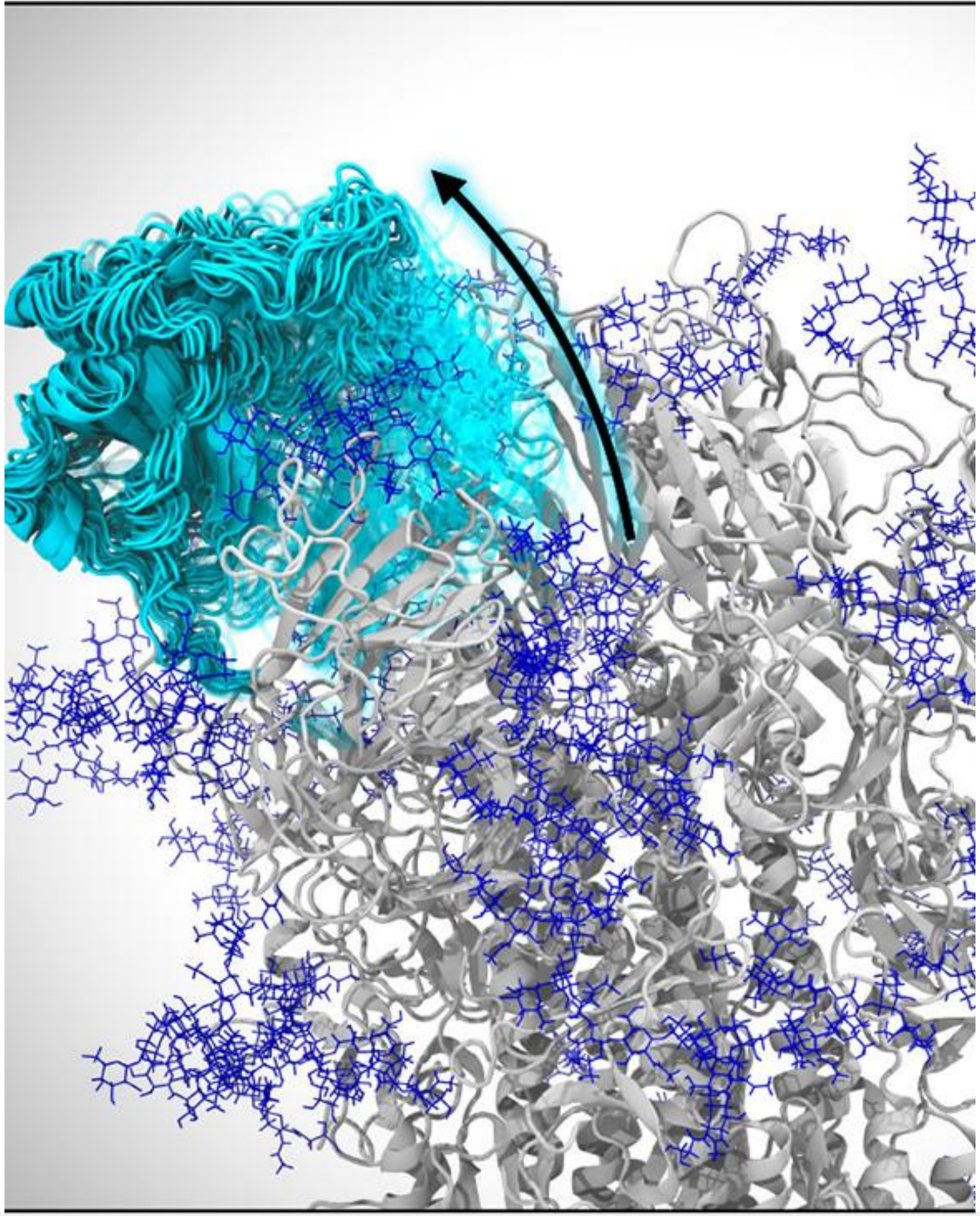
AI-driven Multiphysics HPC Applications: bridging the gap between HPC and ML

Dr. Massoud Rezavand, AI Software Solutions Engineer

Motivation

Motivation

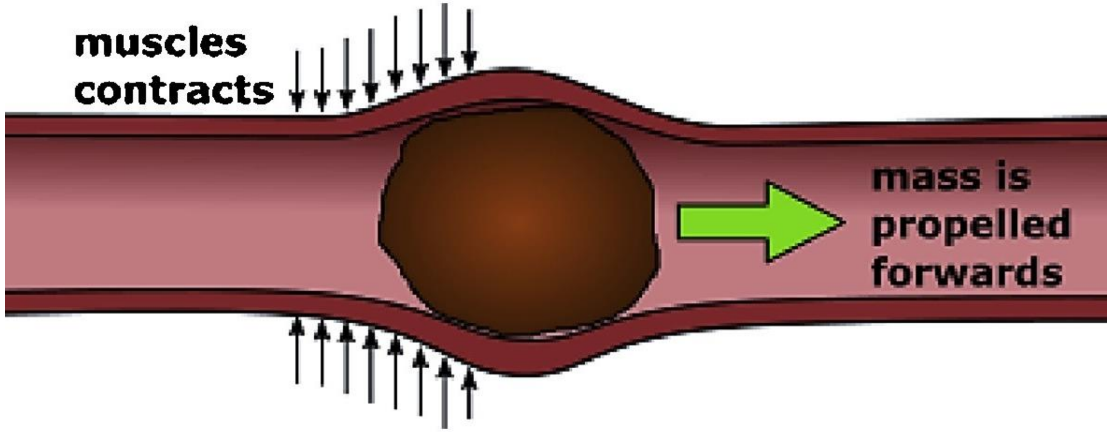
- HPC generates large amounts of data
  - typically, up to hundreds of TBs for atomistic level simulations [1]
- Resource-intensive simulations could be computationally prohibitive
  - 19.64 days on 100 GPUs for ~51.5 μs of comprehensive sampling of spike protein opening of the SARS-CoV-2 virus [1]

