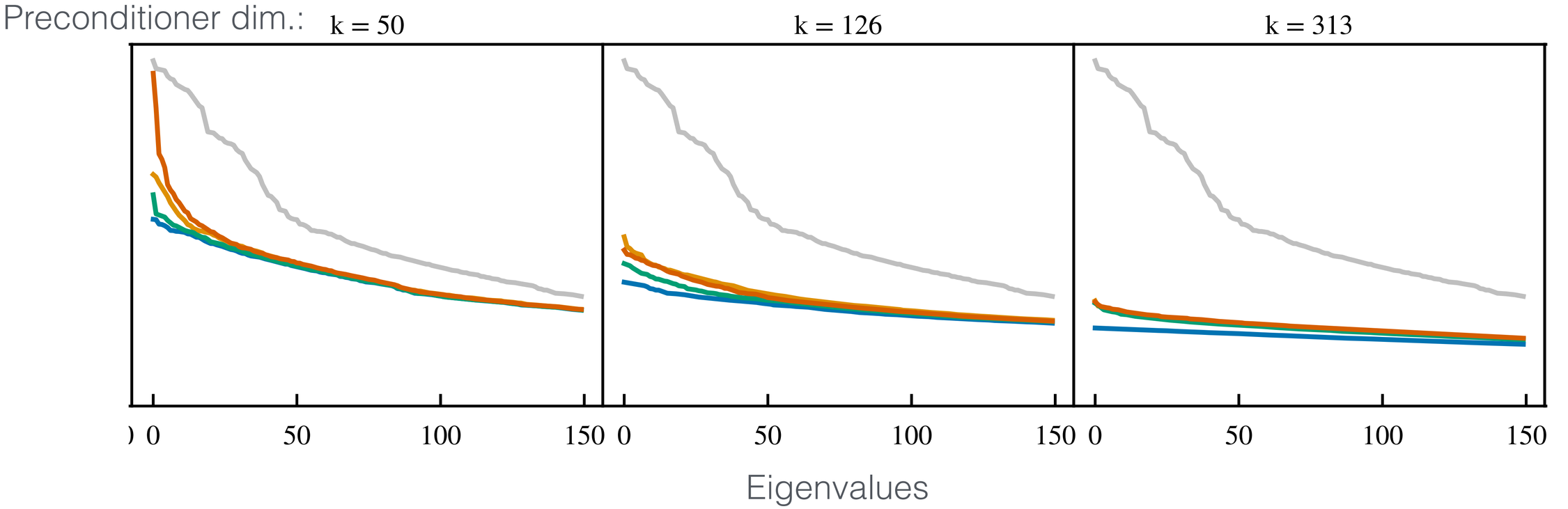
Computational cost / effectiveness trade-off: $\mathbf{K} \approx \hat{\mathbf{K}} + \mathbb{I}\lambda$

- Nyström approximation
  $\hat{\mathbf{K}} = \mathbf{K}_{mk}\mathbf{K}_{kk}^{-1}\mathbf{K}_{mk}^\top$
- Incomplete Cholesky decomposition
  $\hat{\mathbf{K}} = \mathbf{L}_{mk}\mathbf{L}_{mk}^\top$
- Eigen-decomposition
  $\hat{\mathbf{K}} = \mathbf{Q}_{mk}\mathbf{\Lambda}_{kk}\mathbf{Q}_{mk}^\top$

Chmiela et al. 2023, Blücher et al. 2023

Effective preconditioners for correlation matrices

Preconditioning methods (MD17 Aspirin)



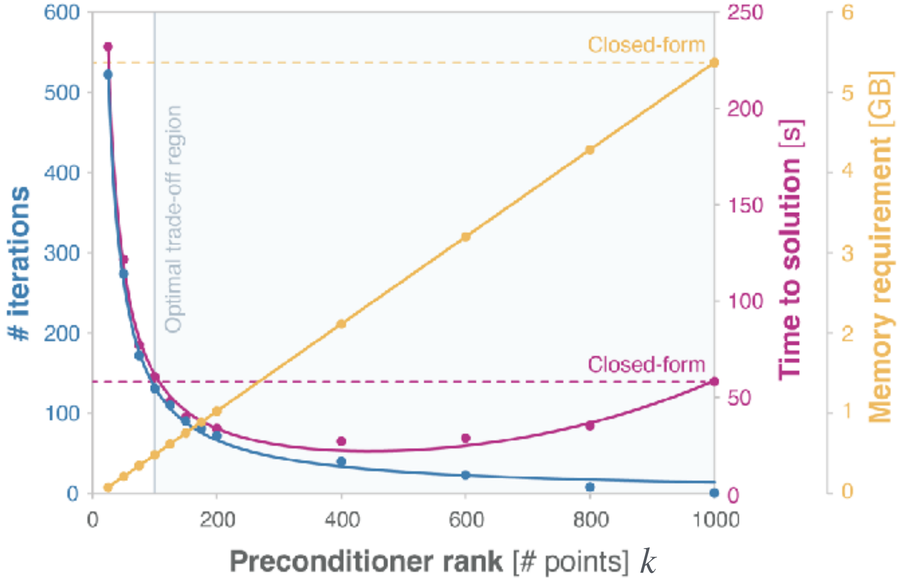
Chmiela et al. 2023, Blücher et al. 2023

Effective preconditioners for correlation matrices

Strongly correlated reference geometries

General idea: $\mathbf{K} = \sum_{t=1}^{k} \mathbf{u}_t \mathbf{u}_t^\top + \sum_{t=k+1}^{n} \mathbf{p}_t \mathbf{p}_t^\top$

Precond. $\bar{\mathbf{K}} = \mathbf{P}^{-1}\mathbf{K}$ — CG decomp. of $\bar{\mathbf{K}}$



Cost ↓

Computational cost / effectiveness trade-off: $\mathbf{K} \approx \hat{\mathbf{K}} + \mathbb{I}\lambda$

