Description of Physically-Informed Neural Network (PINN) potential

January 8, 2021

The *pinn* potential file is formatted as follows:

- Comment - includes tags such as DATE, UNITS, CONTRIBUTOR, and CITATION.
- Comment
- Comment
- $\langle nv \rangle$ $\langle G_0 \rangle$ $\langle nf \rangle$ - format version, reference Gi, type of activation function.
  - Format version (nv): 2 (current version) which means all Gis are transformed by inverse sine hyperbolic function after shifting by the reference value, G_0.
  - Type of activation function (nf): 1 stands for $f(u) = \frac{1}{2}\tanh(\frac{u}{2})$.
- $\langle nel \rangle$ - number of chemical species in the system.
- $\langle symbol \rangle$ $\langle mass \rangle$ - species 1.
- $\langle symbol \rangle$ $\langle mass \rangle$ - species 2.
- ...
- $\langle symbol \rangle$ $\langle mass \rangle$ - nel^{th} species.
- $\langle unused \rangle$ $\langle unused \rangle$ $\langle R_c \rangle$ $\langle d_c \rangle$ $\langle \sigma \rangle$
  - r_c : cutoff distance
  - d_c : cutoff range
  - σ : width of a gaussian
- $\langle nl \rangle$ $\langle P_0 \rangle$ $\langle P_1 \rangle$ ... $\langle P_{nl-1} \rangle$ - number of orders of the Legendre polynomials (nl) followed by their orders in ascending order, e.g. $0, 1, 2, 4, 6$ etc.
- $\langle nS \rangle$ $\{r_0^{(s=1,2,...nS)}\}$ - number of gaussian radii (nS) and list of their values.
- $\langle nNL \rangle$ $\{N_1, N_2, ...N_{nNL}\}$ - number of ANN layers and list of nodes in the input, hidden and output layers respectively.
  - There are two kinds of PINN potentials. The number of nodes in the input layer can be $nel * nel * (nel + 1) * nl * nS/2$ or $nel * (nel + 1) * nl * nS/2$ depending on Kind 1 or Kind 2.

- $\langle 0/1 \rangle$ $\{\chi_i; i = 1, 2, ..., N\}$ - flag and list of base BOP parameters of given chemical species in the system spanned over several lines.
  - Flag: if 0, base BO parameters not used; if 1, use base BO parameters.
  - N: number of nodes in the output layer equals $(4 * nel * nel + nel + 3 * nel * nel * nel)$.
  - As an example, the following is the order of base BOP parameters of a binary system:

$$\begin{pmatrix}
\chi[1] = A_{t_1}^{t_1} & \chi[2] = A_{t_1}^{t_2} & \chi[3] = \alpha_{t_1}^{t_1} & \chi[4] = \alpha_{t_1}^{t_2} \\
\chi[5] = B_{t_1}^{t_1} & \chi[6] = B_{t_1}^{t_2} & \chi[7] = \beta_{t_1}^{t_1} & \chi[8] = \beta_{t_1}^{t_2} \\
\chi[9] = h_{t_1}^{t_1 t_1} & \chi[10] = h_{t_1}^{t_1 t_2} & \chi[11] = h_{t_1}^{t_2 t_1} & \chi[12] = h_{t_1}^{t_2 t_2} \\
\chi[13] = \sigma_{t_1} & & & \\
\chi[14] = a_{t_1}^{t_1 t_1} & \chi[15] = a_{t_1}^{t_1 t_2} & \chi[16] = a_{t_1}^{t_2 t_1} & \chi[17] = a_{t_1}^{t_2 t_2} \\
\chi[18] = \lambda_{t_1}^{t_1 t_1} & \chi[19] = \lambda_{t_1}^{t_1 t_2} & \chi[20] = \lambda_{t_1}^{t_2 t_1} & \chi[21] = \lambda_{t_1}^{t_2 t_2} \\
- & & & \\
\chi[22] = A_{t_2}^{t_1} & \chi[23] = A_{t_2}^{t_2} & \chi[24] = \alpha_{t_2}^{t_1} & \chi[25] = \alpha_{t_2}^{t_2} \\
\chi[26] = B_{t_2}^{t_1} & \chi[27] = B_{t_2}^{t_2} & \chi[28] = \beta_{t_2}^{t_1} & \chi[29] = \beta_{t_2}^{t_2} \\
\chi[30] = h_{t_2}^{t_1 t_1} & \chi[31] = h_{t_2}^{t_1 t_2} & \chi[32] = h_{t_2}^{t_2 t_1} & \chi[33] = h_{t_2}^{t_2 t_2} \\
\chi[34] = \sigma_{t_2} & & & \\
\chi[35] = a_{t_2}^{t_1 t_1} & \chi[36] = a_{t_2}^{t_1 t_2} & \chi[37] = a_{t_2}^{t_2 t_1} & \chi[38] = a_{t_2}^{t_2 t_2} \\
\chi[39] = \lambda_{t_2}^{t_1 t_1} & \chi[40] = \lambda_{t_2}^{t_1 t_2} & \chi[41] = \lambda_{t_2}^{t_2 t_1} & \chi[42] = \lambda_{t_2}^{t_2 t_2}
\end{pmatrix} \quad (1)$$

