

Some notes on extending the SOAP kernel to multi-component systems

Jonathan Vandermause

(Dated: September 25, 2018)

I. REVIEW OF THE SINGLE-COMPONENT SOAP KERNEL

In the SOAP scheme, an atomic environment is represented by an atomic neighbor density function $\rho(\vec{r})$, where

$$\rho(\vec{r}) = \sum_i \exp(-\alpha|\vec{r} - \vec{r}_i|^2), \quad (1)$$

α is a hyperparameter, and \vec{r}_i is the position of atom i in the environment of the central atom. The single-component SOAP kernel between two atomic neighbor densities ρ and ρ' is defined by Eq. (36) of [1] as:

$$K(\rho, \rho') = \left(\frac{k(\rho, \rho')}{\sqrt{k(\rho, \rho)k(\rho', \rho')}} \right)^\xi, \quad (2)$$

where $k(\rho, \rho')$ is a rotationally invariant symmetry kernel given by

$$k(\rho, \rho') = \int d\hat{R} \left| \int \rho(\vec{r}) \rho'(\hat{R}\vec{r}) d\vec{r} \right|^n, \quad (3)$$

and \hat{R} is a 3-D rotation. The outer integral is over all rotations and ensures that the kernel is rotationally symmetric.

In Section IV(B) of [1], it is shown that $k(\rho, \rho')$ is equal to the dot product of the “rotational power spectra” of densities ρ and ρ' , which is efficient to compute. I will review this derivation here, since a similar argument can be applied to multi-component systems.

Letting $g_n(r)$ be an orthonormal set of radial basis functions, so that

$$\int dr r^2 g_n(r) g_{n'}(r) = \delta_{nn'}, \quad (4)$$

an atomic neighbor density $\rho(\vec{r})$ may be expanded into spherical harmonics as

$$\rho(\vec{r}) = \sum_{nlm} c_{nlm} g_n(r) Y_{lm}(\vec{r}). \quad (5)$$

We first compute that

$$\begin{aligned} S(\rho, \hat{R}\rho') &\equiv \int d\vec{r} \rho(\vec{r}) \rho'(\hat{R}\vec{r}) \\ &= \int d\vec{r} \left(\sum_{nlm} c_{nlm}^* g_n(r) Y_{lm}^*(\hat{r}) \right) \left(\sum_{n'l'm'} c'_{n'l'm'} g_{n'}(r) Y_{l'm'}(\hat{R}\vec{r}) \right) \\ &= \sum_{nlmn'l'm'} c_{nlm}^* c'_{n'l'm'} \int dr r^2 g_n(r) g_{n'}(r) \int d\hat{r} Y_{lm}^*(\hat{r}) Y_{l'm'}(\hat{R}\vec{r}) \\ &= \sum_{nlml'm'} c_{nlm}^* c'_{nl'm'} \int d\hat{r} Y_{lm}^*(\hat{r}) Y_{l'm'}(\hat{R}\vec{r}). \end{aligned} \quad (6)$$

Note that

$$\begin{aligned}
Y_{l'm'}(\vec{R}\hat{r}) &= \langle \vec{r} | \hat{R}^\dagger | Y_{l'm'} \rangle \\
&= \sum_{m''} \langle \vec{r} | Y_{l'm''} \rangle \langle Y_{l'm''} | \hat{R}^\dagger | Y_{l'm'} \rangle \\
&= \sum_{m''} \langle \vec{r} | Y_{l'm''} \rangle \langle Y_{l'm'} | \hat{R} | Y_{l'm''} \rangle^* \\
&= \sum_{m''} Y_{l'm''}(\hat{r}) D^*(\hat{R})_{m'm''}^{l'},
\end{aligned} \tag{7}$$

where the Wigner function $D(\hat{R})_{mm'}^l$ is defined as

$$D(\hat{R})_{mm'}^l = \langle Y_{lm} | \hat{R} | Y_{lm'} \rangle. \tag{8}$$

Plugging into Eq. (6) gives

$$\begin{aligned}
S(\rho, \hat{R}\rho') &= \sum_{nlml'm'm''} c_{nlm}^* c'_{nl'm'} D^*(\hat{R})_{m'm''}^{l'} \int d\hat{r} Y_{lm}^*(\hat{r}) Y_{l'm''}(\hat{r}) \\
&= \sum_{nlmm'} c_{nlm}^* c'_{nlm'} D^*(\hat{R})_{m'm}^l.
\end{aligned} \tag{9}$$

The symmetry kernel $k(\rho, \rho')$ with $n = 2$ then takes the form

$$\begin{aligned}
k(\rho, \rho') &= \int d\hat{R} S(\rho, \hat{R}\rho') S^*(\rho, \hat{R}\rho') \\
&= \sum_{nlmm'n'\lambda\mu\mu'} c_{nlm}^* c'_{nlm'} c_{n'\lambda\mu} (c'_{n'\lambda\mu'})^* \int d\hat{R} D^*(\hat{R})_{m'm}^l D(\hat{R})_{\mu'\mu}^{\lambda} \\
&= \sum_{nlmm'n'} \frac{8\pi^2}{2l+1} c_{nlm}^* c'_{nlm'} c_{n'lm} (c'_{n'lm'})^* \\
&= \sum_{nn'l} \frac{8\pi^2}{2l+1} \left(\sum_m c_{nlm}^* c_{n'lm} \right) \left(\sum_{m'} c'_{nlm'} (c'_{n'lm'})^* \right) \\
&= \sum_{nn'l} \frac{8\pi^2}{2l+1} p_{n'nl} p'_{nn'l},
\end{aligned} \tag{10}$$

where $p_{nn'l} \equiv \sum_m c_{nlm} c_{n'lm}^*$ is the rotational power spectrum. Because $p_{n'nl} = p_{nn'l}^*$, $k(\rho, \rho')$ may be expressed as the dot product of the power spectra \vec{p} and \vec{p}' :

$$k(\rho, \rho') = \vec{p} \cdot \vec{p}'. \tag{11}$$

The full SOAP kernel $K(\rho, \rho')$ is the dot product of normalized power spectra raised to a positive power:

$$K(\rho, \rho') = \left(\frac{\vec{p}}{|\vec{p}|} \cdot \frac{\vec{p}'}{|\vec{p}'|} \right)^\xi. \tag{12}$$