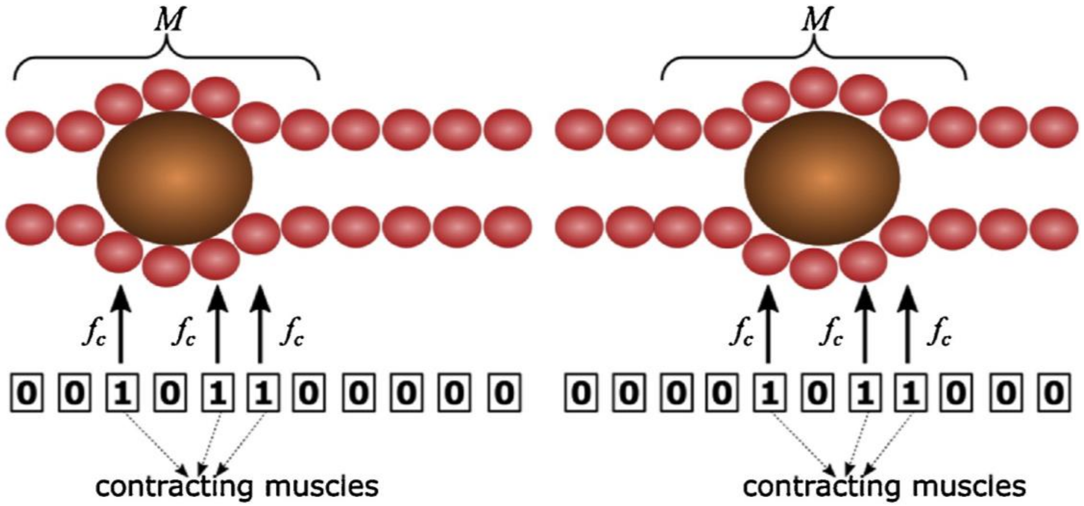


[1] Casalino et al., IJHPCA, 2021, https://doi.org/10.1177%2F10943420211006452

Motivation

- The underlying physics is sometimes poorly known
  - the learning abilities of ML methods from the observations could → address problems with more accurate empirical descriptions [2].
  - Couple DEM/SPH with RL to study Esophagus peristalsis [3]





[2] Alexiadis et al., Sci Rep, 2020, https://doi.org/10.1038/s41598-020-73329-0

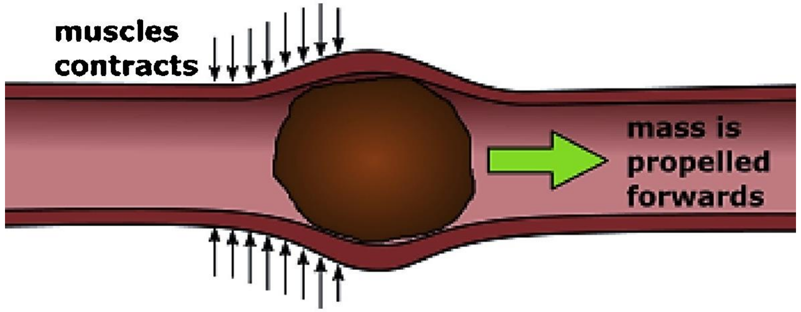
[3] Alexiadis, Artificial Intelligence In Medicine, 2019, https://doi.org/10.1016/j.artmed.2019.06.005

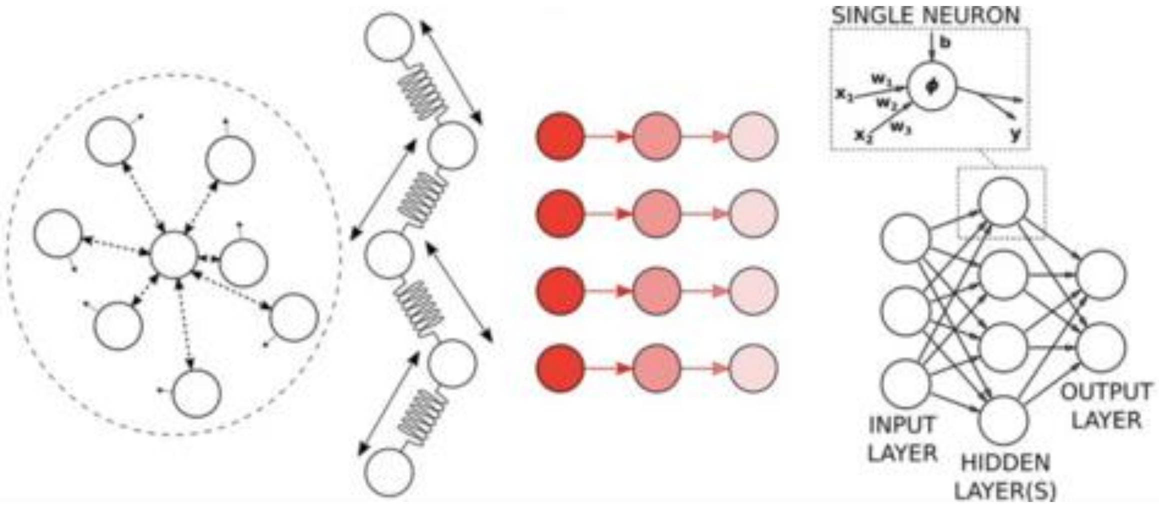
Examples on Intel® XPUs

Deep Multiphysics (DeepMP)

Deep Multiphysics (DeepMP)

- Goal: simulate the intestine peristalsis
- Muscle contractions in HPC is limited to math formulations
- Realistic simulations require interaction with environment
- The particle-neuron duality concept

