Constructing a Moment Tensor Potential (MTP) for the Uranium-Oxygen system

Lance J. Nelson, Lydia Harris, and Eli Harris
Department of Physics, Brigham Young University - Idaho, Rexburg, Idaho 83440, USA
(Dated: June 26, 2019)

Recent advances in materials modeling have produced models with an off-lattice basis. This is a dramatic improvement over previous state-of-the-art methods, which are either confined to a lattice (cluster expansion) or are not a basis expansion model (classical potentials) and therefore cannot be systematically improved. In this work we explore the use of the Moment Tensor Potential (MTP) by building a model for the Ag-Au binary system and later for the Uranium-Oxygen system. The binary Ag-Cu system was chosen as a test case to familiarize us with the model on a system that is expected to have small relaxations and for which first-principles calculations are relatively simple. Results indicate that the MTP models are highly accurate and may be fast enough to be used in many-atom simulations.

I. INTRODUCTION

Technological advances are driven by our ability to discover new, high-performing materials. Due mostly to advances in computing power and computational methods, the role of computation in the discovery process has increased dramatically in recent years. Computational studies help uncover the underlying physics and often narrow the search space for those working to manufacture the technology, saving time and money. Examples of recent materials/technologies whose development and discover have been aided by computational studies include thermoelectrics, solar cells, nuclear fuel, topological insulators, nanoparticles, and others.

Computational studies often involve large-scale searches over many candidate crystal structures to find materials that are both i) thermodynamically stable, and ii) optimal performing. For example Maisel et. al. studied the elastic properties of four binary metallic systems by searching over nearly 1.5 million fcc-based candidate configurations. These searches revealed the ground-state configurations of the system as well as the elastic properties of these ground states.[1] Furthermore, the search revealed important trends about the elastic properties of the materials and helped uncover physical principles that may carry over to other systems.

Another example of computation playing a key role in a technological advancement is in thermoelectric technology. A thermoelectric is a material in which a spontaneous electrical potential appears with the application of a temperature gradient. An optimal thermoelectric material will have high electrical conductivity and low thermal conductivity. Since electrons are partly responsible for conducting heat through a material, a material with high electrical conductivity typically also has high thermal conductivity. Hence, finding a high-performing thermoelectric involves a search for a very specific, rare material. In a 2011 study, Curtarolo et al [2] used high-throughput to construct a database of several thousand candidate materials. Their search revealed important trends in thermoelectric physics and identified the best performer.

Regardless of the material being studied, finding a way to increase the speed of the calculation while maintaining first-principles level accuracy is central. Because of this, much effort and resources have been committed to the task of model building. The moment tensor potential (MTP) is a recently-emerged model that combines the best properties of two long-standing models: cluster expansion and classical potentials. MTP is a basis expansion problem, like the cluster expansion, but is not confined to a parent lattice, like a classical potential. Hence with MTP you get a model that can be systematically improved over a large domain of atomic configurations, including off-lattice configurations. In this work we will investigate the feasibility of using MTP to model uranium dioxide, with the hope that we can achieve greater accuracy and/or a increase in calculation speed.

A. Classical Potentials

Density functional theory (DFT), a mainstay for any computationalist, is long considered the most accurate tool for performing materials calculations. However, even with the advances in computing power, some problems remain impractical for density functional theory. Exhaustive searches involving hundreds of thousands of calculations is impractical due to the shear volume of candidates. Calculations involving hundreds of thousand of atoms (like a grain boundary calculation) are also impractical for DFT due to the large number of atoms being considered.

To circumvent these challenges, a small set of highly-accurate DFT data is typically used to construct a model that can compute much faster. One widely-used class of models are classical potentials, which seek to capture the interaction energies between pairs of atoms using a classical approach/function. For example, the canonical Lennard-Jones potential, shown in figure 1 is given by:

$$4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}\right] \tag{1}$$

Classical potentials typically have a handful of parameters (σ and ϵ for the LJ potential) that are chosen by

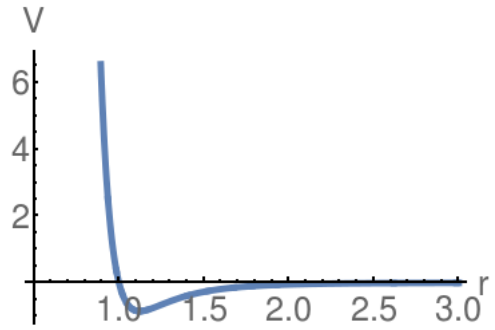

FIG. 1: Lennard-Jones potential

fitting to experimental or computational data. Because classical potentials are not a basis expansion model there is no way to systematically improve the model and expand their applicability. As such, a classical potential is laborious to construct and usually predicts accurately only in a small region of configuration space/ atomic environments.