    Subscripts, $t_i (i = 1, 2, ..., nel)$, represents chemical species.

- $\{\{\mathbf{W}^i\}\{\mathbf{b}^i\}; i = 1, 2, .., nN\}$ - list of all NN weights and biases spanned over several lines; nN is the number of NN layers excluding the input layer. Elements of the first layer of the matrix, $\mathbf{W}^i$, are listed first.

Notes on feature vectors (Gis):

- Definition of structural parameters is described in "G. P. Purja Pun, V. Yamakov , J. Hickman , E. H. Glaessgen and Y. Mishin, Phys Rev Mat 4, 113807 (2020)".

- An input vector of structural parameters (fingerprints) follows a strict order. For a single component system with the Legendre polynomial orders, $l = 0, 1, 2, 4, 6$, and Gaussian radii, $r_0^{(s)} (s = 1, 2, 3, 4, 5, 6)$, the order is $G_i^{(0,1)}$ $G_i^{(0,2)}$ $G_i^{(0,3)}$ $G_i^{(0,4)}$ $G_i^{(0,5)}$ $G_i^{(0,6)}$ $G_i^{(1,1)}$ $G_i^{(1,2)}$ $G_i^{(1,3)}$ $G_i^{(1,4)}$ $G_i^{(1,5)}$ $G_i^{(1,6)}$ ... $G_i^{(6,1)}$ $G_i^{(6,2)}$ $G_i^{(6,3)}$ $G_i^{(6,4)}$ $G_i^{(6,5)}$ $G_i^{(6,6)}$.s

- Feature vector in multicomponent system can be spanned based on chemical types of center atom, i, and its neighbors, j and k. For example, a binary system with chemical types A and B has a feature vector defined as

$$\begin{array}{rl}
atom\ 1 \rightarrow & \boxed{A_iA_jA_k}\boxed{A_iA_jB_k}\boxed{A_iB_jB_k}\boxed{(0)}\boxed{(0)}\boxed{(0)} \\
atom\ 2 \rightarrow & \boxed{(0)}\boxed{(0)}\boxed{(0)}\boxed{B_iA_jA_k}\boxed{B_iA_jB_k}\boxed{B_iB_jB_k} \\
\ldots & \\
atom\ N \rightarrow & \boxed{(0)}\boxed{(0)}\boxed{(0)}\boxed{B_iA_jA_k}\boxed{B_iA_jB_k}\boxed{B_iB_jB_k}
\end{array}$$

  This description is based on PINN Kind 1. Each atom requires $n \times n \times (n+1)/2$ blocks of Gis, where n is the number of chemical species. For the binary system n equals 2. There are $n_s \times n_l$ elements

in each block of Gis where n_s and n_l represent numbers of gaussian positions, $r_0^{(s)}$, and Legendre polynomial orders, l, respectively.

For the same system, the Gis in PINN Kind 2 look like

$$
\begin{array}{rl}
atom\ 1 \rightarrow & \boxed{A_iA_jA_k}\boxed{A_iA_jB_k}\boxed{A_iB_jB_k} \\
atom\ 2 \rightarrow & \boxed{B_iA_jA_k}\boxed{B_iA_jB_k}\boxed{B_iB_jB_k} \\
\cdots & \\
atom\ N \rightarrow & \boxed{B_iA_jA_k}\boxed{B_iA_jB_k}\boxed{B_iB_jB_k}
\end{array}
$$