SPARC - MLFF

Machine-learned force fields

User guide

Material Physics & Mechanics Group
PI: Phanish Suryanarayana
Georgia Institute of Technology

Contributors
Citation
Acknowledgements

Contributors

- **Phanish Suryanarayana** (PI)
  - **Shashikant Kumar**: Code infrastructure, SOAP, Bayesian linear regression, UQ
  - **Abhiraj Sharma**: Boundary conditions, cyclix MLFF
- **Andrew J. Medford** (co-PI)
- **John E. Pask** (co-PI)

Citation

If you publish work using/regarding MLFF in SPARC, please cite some of the following articles, particularly those that are most relevant to your work:

- https://doi.org/10.1063/5.0180541,
  https://doi.org/10.1063/5.0204229 (SQ),
  https://doi.org/10.48550/arXiv.2408.07554 (Cyclix),
  https://doi.org/10.48550/arXiv.2407.15290 (internal energy)

Acknowledgements

- U.S. Department of Energy, Office of Science: DE-SC0023445

Comments

The code will fail with the following options and the related input options are listed.

- Relaxation algorithms specified using: `RELAX_FLAG`

Input file options

MLFF

MLFF_FLAG | MLFF_RCUT_SOAP | MLFF_RADIAL_BASIS | MLFF_ANGULAR_BASIS | MLFF_INITIAL_STEPS_TRAIN | MLFF_MODEL_FOLDER | MLFF_IF_ATOM_DATA_AVAILABLE | MLFF_REGUL_MIN | MLFF_MAX_STR_STORE | MLFF_MAX_CONFIG_STORE | MLFF_FACTOR_MULTIPLY_SIGMATOL | MLFF_IF_SPARSIFY_BEFORE_TRAIN | MLFF_EXPONENT_SOAP | MLFF_SPLINE_NGRID | MLFF_SCALE_FORCE | MLFF_SCALE_STRESS MLFF_PRINT_FLAG | MLFF_DFT_FQ |

MLFF

MLFF_FLAG

Type

Integer

Unit

No unit

Default

0

Example

```
MLFF_FLAG: 1
```

Description

Flag to turn on MLFF in SPARC. There are three options currently available. (1) MLFF_FLAG: 1 to perform on-the-fly MD with no existing model, (2) MLFF_FLAG: 21 to perform only predictions from an existing model, and (3) MLFF_FLAG: 22 to perform on-the-fly MD building on top of an existing model

MLFF_RCUT_SOAP

Type

Double

Unit

Bohr

Default

10.0

Example

```
MLFF_RCUT_SOAP: 8.0
```

Description

The cutoff radius used to calculate the SOAP descriptor.

MLFF_SIGMA_ATOM_SOAP

Type

Double

Unit

Bohr

Default

1.0

Example

```
MLFF_SIGMA_ATOM_SOAP: 1.0
```

Description

The width of Gaussians in SOAP descriptor.

MLFF_RADIAL_BASIS

Type

Integer

Unit

No unit

Default

8

Example

```
MLFF_RADIAL_BASIS: 7
```

Description

The number of radial basis functions used in SOAP descriptors.

Remark

8-10 radial basis functions should be sufficient for most applications.

MLFF_ANGULAR_BASIS

Type

Integer

Unit

No unit

Default

6

Example

```
MLFF_ANGULAR_BASIS: 7
```

Description

The maximum angular momentum quantum number (L) of spherical harmonics used in the SOAP descriptor.

Remark

6-8 for SOAP works well for most applications.

MLFF_INITIAL_STEPS_TRAIN

Type

Integer

Unit

No unit

Default

10

Example

```
MLFF_INITIAL_STEPS_TRAIN:
20
```

Description

The number of DFT steps performed in the beginning of on-the-fly MLFF simulation.

MLFF_MODEL_FOLDER

Type

String

Unit

No unit

Default

No default

Example

```
MLFF_MODEL_FOLDER: ././
```

Description

The location of folder where the trained model file is located.

MLFF_IF_ATOM_DATA_AVAILABLE

Type

Integer

Unit

No unit

Default

0

Example

```
MLFF_IF_ATOM_DATA_AVAILABLE:
1
```

Description

Flag to indicate if an pretrained MLFF model is available.