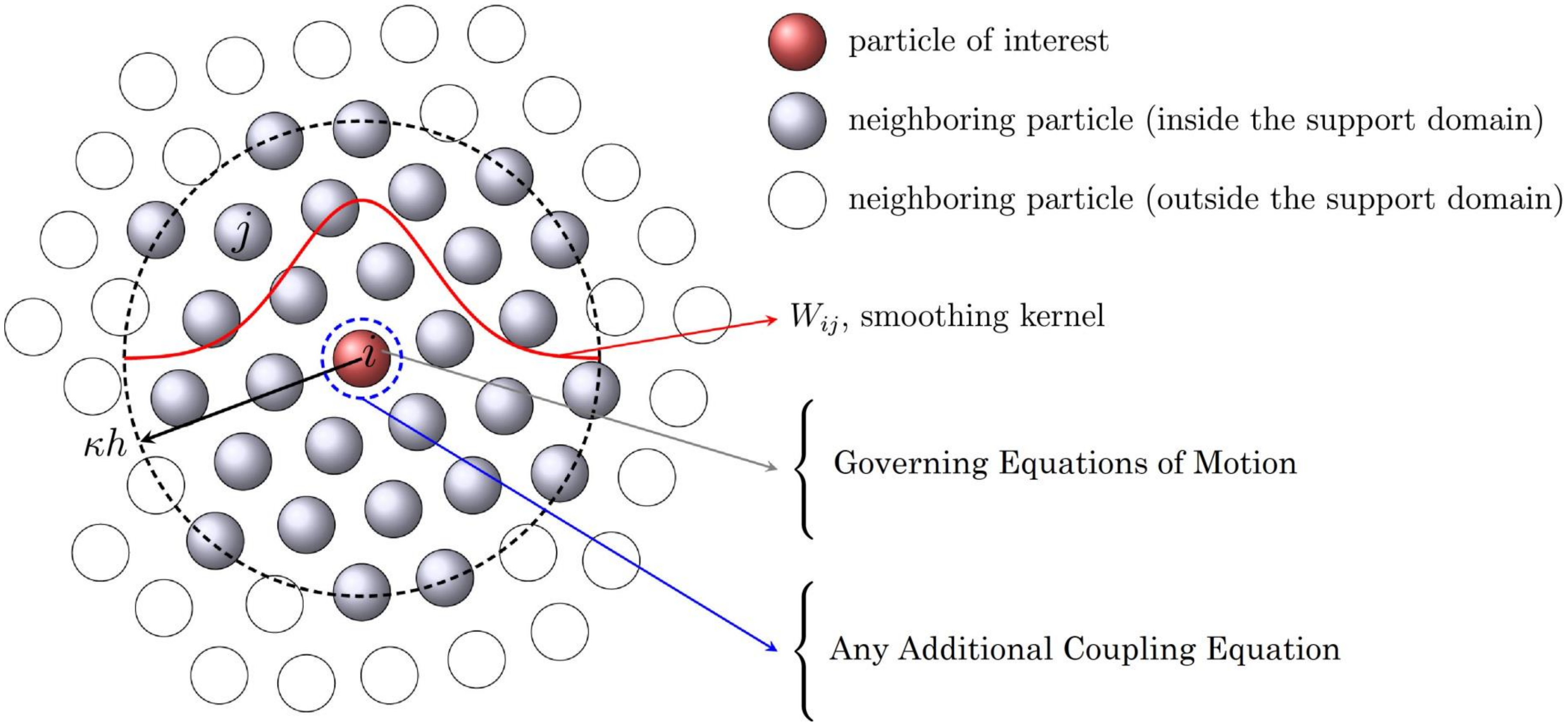




Alexiadis et al., Sci Rep, 2020, https://doi.org/10.1038/s41598-020-73329-0

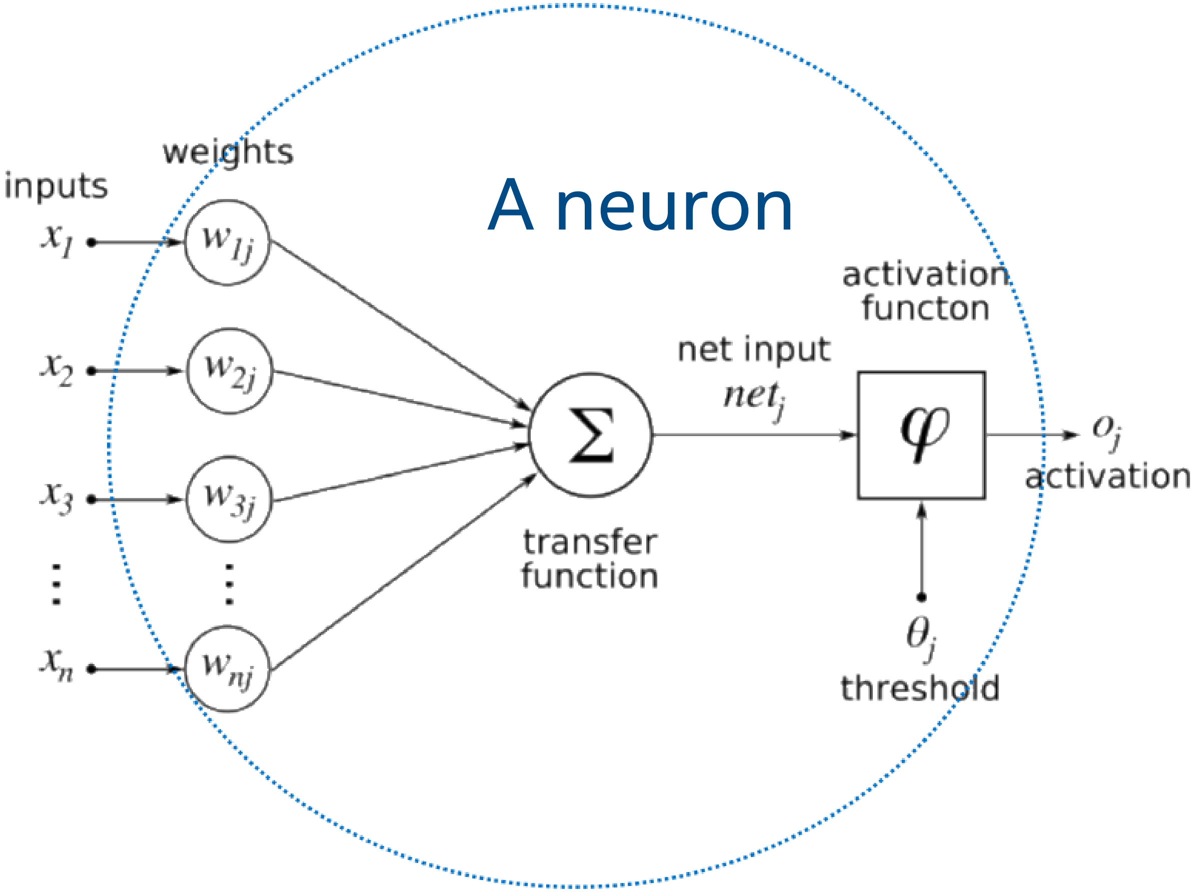
Alexiadis, Artificial Intelligence In Medicine, 2019, https://doi.org/10.1016/j.artmed.2019.06.005

Particles in the SPH methods (HPC)



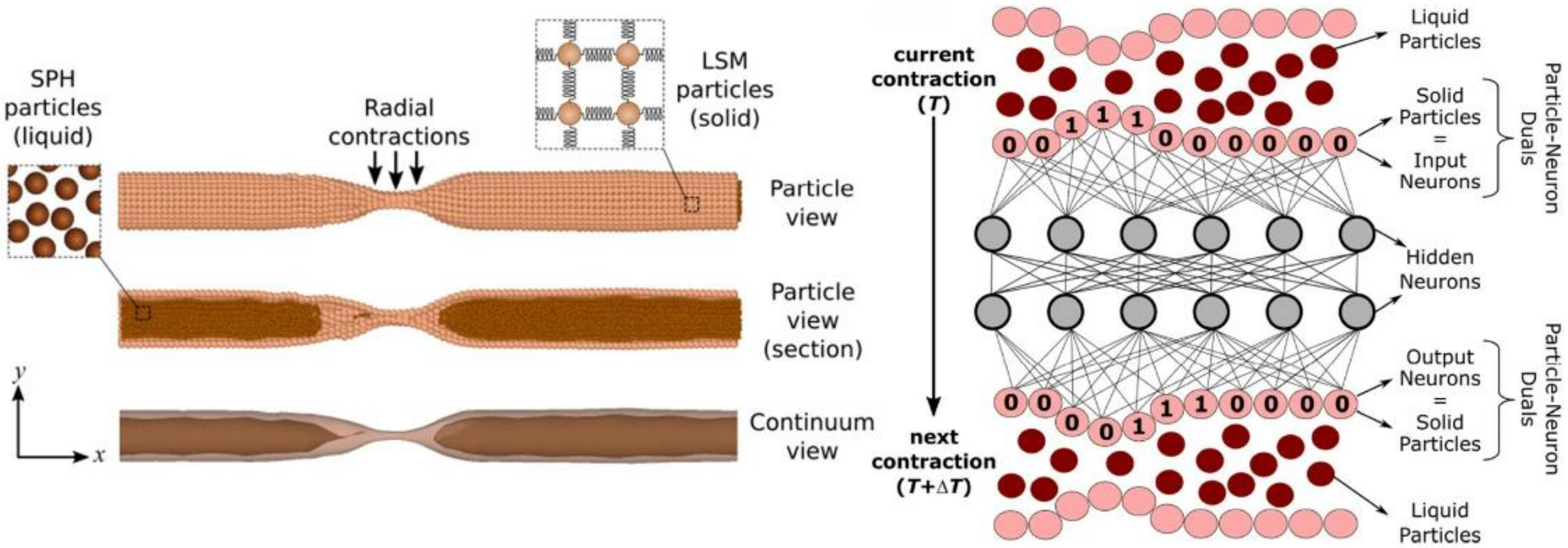
Rezavand et al., C&F, 2019, https://doi.org/10.1016/j.compfluid.2019.01.024

