Multi-component kernel notes

Jonathan Vandermause

(Dated: October 3, 2018)

I. A MULTI-COMPONENT N-BODY KERNEL

We define the N-body kernel

$$k_N(\rho, \rho') = \sum_{\substack{c \in C(\rho, N-1) \\ p \in P(\rho', N-1)}} \sigma^2 \kappa_{c,p}^{\rho,\rho'} \exp(-||\vec{q}_c - \vec{q}_p||^2/2\ell^2) \tag{1}$$

where $C(\rho, N-1)$ and $P(\rho, N-1)$ are all the combinations and permutations, respectively, of $N-1$ atoms from environment ρ. The kernel is N-body because it includes the central atom, accounting for N atoms in total. Summing over all combinations and permutations ensures that the kernel is permutationally invariant. $\kappa_{c,p}^{\rho,\rho'}$ is defined as

$$\kappa_{c,p}^{\rho,\rho'} = \kappa_{\rho\rho'} \prod_{n=1}^{N} \kappa_{c_n p_n}, \tag{2}$$

where κ_{ab} is the ICM coefficient assigned to atom types a and b. $\kappa_{\rho\rho'}$ denotes the comparison of the central atoms of environments ρ and ρ'.

The vectors $\vec{q}_c$ and $\vec{q}_p$ contain all the interatomic distances of the $N-1$ atoms in c and p, respectively. The vector norm appearing in the kernel is

$$||\vec{q}_c - \vec{q}_p||^2 = \sum_{q=1}^{N-1}(r_{cq} - r_{pq})^2 + \sum_{q=1}^{N-2}\sum_{s=q+1}^{N-1}(r_{cq,cs} - r_{pq,ps})^2. \tag{3}$$

Notice that the norm is split into two terms. The first compares bonds containing the central atom, and the second compares bonds involving environment atoms.

II. SOAP IS A THREE-BODY KERNEL

Consider the multi-component SOAP kernel defined in Ref. [1],

$$k(\rho_1, \rho_2) = \int d\hat{R} \left(\int d\vec{r} \sum_{\alpha\alpha'} \kappa_{\alpha\alpha'} \rho_1^\alpha(\vec{r}) \rho_2^{\alpha'}(\hat{R}\vec{r}) \right)^2. \tag{4}$$

Recalling that

$$\rho_1^\alpha(\vec{r}) = \sum_{i \in \rho_1^\alpha} \exp\left(-\frac{(\vec{x}_i - \vec{r})^2}{2\sigma^2}\right) \tag{5}$$

and performing the inner integral gives

$$\int d\vec{r} \sum_{\alpha\alpha'} \kappa_{\alpha\alpha'} \rho_1^\alpha(\vec{r}) \rho_2^{\alpha'}(\hat{R}\vec{r}) = \sum_{\alpha\alpha'} \sum_{\substack{i \in \rho_1^\alpha \\ j \in \rho_2^{\alpha'}}} \kappa_{\alpha\alpha'} \pi^{3/2} \sigma^3 \exp\left(-\frac{(\vec{x}_i - \hat{R}\vec{x}_j)}{4\sigma^2}\right). \tag{6}$$

The full kernel may therefore be written as

$$k(\rho_1, \rho_2) = \pi^3 \sigma^6 \sum_{\substack{\alpha\alpha' \\ \beta\beta'}} \kappa_{\alpha\alpha'} \kappa_{\beta\beta'} \sum_{\substack{i \in \rho_1^\alpha \\ j \in \rho_2^{\alpha'}}} \sum_{\substack{k \in \rho_1^\beta \\ l \in \rho_2^{\beta'}}} \int d\hat{R} \exp\left(-\frac{(\vec{x}_i - \hat{R}\vec{x}_j)^2}{4\sigma^2}\right) \exp\left(-\frac{(\vec{x}_k - \hat{R}\vec{x}_l)^2}{4\sigma^2}\right).$$

$$(7)$$

This kernel captures angular information between at most two atoms in the environment. It is therefore not expected to be any more accurate than Glielmo's three-body kernel [2], which directly compares pairs and triplets of distances without performing an integral over SO(3). Normalizing the SOAP kernel and raising to a positive power may allow the model to capture many-body interactions, but the same procedure can be applied to the two- and three-body kernels of Ref. [2], which are much easier to compute.