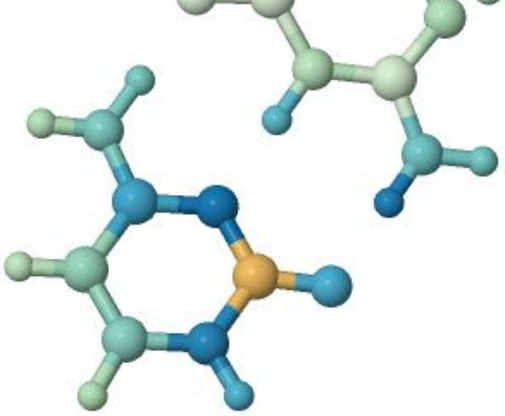
- Nyström approximation
  $\hat{\mathbf{K}} = \mathbf{K}_{mk}\mathbf{K}_{kk}^{-1}\mathbf{K}_{mk}^\top$ → Woodbury identity:
  $\mathbf{P}^{-1} = \lambda^{-1}\left[\mathbb{I} - \mathbf{K}_{mk}\left(\lambda\mathbf{K}_{kk} + \mathbf{K}_{mk}^\top\mathbf{K}_{mk}\right)^{-1}\mathbf{K}_{mk}^\top\right]$
- Incomplete Cholesky decomposition
  $\hat{\mathbf{K}} = \mathbf{L}_{mk}\mathbf{L}_{mk}^\top$
- Eigen-decomposition
  $\hat{\mathbf{K}} = \mathbf{Q}_{mk}\mathbf{\Lambda}_{kk}\mathbf{Q}_{mk}^\top$

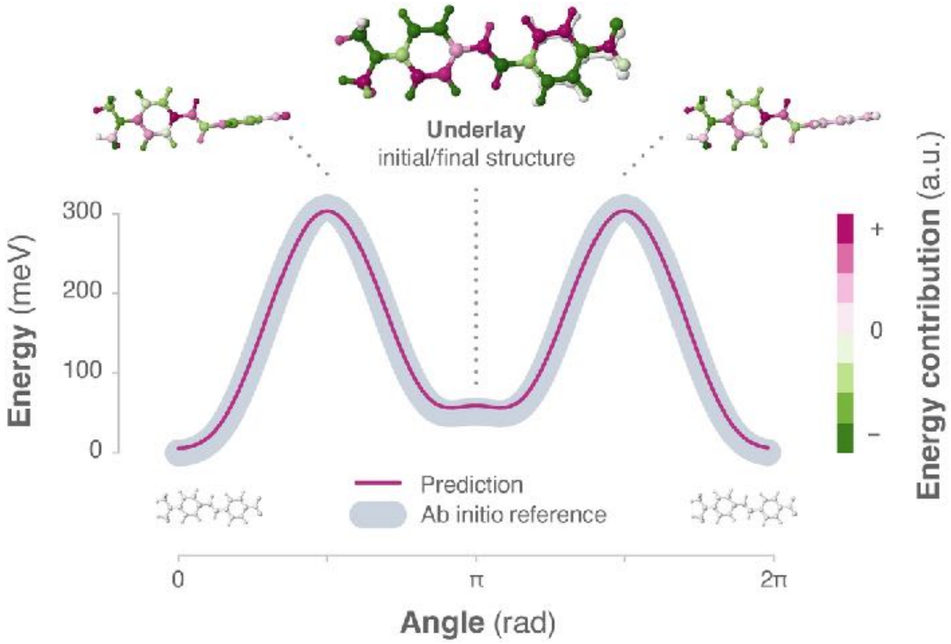
Stable decomposition into Cholesky factors, indirectly using $\mathbf{K}_{mk}^\top\mathbf{K}_{mk} = \mathbf{R}^\top\mathbf{Q}^\top\mathbf{Q}\mathbf{R} = \mathbf{Q}\mathbf{Q}^\top$

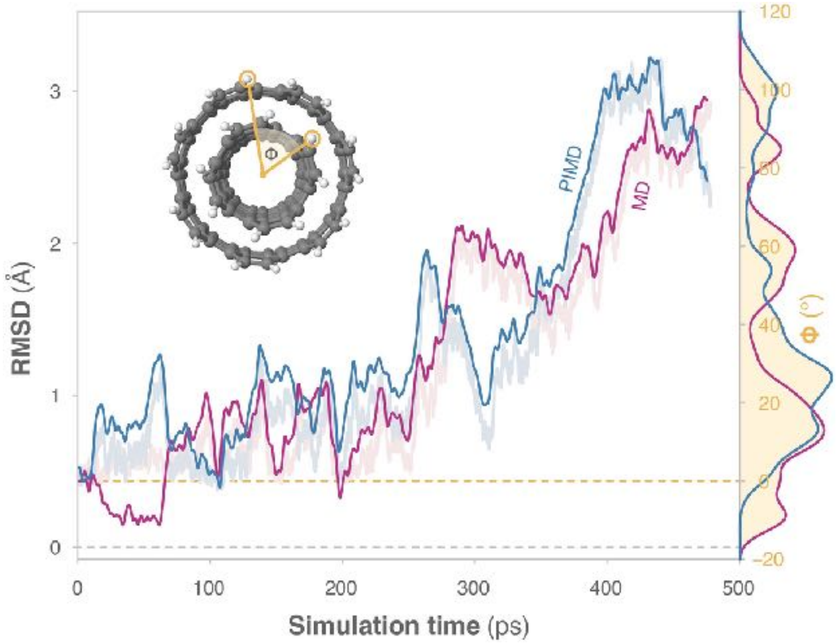
Chmiela et al. 2023, Blücher et al. 2023

Global force fields with hundreds of atoms

A new frontier for machine learning







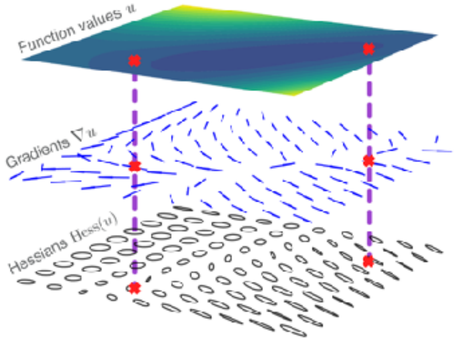
Examples: Conjugated system with delocalized electrons.

Long timescale (PI)MD of supramolecular complexes

www.sgdml.org

Chmiela, S., Vassilev-Galindo, V., Unke, O. T., Kabylda, A., Sauceda, H. E., Tkatchenko, A., & Müller, K. R. (2023). Accurate global machine learning force fields for molecules with hundreds of atoms. Science Advances, 9(2), eadf0873.

B



Efficient contraction of constrained models

Differential constraints in ML models with AD

Stronger inductive biases

Differential equations: $\mathscr{L}u(\mathbf{x}) = \mathbf{f}(\mathbf{x})$,
where $\mathscr{L} = \sum_{|j| \leq n} a_j \mathscr{D}^j$ is a finite linear combination of differential operators $\mathscr{D}^j$ or order n.