Artificial neurons (DL)



https://en.wikibooks.org/wiki/Artificial_Neural_Networks/Print_Version

Duality: Particles/Neurons



Alexiadis et al., Sci Rep, 2020, https://doi.org/10.1038/s41598-020-73329-0

HPC and AI Frameworks

Powered by oneAPI

- LAMMPS
  - Intel® Xeon® CPUs: https://docs.lammps.org/Build_basics.html
  - Intel X^e GPUs (Ponte Vecchio)
- Intel Python
- oneDNN
- Intel TensorFlow

Examples on Intel® XPUs

AI-driven MD (DeepDriveMD)

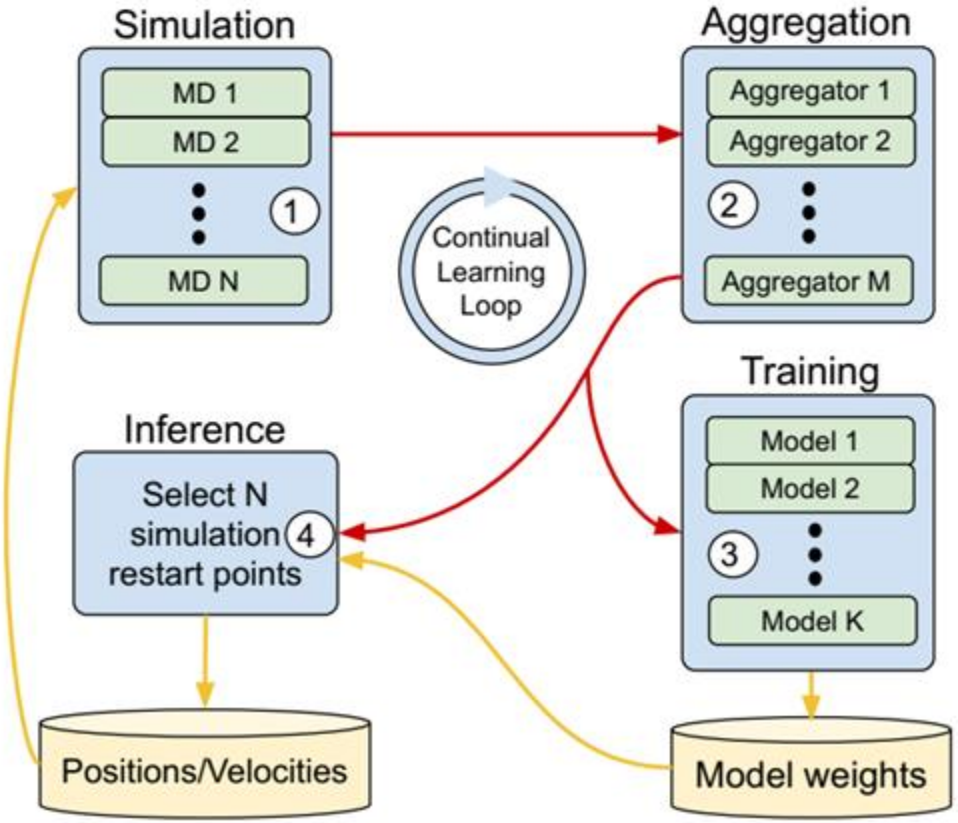
AI-driven HPC ensembles

Steer ensemble-based simulations

Selection of more scientifically meaningful computations

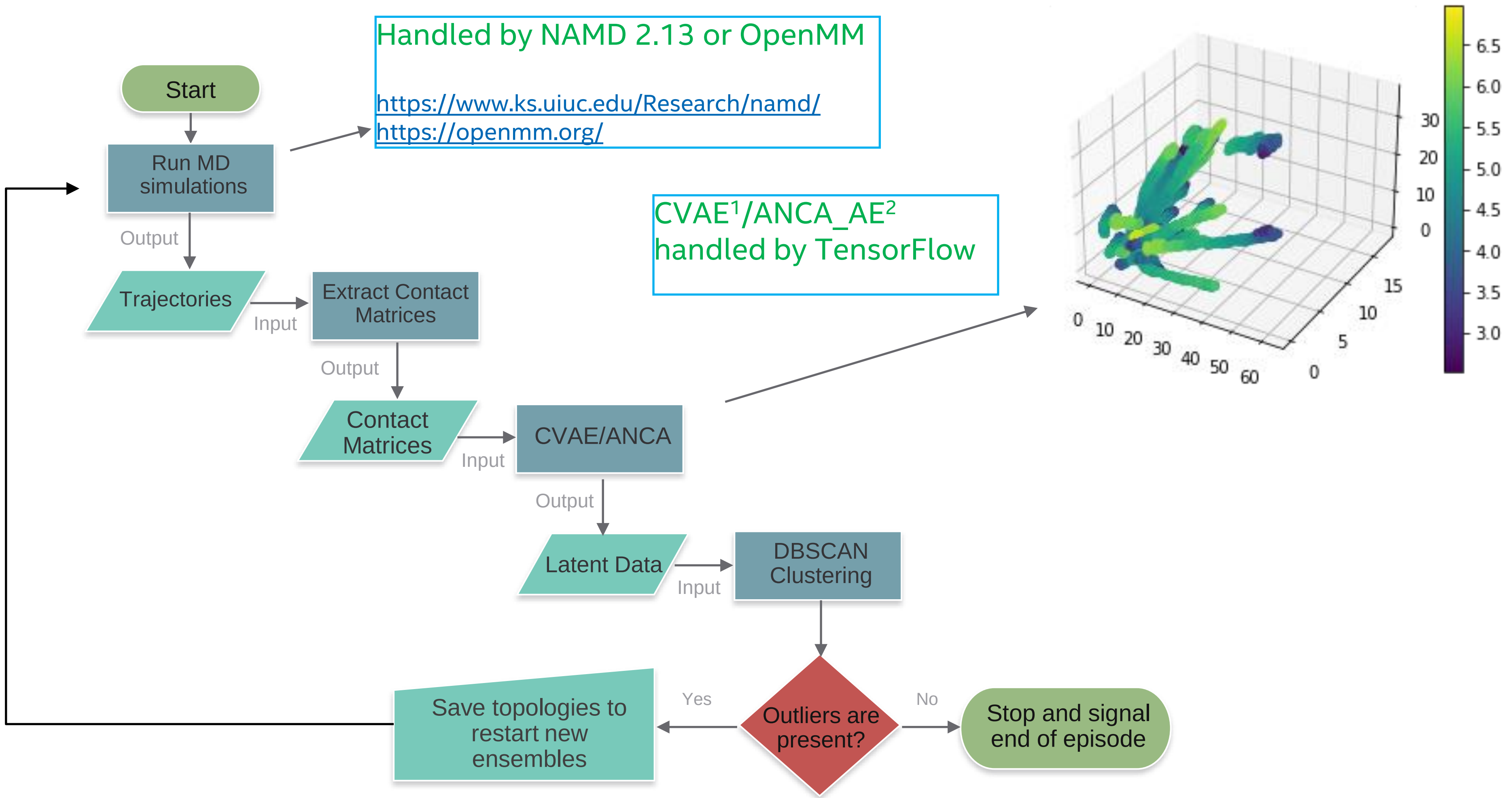
Restart the ensembles accordingly

10^2-10^3X faster MD simulations



Brace et al., arXiv, 2021, https://arxiv.org/abs/2104.04797

Implementation



Start

Run MD simulations

Output

Trajectories

Input

Extract Contact Matrices

Output

Contact Matrices

Input

CVAE/ANCA

Output

Latent Data

Input

DBSCAN Clustering

Outliers are present?

Yes