Note that the multi-component SOAP kernel handles different species by assigning ICM coeffients to pairs of atoms rather than pairs of triplets. This significantly reduces the number of ICM hyperparameters needed.

III.   COMPUTING THE POWER SPECTRUM

We wish to efficiently compute the power spectrum

$$p_{n\nu l} = \sqrt{\frac{8\pi^2}{2l+1}} \sum_m c_{nlm} c_{\nu lm}^*, \tag{8}$$

where c_{nlm} are expansion coefficients determined by

$$\rho(\vec{r}) = \sum_{nlm} c_{nlm} g_{nl}(r) Y_{lm}(\hat{r}). \tag{9}$$

$\rho(\vec{r})$ is the atomic neighbor density

$$\rho(\vec{r}) = \sum_i \exp(-\alpha|\vec{r} - \vec{r}_i|^2) \tag{10}$$

and $g_{nl}(r)$ are an orthonormal set of radial basis functions

$$\int dr\, r^2 g_{nl}(r) g_{n'l}(r) = \delta_{nn'}. \tag{11}$$

Note that we allow the radial basis set to depend on l.

From [3], we have the following exact formula for $\rho(\vec{r})$:

$$\rho(\vec{r}) = \sum_{ilm} c_{lm}^i(r) Y_{lm}(\hat{r}) \tag{12}$$

3

where

$$c_{lm}^i(r) = 4\pi \exp(-\alpha(r^2 + r_i^2))i_l(2\alpha r r_i)Y_{lm}^*(\hat{r}_i). \tag{13}$$

Here $i_l(x)$ is the modified spherical Bessel function of the first kind and is related to the modified Bessel function of the first kind $I_l(x)$ according to

$$i_l(x) = \sqrt{\frac{\pi}{2x}}I_{l+1/2}(x). \tag{14}$$

Let $\phi_{nl}(r)$ be a set of functions and let

$$g_{nl}(r) = \sum_{n'} U_{nn'}\phi_{n'}(r) \tag{15}$$

be an orthonormal basis constructed out of $\phi_n(r)$. Then

$$c_{nlm} = \sum_{in'} 4\pi Y_{lm}^*(\hat{r}_i)U_{nn'}I_{n'l}(r_i) \tag{16}$$

where

$$I_{n'l} \equiv \int_0^\infty dr r^2 \exp(-\alpha(r^2 + r_i^2))i_l(2\alpha r r_i)\phi_{n'l}(r). \tag{17}$$

The integral identity [4]

$$\int_0^\infty t^{\nu+1}I_\nu(bt)\exp(-p^2t^2)dt = \frac{b^\nu}{(2p^2)^{\nu+1}}\exp\left(\frac{b^2}{4p^2}\right) \tag{18}$$

suggests that we choose the functions

$$\phi_{nl}(r) = r^l \exp\left(-\beta_n r^2\right), \tag{19}$$

which gives

$$I_{nl}(r_i) = \frac{1}{4}\sqrt{\frac{\pi}{\alpha r_i}}\exp\left(-\frac{\alpha\beta_n r_i^2}{\alpha + \beta_n}\right)\frac{(\alpha r_i)^{l+1/2}}{(\alpha + \beta_n)^{l+3/2}}. \tag{20}$$

Defining

$$a_{nl}(r_i) \equiv \sum_{n'} U_{nn'}I_{n'l}(r_i), \tag{21}$$

we may write the expansion coefficients c_{nlm} as

$$c_{nlm} = \sum_i 4\pi a_{nl}(r_i)Y_{lm}^*(\hat{r}_i). \tag{22}$$

The power spectrum then takes the form

$$p_{n\nu l} = \frac{32\pi^3\sqrt{2}}{\sqrt{2l+1}}\sum_{mii'} a_{nl}(r_i)a_{\nu l}(r_{i'})Y_{lm}^*(\hat{r}_i)Y_{lm}(\hat{r}_{i'})$$

$$= \frac{32\pi^3\sqrt{2}}{\sqrt{2l+1}}\sum_{ii'} a_{nl}(r_i)a_{\nu l}(r_{i'})\left(\sum_m Y_{lm}^*(\hat{r}_i)Y_{lm}(\hat{r}_{i'})\right). \tag{23}$$