- PDE's describe **causal systems** with **interactions**
- Learn from **few examples** and **generalize**

Challenges:
- Tedious manual implementation: "*one new constraint per paper*"
- Constraints often increase model complexity: e.g. $\mathscr{L} = \nabla_{\mathbf{x}}$

C_u: time complexity of function u

Algorithmic differentiation: Express operators in terms of JVPs $\mathbf{J}_u(\mathbf{x})\mathbf{v}$ and VJPs $\mathbf{J}_u^\mathsf{T}(\mathbf{x})\mathbf{w}$ at $\mathcal{O}(C_u)$!

Example: Hessian-vector products in only $\mathcal{O}(C_u)$!

Schmitz, N. F., Müller, K. R., & Chmiela, S. (2022). Algorithmic Differentiation for Automated Modeling of Machine Learned Force Fields. The Journal of Physical Chemistry Letters, 13(43), 10183-10189.

Differential constraints in ML models with AD

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}\left(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\right)$$

Gaussian process example

Constrained GP: $\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}\left(\mathscr{L}\mu(\mathbf{x}), \mathscr{L}_{\mathbf{x}} \otimes \mathscr{L}_{\mathbf{x}'}^{\top} k(\mathbf{x}, \mathbf{x}')\right)$

mean prior

covariance prior

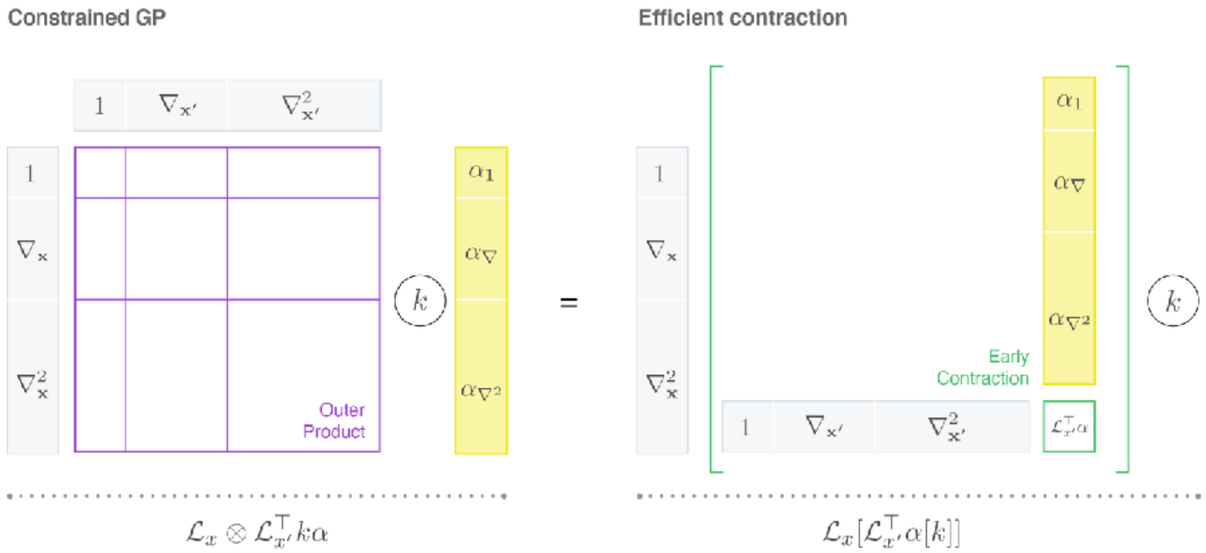
Example: $\mathscr{L} = (1, \nabla, \nabla^2)$

Constraints: function values, gradients & Hessians

$$\mathcal{L}_{\mathbf{x}} \otimes \mathcal{L}_{\mathbf{x}'}^{\top} k = \left[\begin{array}{c|c|c} k & \nabla_{\mathbf{x}'} k & \nabla_{\mathbf{x}'}^2 k \\ \hline \nabla_{\mathbf{x}} k & \nabla_{\mathbf{x}} \otimes \nabla_{\mathbf{x}'} k & \nabla_{\mathbf{x}} \otimes \nabla_{\mathbf{x}'}^2 k \\ \hline \nabla_{\mathbf{x}}^2 k & \nabla_{\mathbf{x}}^2 \otimes \nabla_{\mathbf{x}'} k & \nabla_{\mathbf{x}}^2 \otimes \nabla_{\mathbf{x}'}^2 k \end{array}\right]$$

large kernel matrix (but with structure)

tensor-product structure



Full instantiation of kernel can be avoided with AD:

$$\mathbf{f}(\mathbf{x}) = \sum_i \mathscr{L}_{\mathbf{x}}\left[\mathscr{L}_{\mathbf{x}_i}^{\top} \alpha_i k(\mathbf{x}, \mathbf{x}_i)\right]$$

Schmitz et al. 2022

Differential constraints in ML models with AD

Gaussian process example

Constrained GP: $\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}\left(\mathscr{L}\mu(\mathbf{x}), \mathscr{L}_{\mathbf{x}} \otimes \mathscr{L}_{\mathbf{x}'}^{\mathsf{T}} k\left(\mathbf{x}, \mathbf{x}'\right)\right)$

Factor N reducing in complexity after contraction!

$\mathcal{L}_{\mathbf{x}}$	$\mathcal{L}_{\mathbf{x}'}$	$\left[\mathcal{L}_{\mathbf{x}} \otimes \mathcal{L}_{\mathbf{x}'}^{\mathsf{T}} k(\mathbf{x}, \mathbf{x}')\right]$	cost	$\mathcal{L}_{\mathbf{x}}\left[\mathcal{L}_{\mathbf{x}'}^{\mathsf{T}} \boldsymbol{\alpha}\left[k(\mathbf{x}, \mathbf{x}')\right]\right]$	cost
∇	∇	$\mathrm{Jac}_{\mathbf{x}}(\mathrm{Grad}_{\mathbf{x}'}(k))^{\mathsf{T}}$	NC_k	$\mathrm{Grad}_{\mathbf{x}}(\mathrm{Grad}_{\mathbf{x}'}(k)\boldsymbol{\alpha})$	C_k
∇	∇^2	$\mathrm{Jac}_{\mathbf{x}}(\mathrm{Hess}_{\mathbf{x}'}(k))^{\mathsf{T}}$	N^2C_k	$\mathrm{Grad}_{\mathbf{x}}(\mathrm{Hess}_{\mathbf{x}'}(k)\boldsymbol{\alpha})$	NC_k
∇^2	∇^2	$\mathrm{Hess}_{\mathbf{x}}(\mathrm{Hess}_{\mathbf{x}'}(k))^{\mathsf{T}}$	N^3C_k	$\mathrm{Hess}_{\mathbf{x}}(\mathrm{Hess}_{\mathbf{x}'}(k)\boldsymbol{\alpha})$	N^2C_k

Cost of evaluating the (scalar) base kernel.