Save topologies to restart new ensembles

No

Stop and signal end of episode

Handled by NAMD 2.13 or OpenMM

https://www.ks.uiuc.edu/Research/namd/
https://openmm.org/

CVAE[1]/ANCA_AE[2] handled by TensorFlow

---

[1] Convolutional Variational Autoencoder (CVAE))
[2] Anharmonic conformational Analysis Enabled Autoencoders (ANCA-AE)

HPC and AI Frameworks

Powered by oneAPI

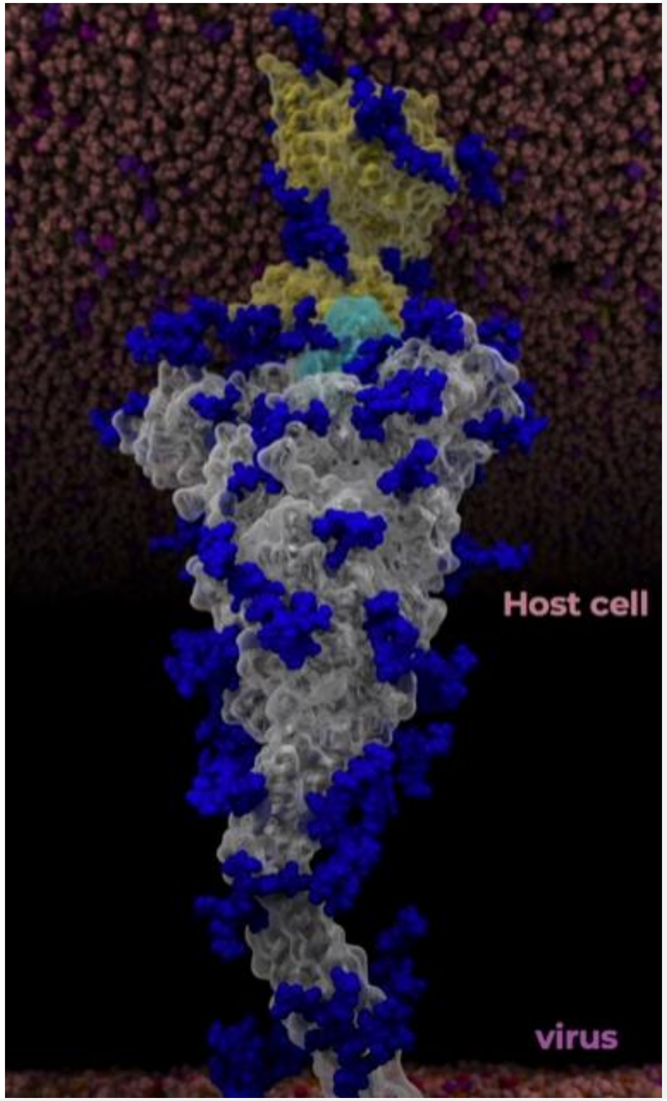
- MD ensembles:
  - NAMD 2.13: https://www.ks.uiuc.edu/Research/namd/
  - OpenMM: https://openmm.org/
- AI hierarchy:
  - oneDNN
  - Intel TensorFlow
  - Intel PyTorch
- Intel® Xeon® CPUs
- Intel X^e GPUs (Ponte Vecchio)

Examples on Intel® XPUs

Intelligent resolution

SARS-COV-2 Spike protein or RTC?

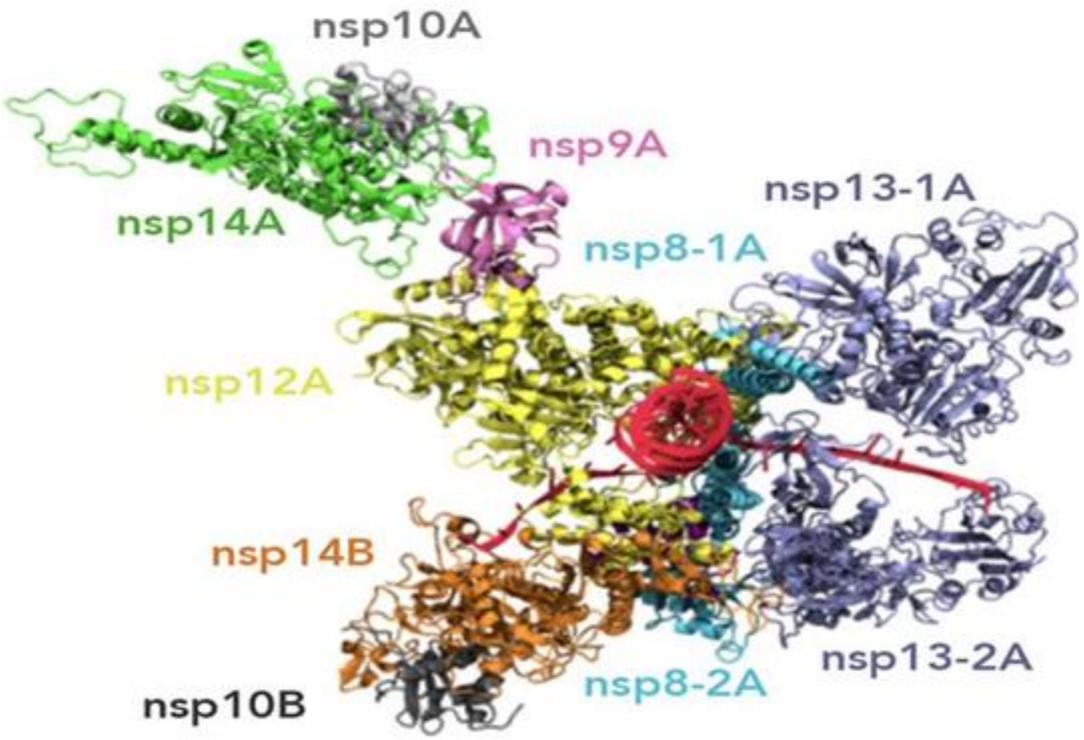
- Simulation of the Spark protein opening has shed light on COVID-19 nature
- However, the virus is still infectious
- Spike protein is subject to mutations
  - Need for continuously updated treatments
- RTC: Mutations much less likely