4

By the addition theorem

$$\sum_m Y_{lm}(\hat{x})Y_{lm}^*(\hat{y}) = \frac{2l+1}{4\pi}P_l(\hat{x}\cdot\hat{y}),\tag{24}$$

where P_l is the Legendre polynomial of degree l, the expression for the power spectrum becomes

$$\boxed{p_{n\nu l} = 8\pi^2\sqrt{4l+2}\sum_{ii'}a_{nl}(r_i)a_{\nu l}(r_{i'})P_l(\hat{r}_i\cdot\hat{r}_{i'})}\tag{25}$$

IV.   COMPUTING THE DERIVATIVE OF THE POWER SPECTRUM

To predict forces, the derivative of the power spectrum with respect to the Cartesian coordinates of the central atom is needed. Let ξ_{ki} denote the kth Cartesian component of atom i and ξ_{k0} denote the kth Cartesian component of the central atom. The derivative of the power spectrum is given by

$$\boxed{\begin{aligned}\frac{\partial p_{n\nu l}}{\partial \xi_{k0}} = 8\pi^2\sqrt{4l+2}\sum_{ii'}\Bigg(&\frac{\partial a_{nl}(r_i)}{\partial r_i}\frac{\partial r_i}{\partial \xi_{k0}}a_{\nu l}(r_{i'})P_l(\hat{r}_i\cdot\hat{r}_{i'})+\\&a_{nl}(r_i)\frac{\partial a_{\nu l}(r_{i'})}{\partial r_{i'}}\frac{\partial r_{i'}}{\partial \xi_{k0}}P_l(\hat{r}_i\cdot\hat{r}_{i'})+\\&a_{nl}(r_i)a_{\nu l}(r_{i'})\frac{\partial P_l(\hat{r}_i\cdot\hat{r}_{i'})}{\partial(\hat{r}_i\cdot\hat{r}_{i'})}\frac{\partial(\hat{r}_i\cdot\hat{r}_{i'})}{\partial \xi_{k0}}\Bigg)\end{aligned}}\tag{26}$$

We now show how each of the derivatives appearing in this expression can be efficiently evaluated. First, note that

$$a'_{nl}(r_i)\equiv\frac{\partial a_{nl}}{\partial r_i}=\sum_{n'}U_{nn'}I'_{n'l}(r_i),\tag{27}$$

where

$$I'_{nl}(r_i)\equiv\frac{\partial I_{nl}}{\partial r_i}=\frac{\sqrt{\pi}\alpha}{4}\left(\frac{1}{\alpha r_i}\right)^{3/2}\frac{(\alpha r_i)^{l+1/2}}{(\alpha+\beta_n)^{l+5/2}}(-2r_i^2\alpha\beta_n+l(\alpha+\beta_n)).\tag{28}$$

Next, note that

$$\frac{\partial r_i}{\partial \xi_{k0}}=-\frac{\xi_{ki}}{r_i},\tag{29}$$

assuming the central atom is fixed at the origin. The derivative of the Legendre polynomial may be efficiently evaluated using the identity

$$\frac{dP_n(x)}{dx}=\frac{n}{x^2-1}(xP_n(x)-P_{n-1}(x)).\tag{30}$$

Finally, it is easy to show that

$$\frac{\partial(\hat{r}_i \cdot \hat{r}_j)}{\partial \xi_0} = \frac{-(\xi_{ki} + \xi_{kj})r_i r_j + (\hat{r}_i \cdot \hat{r}_j)(r_j^2 \xi_{ki} + r_i^2 \xi_{kj})}{r_i^2 r_j^2}, \tag{31}$$

again assuming that the central atom is fixed at the origin.