II. MULTI-COMPONENT OVERLAP OF POSITIONS (MOP)

Consider a two-component system. Let ρ and ρ' denote two atomic environments, with ρ_A, ρ_B the atomic neighbor densities of atom types A and B in environment ρ , and ρ'_A, ρ'_B the densities of A and B in environment ρ' . We define the MOP kernel between ρ and ρ' by direct analogy to the SOAP kernel:

$$K_{\text{MOP}}(\rho, \rho') = \left(\frac{k_{\text{MOP}}(\rho, \rho')}{\sqrt{k_{\text{MOP}}(\rho, \rho)k_{\text{MOP}}(\rho', \rho')}} \right)^\xi, \quad (13)$$

where $k_{\text{MOP}}(\rho, \rho')$ is a rotationally invariant symmetry kernel given by

$$\begin{aligned} k_{\text{MOP}}(\rho, \rho') &= \int d\hat{R} \left| \int \left(\rho_A(\vec{r})\rho'_A(\hat{R}\vec{r}) + \rho_B(\vec{r})\rho'_B(\hat{R}\vec{r}) \right) d\vec{r} \right|^2 \\ &= \int d\hat{R} \left| \int \rho_A(\vec{r})\rho'_A(\hat{R}\vec{r}) d\vec{r} \right|^2 + \int d\hat{R} \left| \int \rho_B(\vec{r})\rho'_B(\hat{R}\vec{r}) d\vec{r} \right|^2 \\ &\quad + \int d\hat{R} \left(\int d\vec{r} \rho_A(\vec{r})\rho'_A(\hat{R}\vec{r}) \right) \left(\int d\vec{r} \rho_B(\vec{r})\rho'_B(\hat{R}\vec{r}) \right)^* \\ &\quad + \int d\hat{R} \left(\int d\vec{r} \rho_A(\vec{r})\rho'_A(\hat{R}\vec{r}) \right)^* \left(\int d\vec{r} \rho_B(\vec{r})\rho'_B(\hat{R}\vec{r}) \right) \\ &= k(\rho_A, \rho'_A) + k(\rho_B, \rho'_B) + k_{\text{cross}}(\rho, \rho') + k_{\text{cross}}^*(\rho, \rho'), \end{aligned} \quad (14)$$

which is the sum of the SOAP kernels between the two density types plus a cross term k_{cross} , which we now analyze in detail.

From the derivation in part 1, we have that

$$\int d\vec{r} \rho_A(\vec{r})\rho'_A(\hat{R}\vec{r}) = \sum_{nlmm'} (c_{nlm}^A)^* c_{nlm}'^A D^*(\hat{R})_{m'm}^l, \quad (15)$$

$$\int d\vec{r} \rho_B(\vec{r})\rho'_B(\hat{R}\vec{r}) = \sum_{n'\lambda\mu\mu'} (c_{n'\lambda\mu}^B)^* c_{n'\lambda\mu}'^B D^*(\hat{R})_{\mu'\mu}^l, \quad (16)$$

and thus

$$\begin{aligned} k_{\text{cross}}(\rho, \rho') &= \int d\hat{R} S(\rho_A, \hat{R}\rho'_A) S^*(\rho_B, \hat{R}\rho'_B) \\ &= \sum_{nlmm'n'\lambda\mu\mu'} (c_{nlm}^A)^* c_{nlm}'^A c_{n'\lambda\mu}^B (c_{n'\lambda\mu}')^* \int d\hat{R} D^*(\hat{R})_{m'm}^l D(\hat{R})_{\mu'\mu}^\lambda \\ &= \sum_{nlmm'n'} \frac{8\pi^2}{2l+1} (c_{nlm}^A)^* c_{nlm}'^A c_{n'lm}^B (c_{n'lm}')^* \\ &= \sum_{nn'l} \frac{8\pi^2}{2l+1} \left(\sum_m (c_{nlm}^A)^* c_{n'lm}^B \right) \left(\sum_m c_{nlm}'^A (c_{n'lm}')^* \right) \\ &= \vec{p}_{AB} \cdot \vec{p}_{AB}', \end{aligned} \quad (17)$$

which is the dot product of the cross power spectra of the two environments. The symmetric MOP kernel $k_{\text{MOP}}(\rho, \rho')$ before normalization therefore takes the form

$$k_{\text{MOP}}(\rho, \rho') = \vec{p}_A \cdot \vec{p}_A + \vec{p}_B \cdot \vec{p}_B + 2\vec{p}_{AB} \cdot \vec{p}_{AB}. \quad (18)$$

Notice that if one of the densities vanishes, e.g. $\vec{p}_A = 0$, the MOP kernel reduces to the SOAP kernel. MOP may therefore be viewed as a generalization of SOAP that is suitable for multicomponent systems.

[1] A. P. Bartók, R. Kondor, and G. Csányi, Physical Review B **87**, 184115 (2013).