Schmitz, N. F., Müller, K. R., & Chmiela, S. (2022). Algorithmic Differentiation for Automated Modeling of Machine Learned Force Fields. The Journal of Physical Chemistry Letters, 13(43), 10183-10189.

Differential constraints in ML models with AD

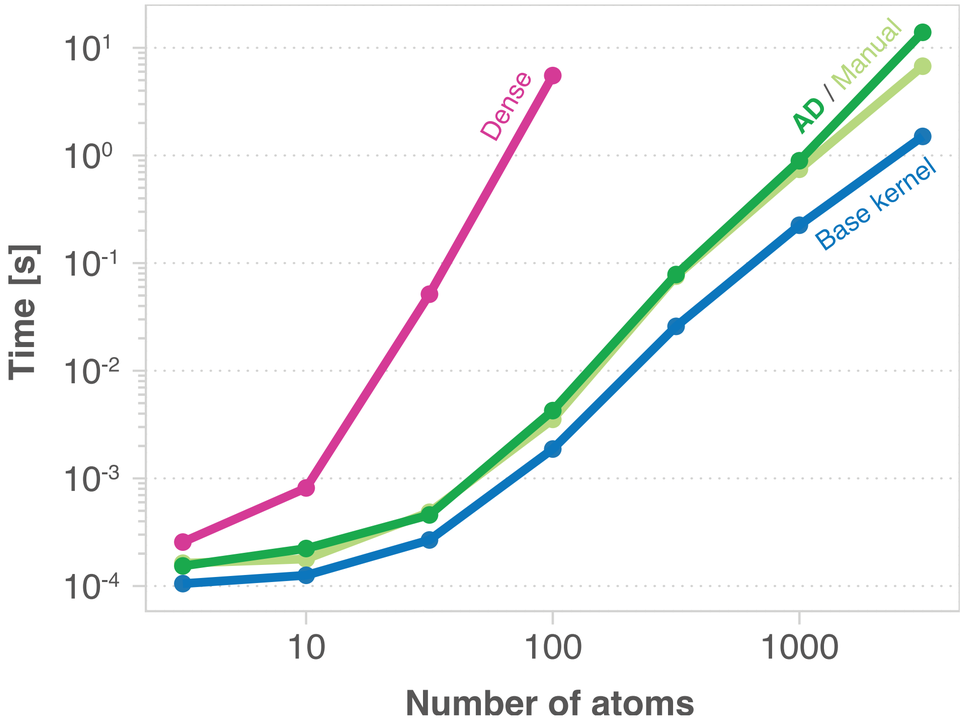
Gaussian process example

Constrained GP: $\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}\left(\mathcal{L}\mu(\mathbf{x}), \mathcal{L}_{\mathbf{x}} \otimes \mathcal{L}_{\mathbf{x}'}^{\top} k\left(\mathbf{x}, \mathbf{x}'\right)\right)$

$\mathcal{L}_{\mathbf{x}}$	$\mathcal{L}_{\mathbf{x}'}$	$\left[\mathcal{L}_{\mathbf{x}} \otimes \mathcal{L}_{\mathbf{x}'}^{\top} k(\mathbf{x}, \mathbf{x}')\right]$	cost	$\mathcal{L}_{\mathbf{x}}\left[\mathcal{L}_{\mathbf{x}'}^{\top} \boldsymbol{\alpha}\left[k(\mathbf{x}, \mathbf{x}')\right]\right]$	cost
∇	∇	$\mathrm{Jac}_{\mathbf{x}}(\mathrm{Grad}_{\mathbf{x}'}(k))^{\top}$	NC_k	$\mathrm{Grad}_{\mathbf{x}}(\mathrm{Grad}_{\mathbf{x}'}(k)\boldsymbol{\alpha})$	C_k
∇	∇^2	$\mathrm{Jac}_{\mathbf{x}}(\mathrm{Hess}_{\mathbf{x}'}(k))^{\top}$	N^2C_k	$\mathrm{Grad}_{\mathbf{x}}(\mathrm{Hess}_{\mathbf{x}'}(k)\boldsymbol{\alpha})$	NC_k
∇^2	∇^2	$\mathrm{Hess}_{\mathbf{x}}(\mathrm{Hess}_{\mathbf{x}'}(k))^{\top}$	N^3C_k	$\mathrm{Hess}_{\mathbf{x}}(\mathrm{Hess}_{\mathbf{x}'}(k)\boldsymbol{\alpha})$	N^2C_k

Gradient constraints in GPs:

Same complexity class as the unconstrained model!