http://www.ks.uiuc.edu/Research/gordonbell2020/

SARS-COV-2 RTC

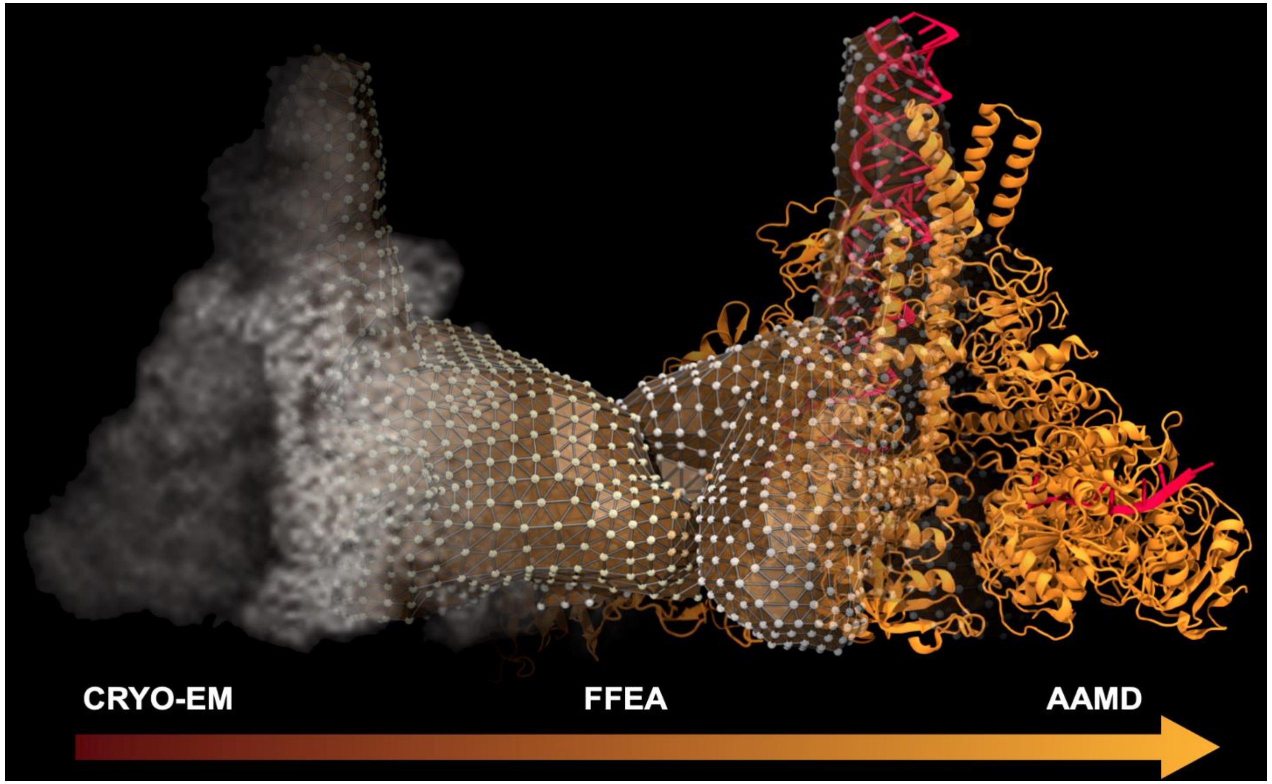
- Replication Transcription Complex (RTC): a multi-domain protein responsible for replicating and transcribing the viral RNA inside a human cell.
- RTC acts like a photocopier that assembles and replicates the viral RNA in large quantities.
- Stopping the RTC function can be a pathway to treat Covid-19.
- Viral RNA replication is complex: its dynamic nature needs to be understood



Trifan et al., bioRxiv, 2021, https://doi.org/10.1101/2021.10.09.463779

A multi-resolution problem

From Cryo-EM to FFEA (continuum) AAMD (automictic)



http://www.ks.uiuc.edu/Research/gb21intelligentresolution

30% of the collected images for RTC remain unused, due to lack of sufficient ability to fit the data into 3D conformers.

Data consists of hidden states or functionally relevant information for drug discovery.

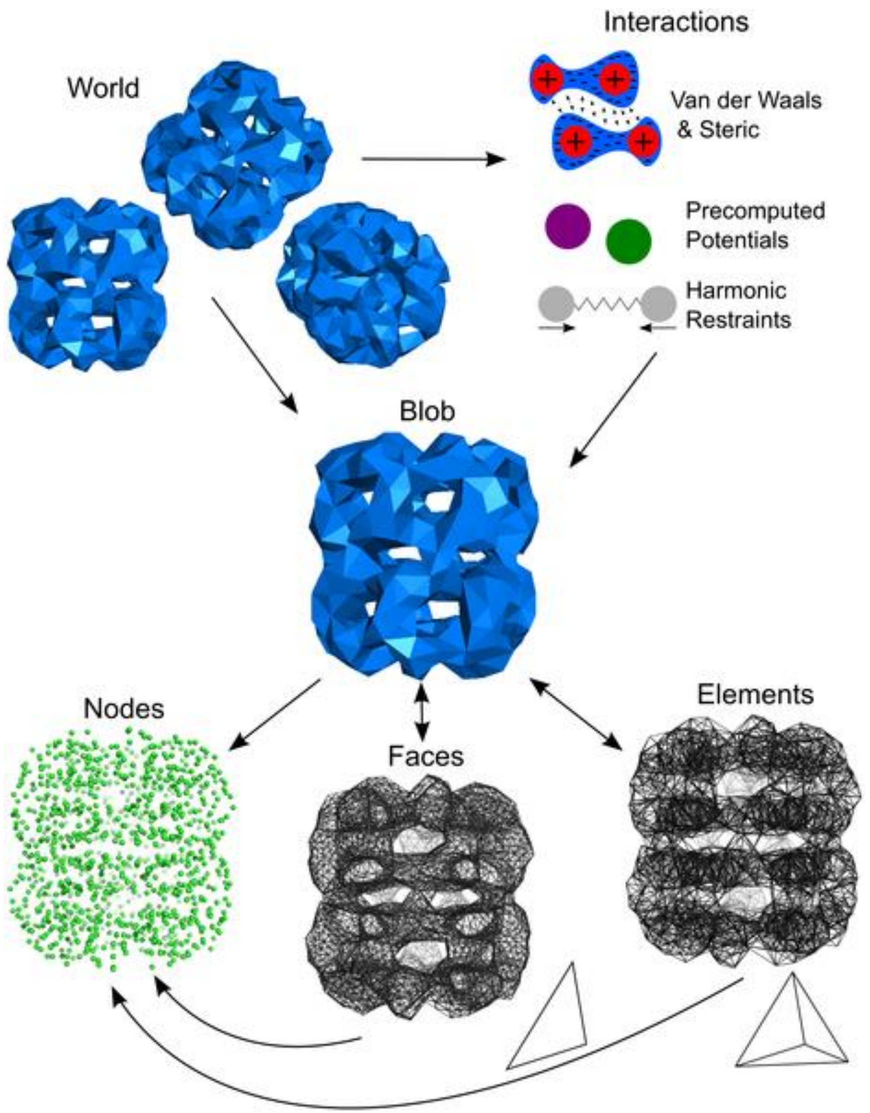
Hence, we need to move from Cryo-EM to volumetric representations (continuum FFEA) and then AAMD simulations for atomistic details.

