Supplemental Materials for "Many-Body Descriptors for Predicting Molecular Properties with Machine Learning: Analysis of Pairwise and Three-Body Interactions in Molecules"

Wiktor Pronobis,[†] Alexandre Tkatchenko,[*,‡] and Klaus-Robert Müller[*,†,¶,§]

†*Machine Learning Group, Technische Universität Berlin, 10587 Berlin, Germany*

‡*Physics and Materials Science Research Unit, University of Luxembourg, Luxembourg L-1511, Luxembourg*

¶*Max Planck Institute for Informatics, 66123 Saarbrücken, Germany*

§*Department of Brain and Cognitive Engineering, Korea University, Seoul 136-713, South Korea*

E-mail: alexandre.tkatchenko@uni.lu; klaus-robert.mueller@tu-berlin.de

Invariant Many-Body Interaction Descriptors

The pseudo code for calculating the two-body and three-body descriptors is provided in Alg. 1 and Alg. 2, respectively.

Algorithm 1 Pairwise Descriptors

```
procedure F_2B(molecule)
    M ← 15
    f(·) ← 0
    for i, j ← atoms of molecule do
        r_ij ← ‖r_i − r_j‖    ▷ each atom is represented by position r and atomic number Z
        Z ← sorted tuple (Z_i, Z_j)
        for m ← 1, M do
            f(Z, m) += r_ij^(−m)    ▷ feature entries are indexed by (Z, m)
```

Algorithm 2 Three-Body Descriptors

```
procedure F_3B(molecule)
    G(3, 6) ← sequences without repetition of 3 elements from the set {1, 2, ..., 6}
    B(·) ← 1.1 · L(·)    ▷ the bond length function L is defined in the main article
    f(·) ← 0
    for i, j, k ← atoms of molecule do
                    ▷ each atom is represented by position r and atomic number Z
        r_ij ← ‖r_i − r_j‖
        r_ik ← ‖r_i − r_k‖
        r_jk ← ‖r_j − r_k‖
        hasAngle ← False
        if r_ij < B(Z_i, Z_j) and r_ik < B(Z_i, Z_k) then
            hasAngle ← True
        else if r_ik < B(Z_i, Z_k) and r_jk < B(Z_j, Z_k) then
            hasAngle ← True
        else if r_ij < B(Z_i, Z_j) and r_jk < B(Z_j, Z_k) then
            hasAngle ← True
        if not hasAngle then
            continue
        Z ← sorted tuple (Z_i, Z_j, Z_k)
        for m_1, m_2, m_3 ← G(3, 6) do
            f(Z, m_1, m_2, m_3) += r_ij^(−m_1) · r_ik^(−m_2) · r_jk^(−m_3)
                    ▷ feature entries are indexed by (Z, m_1, m_2, m_3)
```