B. The cluster expansion

Another commonly-used model for exploring substitutional order in materials is the cluster expansion, which provides a fast, accurate way to compute the total energy of all atomic configurations on a parent lattice.[3–5] The cluster expansion is constructed by first assigning each atomic type a pseudo-"spin" variable. Any atomic configuration on the parent lattice can then be specified using a vector of pseudo-spin variables. The physical quantity of interest is then expressed as a linear combination of basis functions, an idea very analogous to a Taylor or Fourier expansion.

$$E(\sigma) = E_0 + \sum_f \bar{\Pi}_f(\sigma) J_f, \tag{2}$$

Here the argument to the function, σ, is a vector of pseudo-spin variables indicating the atomic occupation on the parent lattice sites. The vector σ represents a specific structure (unit cell and atomic configuration). The $\bar{\Pi}_f$ are the basis functions, often referred to as cluster functions, with each function corresponding to a cluster of lattice sites. For binary systems, these basis functions are evaluated by averaging over products of pseudo-spin variables (For higher component systems, the basis is more complex). The expansion coefficients J_f are called effective cluster interactions (ECI's) and finding their values is the critical task when constructing a cluster expansion.

The cluster expansion is essentially a linear algebra problem:

$$\bar{\mathbf{\Pi}}\mathbf{J} = \mathbf{E} \tag{3}$$

with $\mathbf{E}$ containing the first-principles training data, and $\mathbf{J}$ the sought-after coefficients. Early in the development of cluster expansion, the ECI's were found by directly inverting Eq. 3. This so-called structure inversion method (SIM)[6] is conceptually appealing, but in practice the resulting model has poor predictive capability. As the CE method developed the best practice that emerged was to generate more fitting data than fitting variables (more elements in the vector $\mathbf{E}$ than in the ECI's vector $\mathbf{J}$). This results in an overdetermined problem that can be solved, in the least-squares sense, by singular value decomposition or related methods. Before discussing the fitting approaches in more detail, we point out that whatever the details of the fitting procedure are, any method must deal with two difficulties: (1) The expansion given in Eq. 2 must be truncated to a finite (and typically small) number of terms, and (2) a choice must be made about which structures (among a practically infinite set) should be used as training data (to generate the vector $\mathbf{E}$). The expansion must be severely truncated so that it has fewer terms than the number of training structures (maintaining an overdetermined problem), and the training structures should be chosen to minimize the predictive errors. Mathematically speaking, the choice of the training structures is not independent of the truncation.

II. MOMENT TENSOR POTENTIALS

The major drawback of the cluster expansion is that you can only consider atomic configurations on the parent lattice. Of course in real system, the atoms will relax away from those ideal lattice sites, and once that happens, the model fails. Cluster expansion practitioners would typically associate an unrelaxed crystal with the relaxed energy value and hope that there was a one-to-one mapping between the crystal structures and energies. There are many examples of systems where the cluster expansion failed to converge due to excessive relaxations.

The advantage of cluster expansion is it's basis; all of the mathematical framework associated with solving linear systems comes with the cluster expansion and hence the model can be improved to arbitrary precision by adding more basis functions to the expansion.

Classical potentials, in contrast to cluster expansion, are not confined to a lattice but are capable of calculating for any atomic geometry. However, they lack the mathematical foundation that comes with a basis expansion. This usually means that the range of validity is limited. The moment tensor potential combines the strengths of both methods; It is a basis expansion problem, but is not confined to a lattice.

The MTP was developed by Shapeev et al.[7]. What follows next is a summary of the MTP basis with some examples to help illustrate.

A. Definition of the basis

As with any basis expansion problem, we will write the MTP basis expansion as

$$V(\eta) = \sum_{\alpha} \xi_{\alpha} B_{\alpha}(\eta) \tag{4}$$

where $B_{\alpha}(\eta)$ are the basis functions, ξ_{α} are the sought-after coefficients, and η is the atomic environment(crystal) being evaluated.