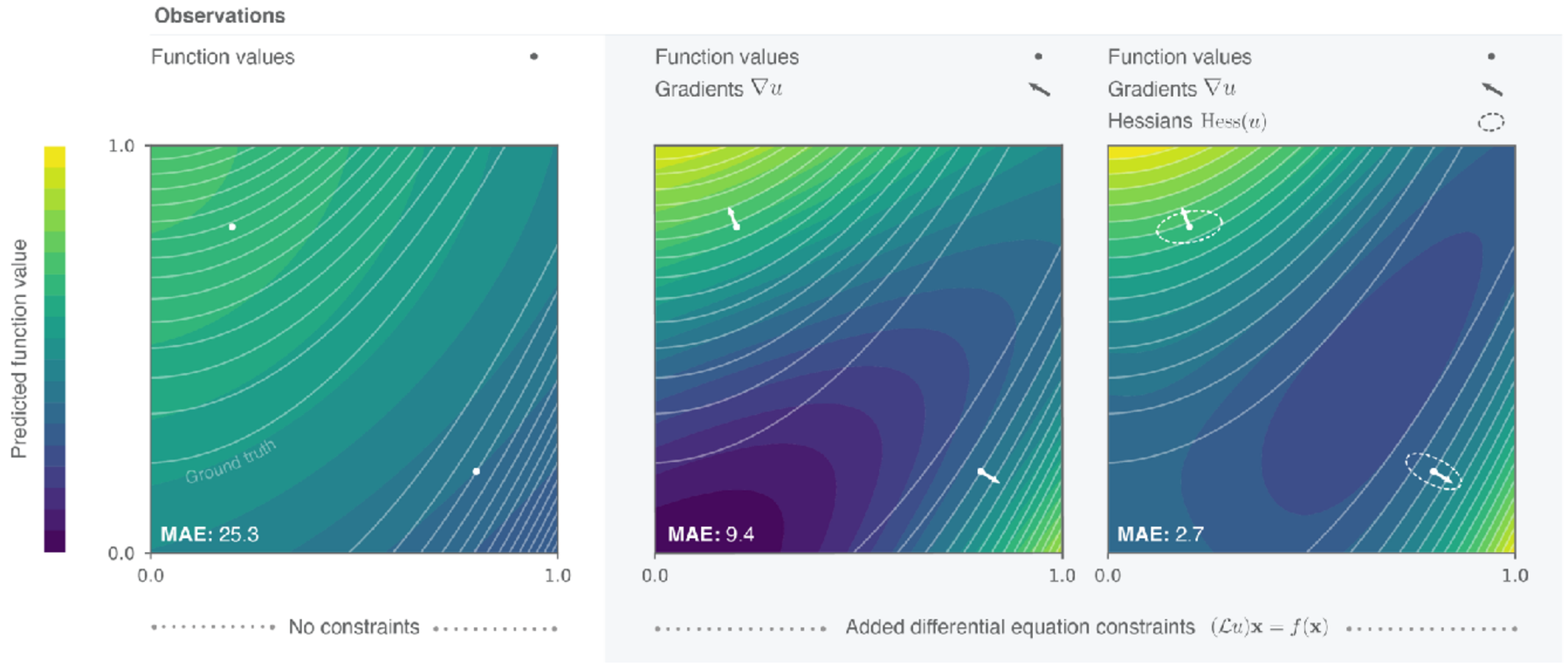


Schmitz, N. F., Müller, K. R., & Chmiela, S. (2022). Algorithmic Differentiation for Automated Modeling of Machine Learned Force Fields. The Journal of Physical Chemistry Letters, 13(43), 10183-10189.

Differential constraints in ML models with AD

Gaussian process example

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}\left(\mathcal{L}\mu(\mathbf{x}), \mathcal{L}_{\mathbf{x}} \otimes \mathcal{L}_{\mathbf{x}'}^{\top} k(\mathbf{x}, \mathbf{x}')\right)$$

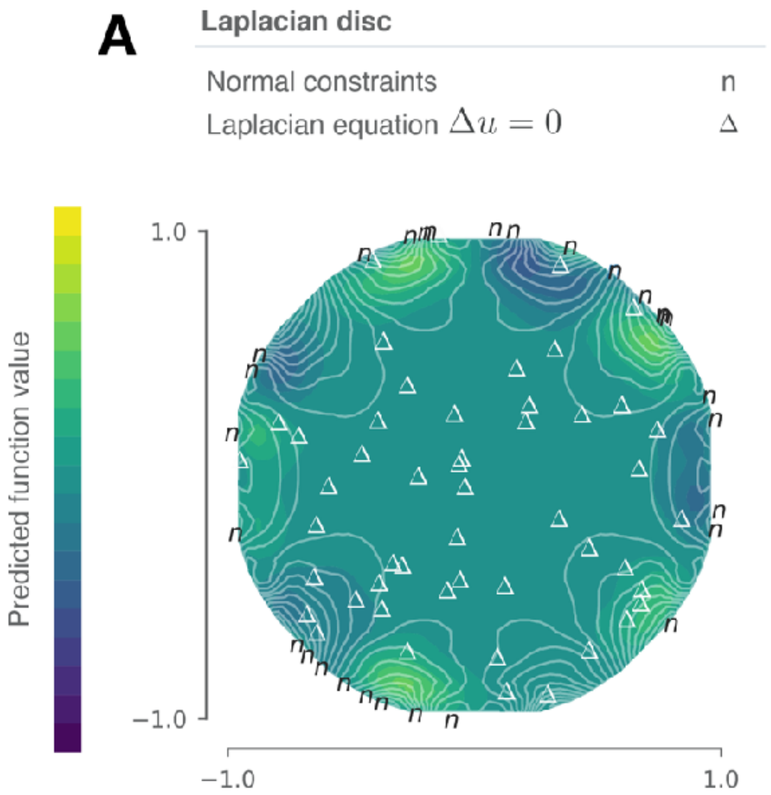


2D Rosenbrock function: $u(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2$

Schmitz, N. F., Müller, K. R., & Chmiela, S. (2022). Algorithmic Differentiation for Automated Modeling of Machine Learned Force Fields. The Journal of Physical Chemistry Letters, 13(43), 10183-10189.

Differential constraints in ML models with AD

Combining differential equation constraints in GP's



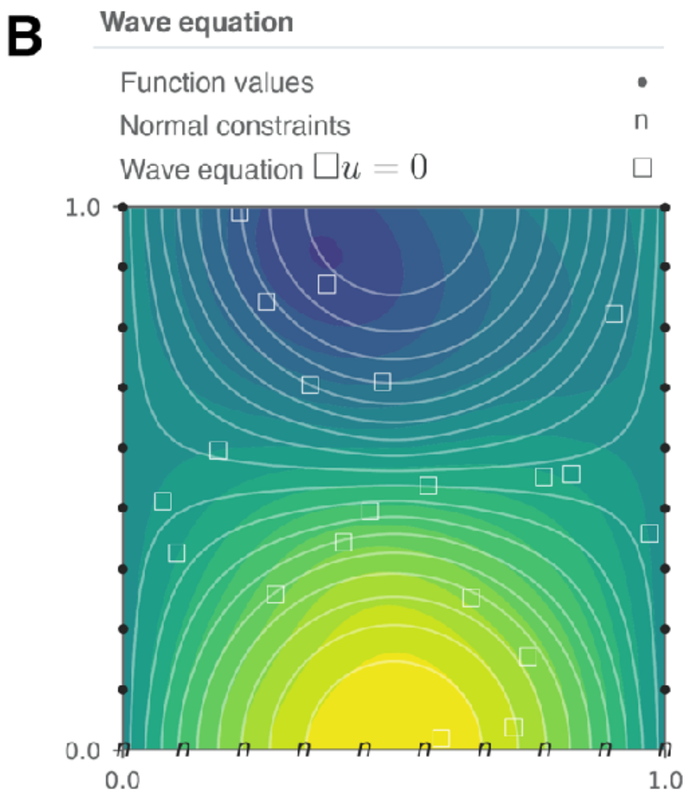
Solving Laplace's equation $\Delta u(x) = 0$ on the unit disk

- Neumann boundary condition $\nabla u(x) \cdot \mathbf{n}(x) = \cos(5\phi)$
- ϕ: radial angle of x
- $\mathbf{n}(x)$: boundary normal vector

Schmitz, N. F., Müller, K. R., & Chmiela, S. (2022). Algorithmic Differentiation for Automated Modeling of Machine Learned Force Fields. The Journal of Physical Chemistry Letters, 13(43), 10183-10189.