V. COMPUTING THE DERIVATIVE OF THE SOAP KERNEL

The full SOAP kernel is the dot product of normalized power spectra raised to a positive power. In our approach, the covariance between local energies ϵ_i is modelled by the SOAP kernel,

$$\langle \epsilon_i \epsilon_j \rangle = K(\hat{p}_i, \hat{p}_j) = (\hat{p}_i \cdot \hat{p}_j)^\xi, \tag{32}$$

where $\hat{p}_i$ denotes the normalized power spectrum of environment i. To model forces, the covariance of the partial derivatives of the local energies is needed:

$$\begin{aligned}
\left\langle \frac{\partial \epsilon_i}{\partial \xi_{ik}}, \frac{\partial \epsilon_j}{\partial \xi_{jl}} \right\rangle &= \frac{\partial^2 \langle \epsilon_i \epsilon_j \rangle}{\partial \xi_{ik} \partial \xi_{jl}} \\
&= \frac{\partial^2 K(\hat{p}_i, \hat{p}_j)}{\partial \xi_{ik} \partial \xi_{jl}} \\
&= \sum_{\alpha\beta} \frac{\partial p_{i\alpha}}{\partial \xi_{ik}} \frac{\partial^2 K(\hat{p}_i, \hat{p}_j)}{\partial p_{i\alpha} \partial p_{j\beta}} \frac{\partial p_{j\beta}}{\partial \xi_{jl}},
\end{aligned} \tag{33}$$

where $p_{i\alpha}$ is component α of the normalized power spectrum $\hat{p}_i$. The derivatives of the power spectra with respect to Cartesian components may be evaluated using the expressions from Section 2. The remaining term is

$$\begin{aligned}
\frac{\partial^2 K(\hat{p}_i, \hat{p}_j)}{\partial p_{i\alpha} \partial p_{j\beta}} &= \frac{\partial}{\partial p_{i\alpha}} \left(\xi(\hat{p}_i \cdot \hat{p}_j)^{\xi-1} p_{i\beta} \right) \\
&= \xi(\xi-1)(\hat{p}_i \cdot \hat{p}_j)^{\xi-2} p_{j\alpha} p_{i\beta} + \xi(\hat{p}_i \cdot \hat{p}_j)^{\xi-1} \delta_{\alpha\beta} \\
&= \xi(\hat{p}_i \cdot \hat{p}_j)^{\xi-2} \left((\xi-1) p_{j\alpha} p_{i\beta} + (\hat{p}_i \cdot \hat{p}_j) \delta_{\alpha\beta} \right).
\end{aligned} \tag{34}$$

Notice that if $\xi = 1$, the double sum collapses into a single sum. Raising the SOAP kernel to a power $\xi > 1$ therefore comes at a significant computational cost. In [5], it is claimed that the role of ξ is to sharpen the selectivity of the similarity measure. However, it seems that the hyperparameter α plays the exact same role. To see this, notice that in the limit of infinite α, the atomic density becomes a sum of Dirac delta functions, and the covariance is nonzero only if the atomic environments are identical. In the limit $\alpha \to 0$, the atomic density becomes a constant, and all atomic densities are treated the same (up to a constant). Tuning α properly may make the ξ hyperparameter unnecessary.

VI. 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

$$\rho(\vec{r}) = \sum_i \exp\left(-\alpha|\vec{r} - \vec{r}_i|^2\right), \tag{35}$$

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 [3] as:

$$K(\rho, \rho') = \left(\frac{k(\rho, \rho')}{\sqrt{k(\rho, \rho)k(\rho', \rho')}}\right)^\xi, \tag{36}$$

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

$$k(\rho, \rho') = \int d\hat{R} \left|\int \rho(\vec{r})\rho'(\hat{R}\vec{r})d\vec{r}\right|^n, \tag{37}$$