The FFEA method

- Fluctuating-FEA treats biomolecules as continuum visco-elastic solids subject to thermal noise, to satisfy the fluctuation-dissipation theorem

$$\rho \frac{D\boldsymbol{u}}{Dt} = \nabla \cdot \left(\boldsymbol{\sigma}^{v} + \boldsymbol{\sigma}^{e} + \boldsymbol{\pi}\right) + \boldsymbol{f}$$

- Harmonic restraints for RNA
- Precomputed Potentials from AAMD
- Physics evaluated on tetrahedral elements and quantities defined on nodes

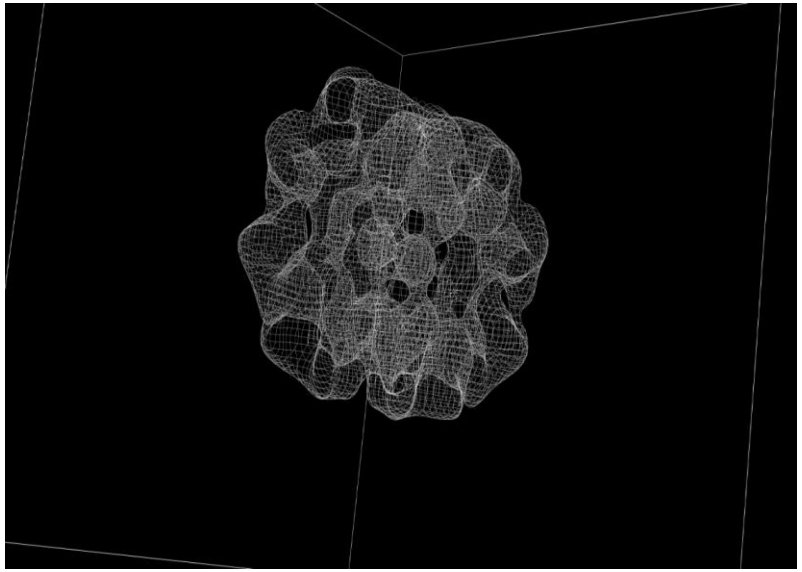


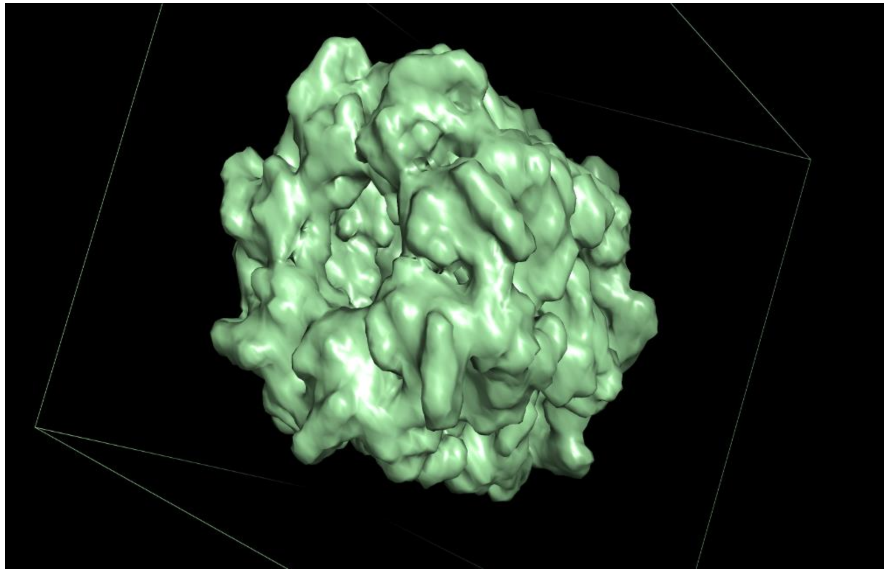
Solernou A, et al., PLOS Computational Biology, 2018, https://doi.org/10.1371/journal.pcbi.1005897