Differential constraints in ML models with AD

Combining differential equation constraints in GP's



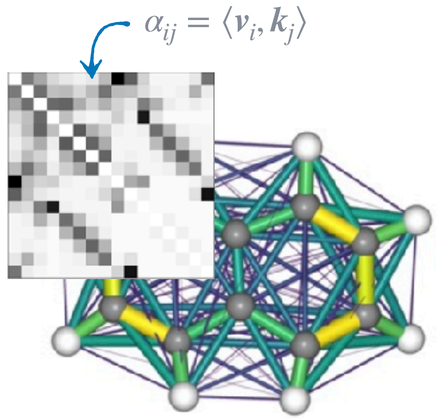
1D wave equation $\square u(x,t) = 0,\ \square = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}$

Dirichlet boundary conditions:
$u(0,t) = u(1,t) = 0,\ u(x,0) = x(1-x)$

Initial Neumann boundary condition:
$\partial_t u(x,0) = 0$

Schmitz, N. F., Müller, K. R., & Chmiela, S. (2022). Algorithmic Differentiation for Automated Modeling of Machine Learned Force Fields. The Journal of Physical Chemistry Letters, 13(43), 10183-10189.

C



Interactions from self-attention

Self-attention mechanism

Basic definition

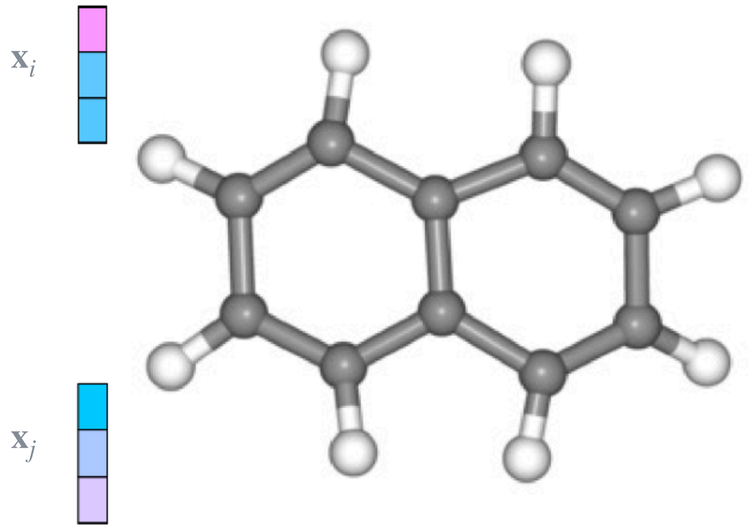
$$\mathbf{q}_i = \mathbf{W}_q \boldsymbol{x}_i \quad \mathbf{k}_i = \mathbf{W}_k \boldsymbol{x}_i \quad \boldsymbol{v}_i = \mathbf{W}_v \boldsymbol{x}_i$$

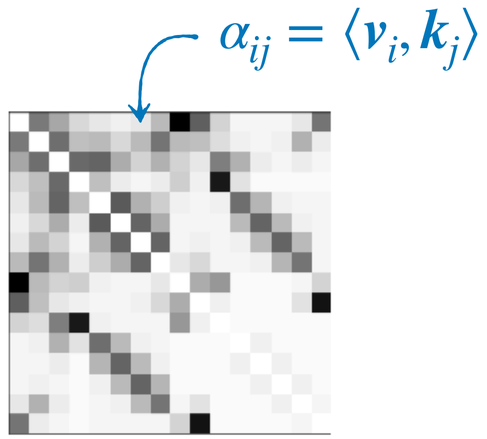
"queries" "keys" "values"

$\mathbf{x}$: atom embeddings
$\mathbf{W}$: learnable weights

$$\mathbf{A} = \text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V}$$

$\mathbf{Q}, \mathbf{K}, \mathbf{V}$: query, key and value embeddings for each atom
d: embedding dimension

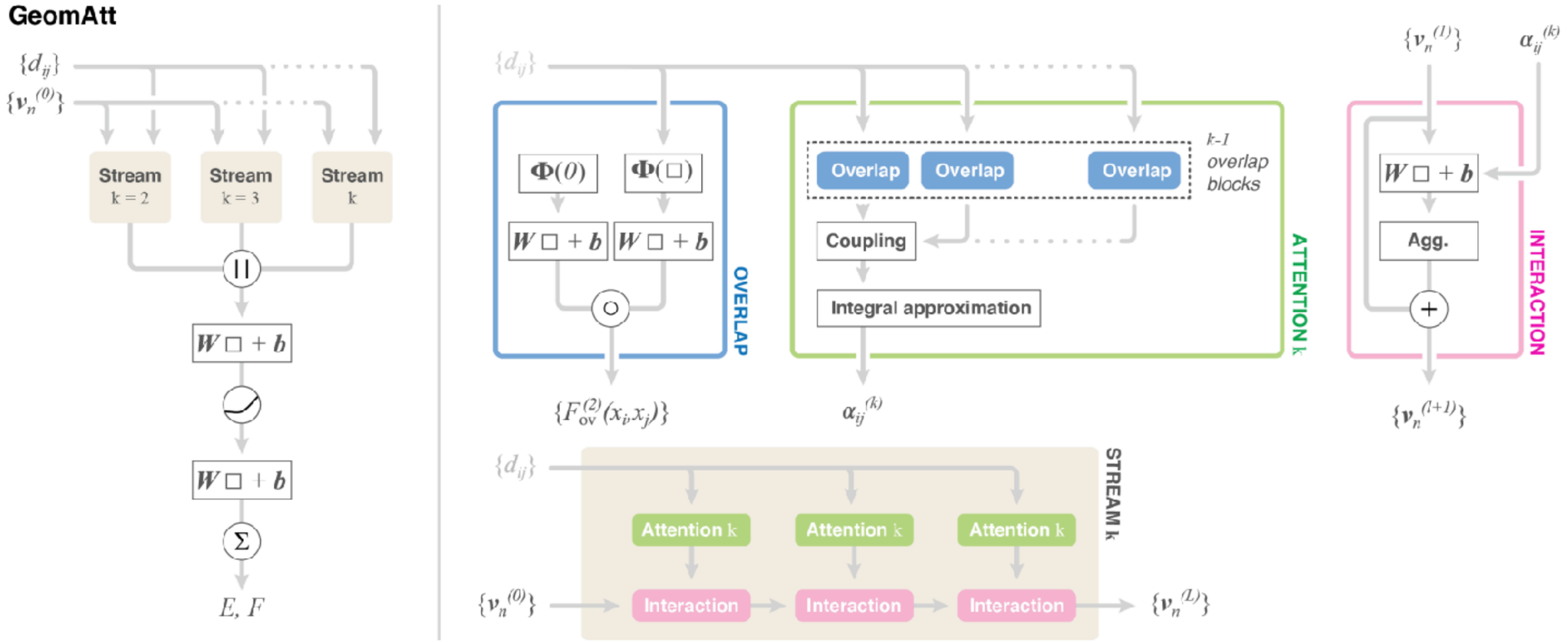




Frank, Thorben et al. "GeoPaTra: An Equivariant Transformer for Atomic Interactions on Arbitrary Length-Scales", 2021