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 [3], 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

$$\int dr r^2 g_n(r)g_{n'}(r) = \delta_{nn'}, \tag{38}$$

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

$$\rho(\vec{r}) = \sum_{nlm} c_{nlm} g_n(r) Y_{lm}(\vec{r}). \tag{39}$$

We first compute that

$$
\begin{aligned}
S(\rho, \hat{R}\rho') &\equiv \int d\vec{r}\rho(\vec{r})\rho'(\hat{R}\vec{r}) \\
&= \int d\vec{r}\left(\sum_{nlm} c^*_{nlm}g_n(r)Y^*_{lm}(\hat{r})\right)\left(\sum_{n'l'm'} c'_{n'l'm'}g_{n'}(r)Y_{l'm'}(\hat{R}\vec{r})\right) \\
&= \sum_{nlmn'l'm'} c^*_{nlm}c'_{n'l'm'}\int dr r^2 g_n(r)g_{n'}(r)\int d\hat{r}Y^*_{lm}(\hat{r})Y_{l'm'}(\hat{R}\vec{r}) \\
&= \sum_{nlml'm'} c^*_{nlm}c'_{nl'm'}\int d\hat{r}Y^*_{lm}(\hat{r})Y_{l'm'}(\hat{R}\vec{r}).
\end{aligned}
\tag{40}
$$

Note that

$$
\begin{aligned}
Y_{l'm'}(\vec{R}\hat{r}) &= \langle \vec{r}|\hat{R}^\dagger|Y_{l'm'}\rangle \\
&= \sum_{m''} \langle \vec{r}|Y_{l'm''}\rangle \langle Y_{l'm''}|\hat{R}^\dagger|Y_{l'm'}\rangle \\
&= \sum_{m''} \langle \vec{r}|Y_{l'm''}\rangle \langle Y_{l'm'}|\hat{R}|Y_{l'm''}\rangle^* \\
&= \sum_{m''} Y_{l'm''}(\hat{r})D^*(\hat{R})^{l'}_{m'm''},
\end{aligned}
\tag{41}
$$

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

$$
D(\hat{R})^l_{mm'} = \langle Y_{lm}|\hat{R}|Y_{lm'}\rangle.
\tag{42}
$$

Plugging into Eq. (40) gives

$$
\begin{aligned}
S(\rho, \hat{R}\rho') &= \sum_{nlml'm'm''} c^*_{nlm}c'_{nl'm'}D^*(\hat{R})^{l'}_{m'm''}\int d\hat{r}\, Y^*_{lm}(\hat{r})Y_{l'm''}(\hat{r}) \\
&= \sum_{nlmm'} c^*_{nlm}c'_{nlm'}D^*(\hat{R})^l_{m'm}.
\end{aligned}
\tag{43}
$$

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

$$
\begin{aligned}
k(\rho, \rho') &= \int d\hat{R}\, S(\rho, \hat{R}\rho')S^*(\rho, \hat{R}\rho') \\
&= \sum_{nlmm'n'\lambda\mu\mu'} c^*_{nlm}c'_{nlm'}c_{n'\lambda\mu}(c'_{n'\lambda\mu'})^* \int d\hat{R}\, D^*(\hat{R})^l_{m'm}D(\hat{R})^\lambda_{\mu'\mu} \\
&= \sum_{nlmm'n'} \frac{8\pi^2}{2l+1}c^*_{nlm}c'_{nlm'}c_{n'lm}(c'_{n'lm'})^* \\
&= \sum_{nn'l} \frac{8\pi^2}{2l+1}\left(\sum_m c^*_{nlm}c_{n'lm}\right)\left(\sum_{m'} c'_{nlm'}(c'_{n'lm'})^*\right) \\
&= \sum_{nn'l} \frac{8\pi^2}{2l+1}p_{n'nl}p'_{nn'l},
\end{aligned}
\tag{44}
$$

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

$$
k(\rho, \rho') = \vec{p} \cdot \vec{p'}.
\tag{45}
$$

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

$$
K(\rho, \rho') = \left(\frac{\vec{p}}{|\vec{p}|} \cdot \frac{\vec{p'}}{|\vec{p'}|}\right)^\xi.
\tag{46}
$$

VII.   REVIEW OF THE MULTI-COMPONENT SOAP KERNEL

We now review the multi-species version of the SOAP kernel proposed in [1]. 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:

$$K_{\mathrm{MOP}}(\rho, \rho') = \left(\frac{k_{\mathrm{MOP}}(\rho, \rho')}{\sqrt{k_{\mathrm{MOP}}(\rho, \rho)k_{\mathrm{MOP}}(\rho', \rho')}} \right)^{\xi}, \tag{47}$$