The FFEA method

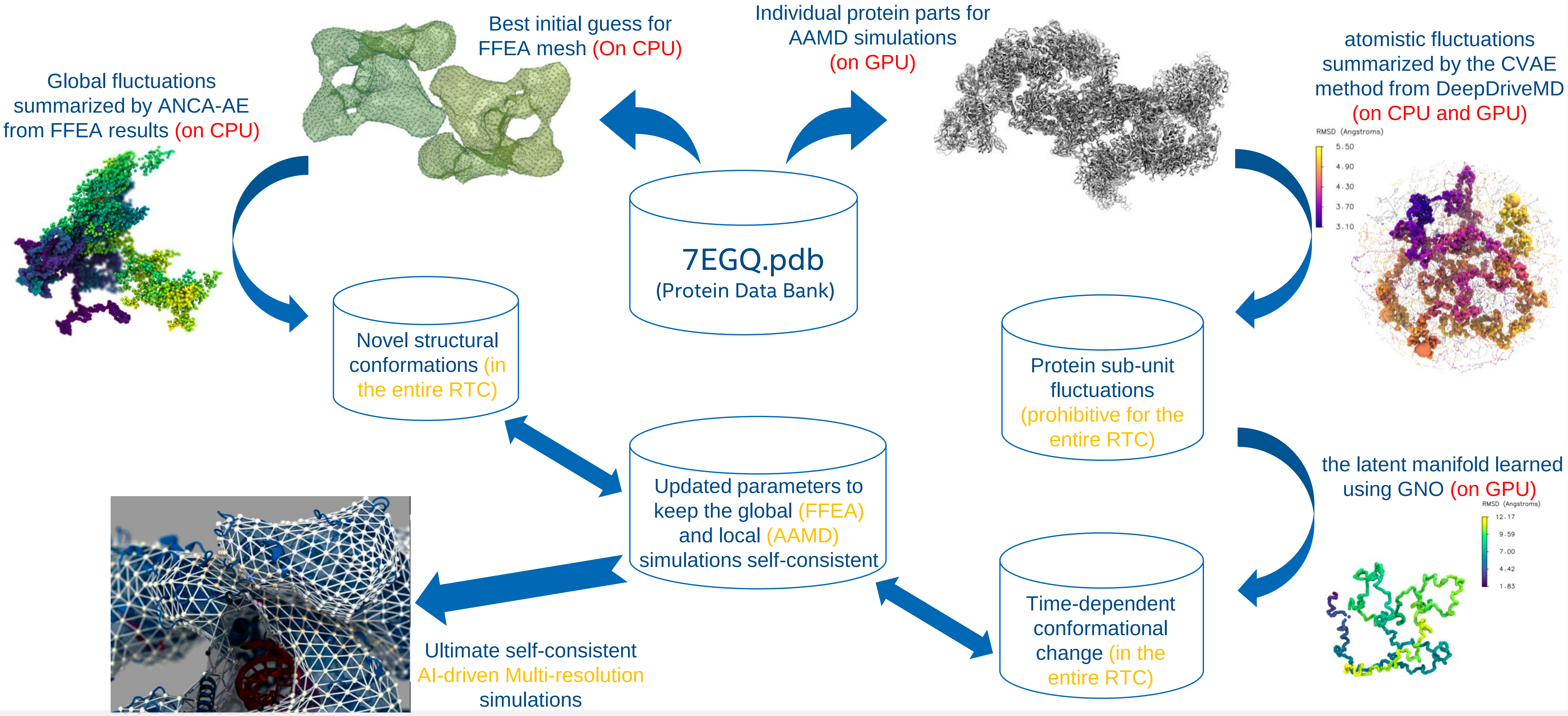
- FFEA framework (Parallelized with OpenMP)
  - https://bitbucket.org/FFEA/ffea/downloads/

- PyMOL visualization of an exemplary protein, 4hel protein:





Implementation



7EGQ.pdb
(Protein Data Bank)

Best initial guess for FFEA mesh (On CPU)

Global fluctuations summarized by ANCA-AE from FFEA results (on CPU)

Novel structural conformations (in the entire RTC)

Individual protein parts for AAMD simulations (on GPU)

atomistic fluctuations summarized by the CVAE method from DeepDriveMD (on CPU and GPU)

Protein sub-unit fluctuations (prohibitive for the entire RTC)

the latent manifold learned using GNO (on GPU)

Time-dependent conformational change (in the entire RTC)

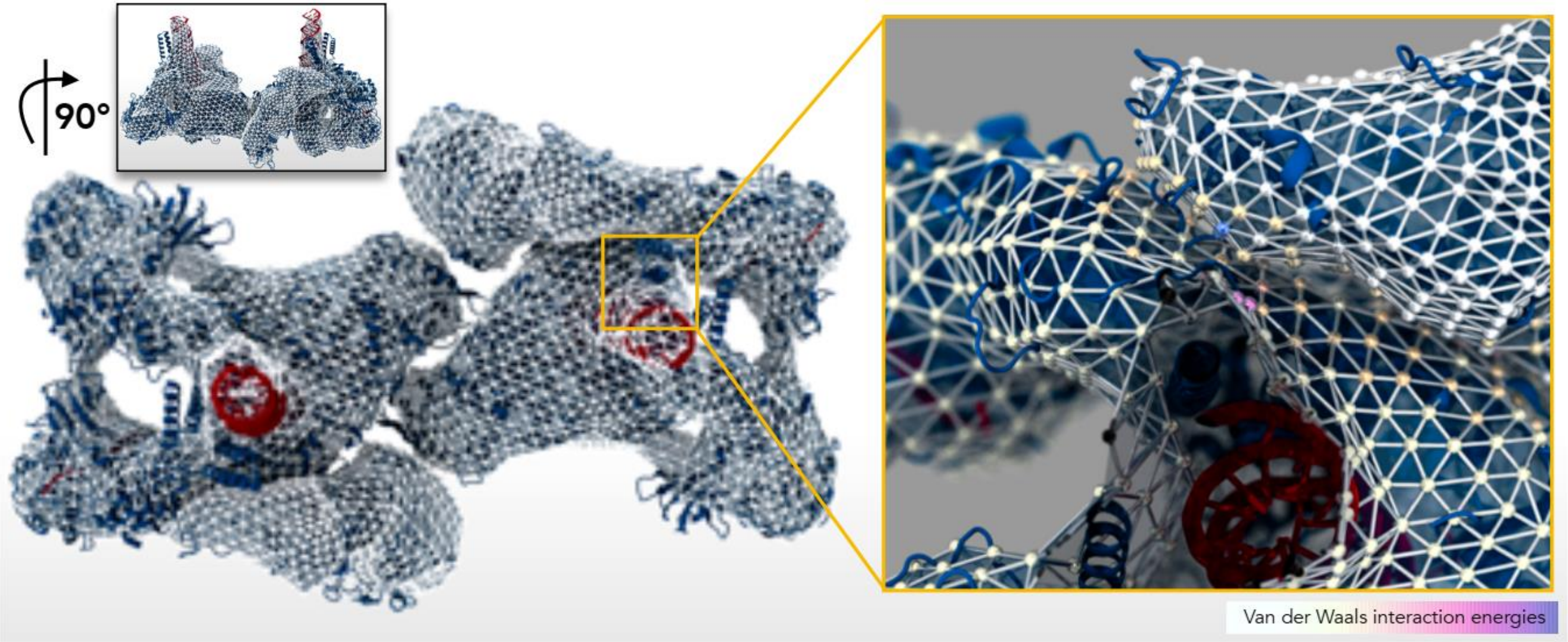
Updated parameters to keep the global (FFEA) and local (AAMD) simulations self-consistent

Ultimate self-consistent AI-driven Multi-resolution simulations

http://www.ks.uiuc.edu/Research/gb21intelligentresolution

http://www.ks.uiuc.edu/Research/gordonbell2020

Hybrid structure of FFEA superimposed with AAMD



Trifan et al., bioRxiv, 2021, https://doi.org/10.1101/2021.10.09.463779

HPC and AI Frameworks

Powered by oneAPI

- MD ensembles:
  - NAMD 2.13: https://www.ks.uiuc.edu/Research/namd/
  - OpenMM: https://openmm.org/
- AI hierarchy:
  - oneDNN
  - Intel TensorFlow
  - Intel PyTorch
- Intel® Xeon® CPUs
- Intel X^e GPUs (Ponte Vecchio)

Takeaways

- AI methods can accelerate HPC applications and effectively use the generated data.
- Various spatial and temporal resolutions are possible to couple using AI methods.
- Complex HPC problems driven by AI methodologies are easier to implement using oneAPI.
  - Open-source and cross-platform
- Essential computations and the communications between Host and Device are efficiently enabled on Intel® XPUs

intel.