MLFF_REGUL_MIN

Type

Double

Unit

No unit

Default

1e-14

Example

MLFF_REGUL_MIN: 1e-10

Description

The mimimum value of inverse of condition number of $K^T K + \lambda I$ matrix which appear in MLFF calculation.

MLFF_REGUL_MIN

Remark

The $K^T K + \lambda I$ matrix needs to be inverted during the training of MLFF. The matrix $K^T K$ is in general ill-conditioned so regularization is used to improve the conditioning. Any number in the range of 1E-10-1E-14 should work. A larger value should result in lesser DFT calls, but also lesser accuracy of MLFF model.

MLFF_MAX_STR_STORE

Type

Integer

Unit

No unit

Default

500

Example

```
MLFF_MAX_STR_STORE: 1000
```

Description

Maximum number of DFT calls that can be made during the on-the-fly simulation.

Remark

A larger number requires larger memory allocation.

MLFF_MAX_CONFIG_STORE

Type

Integer

Unit

No unit

Default

5000

Example

```
MLFF_MAX_CONFIG_STORE:
1000
```

Description

Maximum number of atomic descriptors per element type that will be used in training.

Remark

A larger number requires larger memory allocation.

MLFF_FACTOR_MULTIPLY_SIGMATOL

Type

Double

Unit

No unit

Default

1.01

Example

```
MLFF_FACTOR_MULTIPLY_SIGMATOL:
1.001
```

Description

A prefactor multiplied with maximum of Bayesian errors while updating the σ_{tol}. It has to be greater than 1.

Remark

A larger value will result lesser DFT calls hence might reduce the accuracy. Use any number in (1.0001-1.01) range.

MLFF_IF_SPARSIFY_BEFORE_TRAIN

Type

Integer

Unit

No unit

Default

1

Example

```
MLFF_IF_SPARSIFY_BEFORE_TRAIN:
1
```

Description

Flag to turn on the sparsification before each of the training steps.

MLFF_EXPONENT_SOAP

Type

Double

Unit

No unit

Default

4.0

Example

```
MLFF_EXPONENT_SOAP: 4.0
```

Description

The exponent in the polynomial kernel used in Bayesian linear regression.

Remark

In principle, any number greater than 1 could be used. The code has been tested most for the exponent of 4.

MLFF_TOL_FORCE

Type

Double

Unit

No unit

Default

5e-10

Example

MLFF_TOL_FORCE: 1e-10

Description

The initial value of σ_{tol}.

MLFF_SCALE_FORCE

Type

Double

Unit

No unit

Default

1.0

Example

```
MLFF_SCALE_FORCE: 10
```

Description

The scaling of force terms in the loss function during the training.

Remark

A larger number can result in better accuracy of MLFF forces.

MLFF_SPLINE_NGRID

Type

Integer

Unit

No unit

Default

100

Example

MLFF_SPLINE_NGRID: 100

Description

Number of grid points to use for spline interpolation of hnl.

Remark

A larger number can result in better accuracy of MLFF forces.

MLFF_SCALE_STRESS

Type

Double

Unit

No unit

Default

1.0 1.0 1.0 1.0 1.0 1.0

Example

```
MLFF_SCALE_STRESS: 10 1 1
10 1 10
```

Description

The scaling of stress tensor terms in the loss function during the training.

MLFF_PRINT_FLAG

Type

Integer

Unit

No unit

Default

1

Example

```
MLFF_PRINT_FLAG: 1
```

Description

Flag to turn on the printing of MLFF related log in a file named 'mlff.log' during runtime.

MLFF_INTERNAL_ENERGY_FLAG

Type

Integer

Unit

No unit

Default

0

Example

```
MLFF_INTERNAL_ENERGY_FLAG:
1
```

Description

Flag to turn on the training of internal energy model based on linear regression.

MLFF_PRESSURE_TRAIN_FLAG

Type

Integer

Unit

No unit

Default

0

Example

```
MLFF_PRESSURE_TRAIN_FLAG:
1
```

Description

Flag to turn on the training with only pressure and not all of the stress tensor components.

MLFF_DFT_FQ

Type

Integer

Unit

No unit

Default

100000000

Example

MLFF_DFT_FQ: 100

Description

Input variable to set the frequency at which DFT must be performed during on-the-fly MD.