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

$$
\begin{aligned}
k_{\mathrm{MOP}}(\rho, \rho') &= \int d\hat{R} \left| \int \left(\rho_A(\vec{r})\rho'_A(\hat{R}\vec{r}) + \rho_B(\vec{r})\rho'_B(\hat{R}\vec{r}) \right) d\vec{r} \right|^2 \\
&= \int d\hat{R} \left| \int \rho_A(\vec{r})\rho'_A(\hat{R}\vec{r})d\vec{r} \right|^2 + \int d\hat{R} \left| \int \rho_B(\vec{r})\rho'_B(\hat{R}\vec{r})d\vec{r} \right|^2 \\
&\quad + \int d\hat{R} \left(\int d\vec{r}\rho_A(\vec{r})\rho'_A(\hat{R}\vec{r}) \right) \left(\int d\vec{r}\rho_B(\vec{r})\rho'_B(\hat{R}\vec{r}) \right)^* \\
&\quad + \int d\hat{R} \left(\int d\vec{r}\rho_A(\vec{r})\rho'_A(\hat{R}\vec{r}) \right)^* \left(\int d\vec{r}\rho_B(\vec{r})\rho'_B(\hat{R}\vec{r}) \right) \\
&= k(\rho_A, \rho'_A) + k(\rho_B, \rho'_B) + k_{\mathrm{cross}}(\rho, \rho') + k^*_{\mathrm{cross}}(\rho, \rho'),
\end{aligned}
\tag{48}
$$

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

$$\int d\vec{r}\rho_A(\vec{r})\rho'_A(\hat{R}\hat{r}) = \sum_{nlmm'} (c^A_{nlm})^* c'^A_{nlm} D^*(\hat{R})^l_{m'm}, \tag{49}$$

$$\int d\vec{r}\rho_B(\vec{r})\rho'_B(\hat{R}\hat{r}) = \sum_{n'\lambda\mu\mu'} (c^B_{n'\lambda\mu})^* c'^B_{n'\lambda\mu} D^*(\hat{R})^l_{\mu'\mu}, \tag{50}$$

9

and thus

$$
\begin{aligned}
k_{\text{cross}}(\rho, \rho') &= \int d\hat{R} S(\rho_A, \hat{R}\rho'_A) S^*(\rho_B, \hat{R}\rho'_B) \\
&= \sum_{nlmm'n'\lambda\mu\mu'} (c^A_{nlm})^* c'^A_{nlm'} c^B_{n'\lambda\mu} (c'^B_{n'\lambda\mu'})^* \int d\hat{R} D^*(\hat{R})^l_{m'm} D(\hat{R})^\lambda_{\mu'\mu} \\
&= \sum_{nlmm'n'} \frac{8\pi^2}{2l+1} (c^A_{nlm})^* c'^A_{nlm'} c^B_{n'lm} (c'^B_{n'lm'})^* \\
&= \sum_{nn'l} \frac{8\pi^2}{2l+1} \left(\sum_m (c^A_{nlm})^* c^B_{n'lm} \right) \left(\sum_m c'^A_{nlm'} (c'^B_{n'lm'})^* \right) \\
&= \vec{p}_{AB} \cdot \vec{p}'_{AB},
\end{aligned}
\tag{51}
$$

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

$$
k_{\text{MOP}}(\rho, \rho') = \vec{p}_A \cdot \vec{p}'_A + \vec{p}_B \cdot \vec{p}'_B + 2\vec{p}_{AB} \cdot \vec{p}'_{AB}.
\tag{52}
$$

Notice that if one of the densities vanishes, e.g. $\vec{\rho}_A = 0$, 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] S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, Physical Chemistry Chemical Physics **18**, 13754 (2016).

[2] A. Glielmo, C. Zeni, and A. De Vita, Physical Review B **97**, 184307 (2018).

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

[4] Digital Library of Mathematical Functions: https://dlmf.nist.gov/10.43.

[5] W. J. Szlachta, A. P. Bartók, and G. Csányi, Physical Review B **90**, 104108 (2014).