Interactions from self-attention

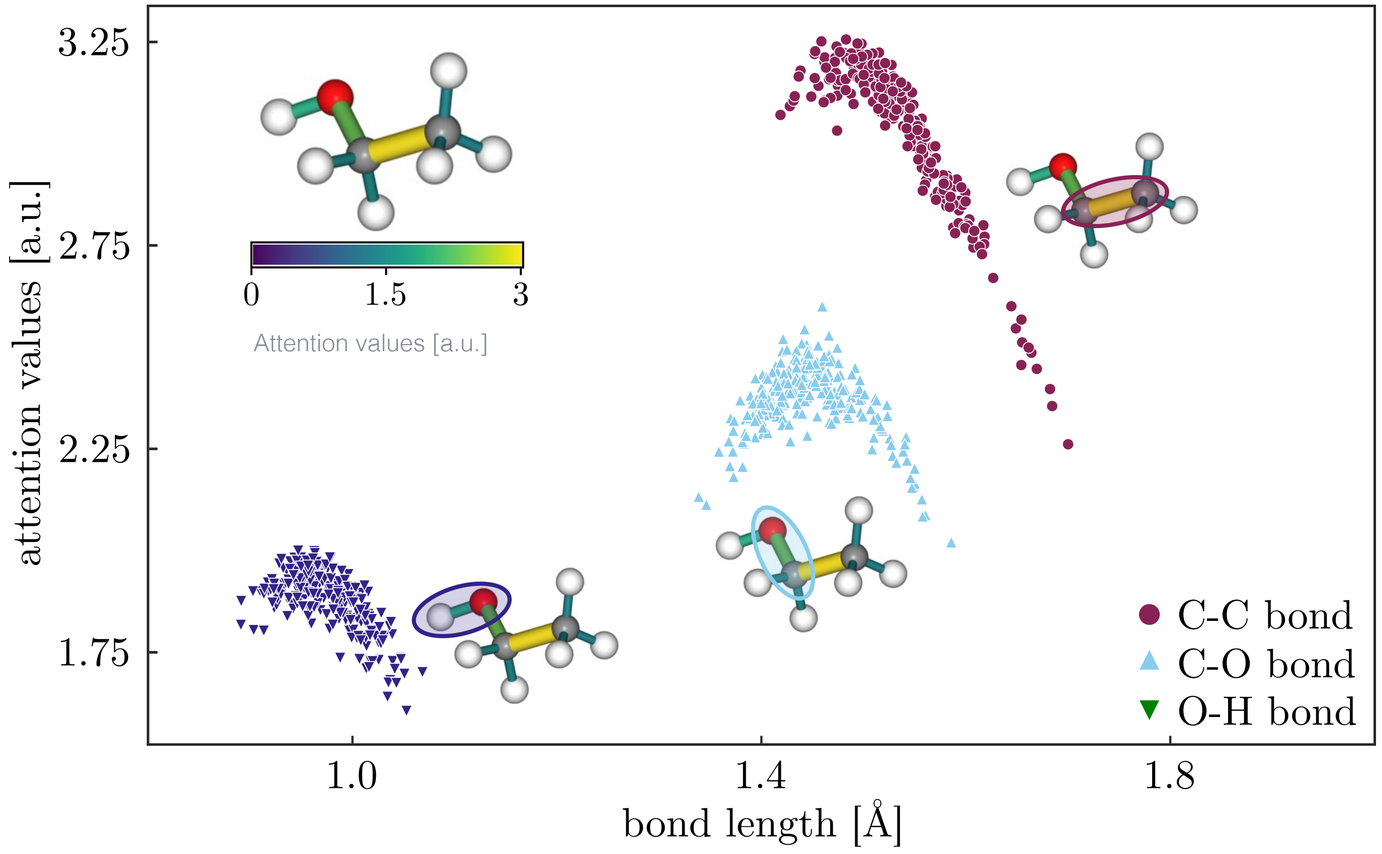
Neural network architecture

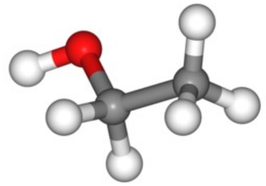


Frank, Thorben et al. "GeoPaTra: An Equivariant Transformer for Atomic Interactions on Arbitrary Length-Scales", 2021

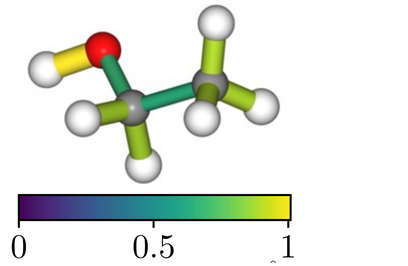
Emerging non-linear couplings between atoms

Attention coefficients vs. bond distances





Covalent bonds

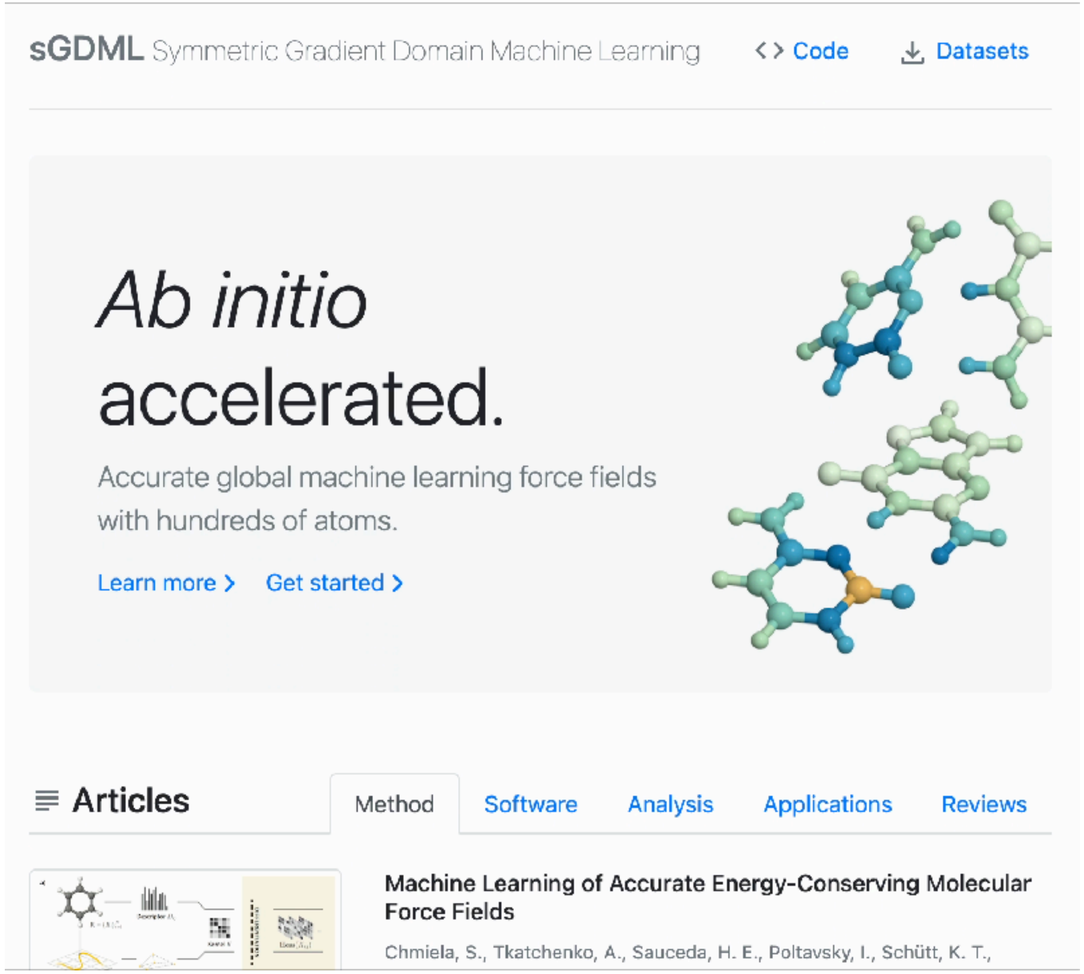


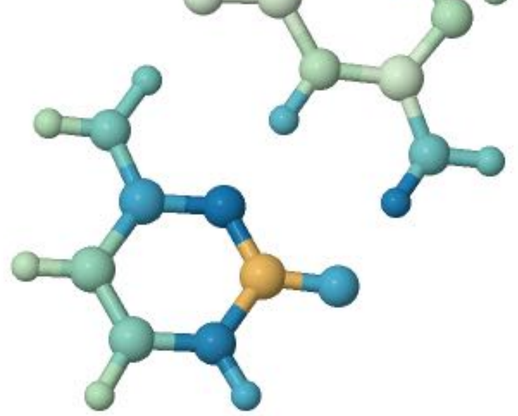
Inverse distances [1/Å]

Frank, Thorben et al. "GeoPaTra: An Equivariant Transformer for Atomic Interactions on Arbitrary Length-Scales", 2021

Want to know more?

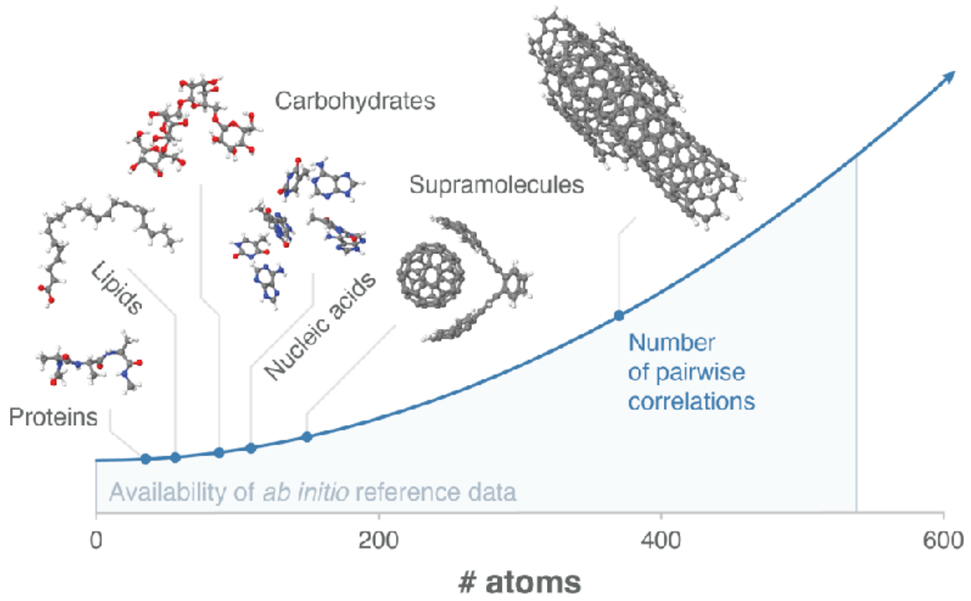
Articles, datasets & code





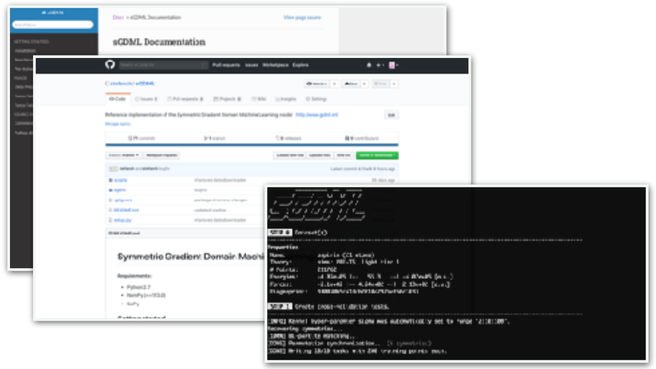
← **MD22 benchmark dataset**

- four major classes of biomolecules and supramolecules
- up to 370 atoms



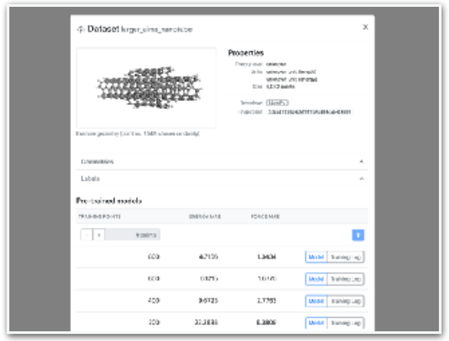
Want to know more?

Articles, datasets & code



Parametrize your own datasets
and use them as a force field.

OR



Upload your dataset and
let us do the training.

www.sgdml.org