To understand how the B_{α} are constructed, we must first learn about the following tensors:

$$M_{\mu,\nu}(\eta_i) = \sum_{j} f_{\mu}(|\vec{r}_{ij}|, z_i, z_j) \underbrace{\mathbf{r}_{ij} \otimes \ldots \mathbf{r}_{ij}}_{\nu \text{ times}} \tag{5}$$

The rank of these tensors is determined by the parameter ν, which specifies how many dyadic products are performed: $\underbrace{\mathbf{r}_{ij} \otimes \ldots \mathbf{r}_{ij}}_{\nu \text{ times}}$. For $\nu = 1$, $\underbrace{\mathbf{r}_{ij} \otimes \ldots \mathbf{r}_{ij}}_{\nu \text{ times}} = \mathbf{r}_{ij}$, for $\nu = 2$, $\underbrace{\mathbf{r}_{ij} \otimes \ldots \mathbf{r}_{ij}}_{\nu \text{ times}} = \mathbf{r}_{ij} \otimes \mathbf{r}_{ij}$, and so forth. To perform the summation in equation 5, a cutoff radius is first chosen and centered on the atom of interest. $\mathbf{r}_{ij}$ is the vector connecting the central atom to one of it's neighbors and the sum is performed over all of the atoms inside the cutoff radius. (see figures 2 and 3)

B. Radial Functions

The scalar radial functions ($f_{\mu}(r_{ij}, z_i, z_j)$) are expressed as an expansion over Chebyshev polynomials:

$$f_{\mu}(|\vec{r}_{ij}|, z_i, z_j) = \sum_{k} c^{k}_{\mu,z_i,z_j} Q^{k}(\rho) \tag{6}$$

where

$$Q(\rho) = T_k(\rho)(R_{\text{cut}} - \rho)^2 \tag{7}$$

where $T_k(\rho)$ are Chebyshev polynomials of order k and $\rho = |r_{ij}|$. The c^{k}_{μ,z_i,z_j} coefficients are unknown a priori and belong to the set of fitting parameters that are extracted during the fitting process. As you can see, the presence of the c^{k}_{μ,z_i,z_j} make this model non-linear. The full set of coefficients to be extracted is $\{\xi_{\alpha}, c^{k}_{\mu,z_i,z_j}\}$.

C. Example

To help illustrate this basis, consider the two-dimensional crystal shown in Figure 2. We'll choose $R_{\text{cut}} = 1.5$, and $k = 5$. We'll also let red atoms be type 1 and green atoms be type 2, so that $z_1 = z_3 = z_4 = z_5 = z_7 = z_9 = z_{10} = z_{11} = z_{13} = 1$ and $z_2 = z_{12} = z_6 = z_8 = 2$. To simplify the illustration, we'll assume that the coefficients c^{k}_{μ,z_i,z_j} are known. Notice that for each choice of μ and k, there is a matrix of coefficients, a 2 x 2 in this case. Arbitrarily, we'll set that matrix to be

$$\begin{pmatrix} c^{k}_{\mu,1,1} & c^{k}_{\mu,1,2} \\ c^{k}_{\mu,2,1} & c^{k}_{\mu,2,2} \end{pmatrix} = \begin{pmatrix} 10 & 20 \\ 20 & 50 \end{pmatrix}$$

for all k and μ. Respectively, the coefficients $c^{k}_{\mu,1,1}, c^{k}_{\mu,1,2}, c^{k}_{\mu,2,1}, c^{k}_{\mu,2,2}$ correspond to a red central atom with a red neighbor, a red central atom with a green neighbor, a green central atom with a red neighbor and a green central atom with a green neighbor.

$\nu = 1$ (Rank-1 Tensors)

For $\nu = 1$, the tensors $M_{\mu,\nu}(\eta_i)$ are of first order (i.e they are vectors). Equation 5 must be evaluated once for each unique atom in the crystal (η_i). Since the crystal in figure 2 has two unique atoms, we'll write down equation 5 twice. First with the red atom at the center:

$$\begin{aligned} M_{\mu,1}(\eta_1) = {} & f_{\mu}(|\vec{r}_{1,2}|, 1, 2) r_{1,2} + f_{\mu}(|\vec{r}_{1,12}|, 1, 2) r_{1,12} \\ & + f_{\mu}(|\vec{r}_{1,8}|, 1, 2) r_{1,8} + f_{\mu}(|\vec{r}_{1,6}|, 1, 2) r_{1,6} \\ & + f_{\mu}(|\vec{r}_{1,3}|, 1, 1) r_{1,3} + f_{\mu}(|\vec{r}_{1,4}|, 1, 1) r_{1,4} \\ & + f_{\mu}(|\vec{r}_{1,5}|, 1, 1) r_{1,5} + f_{\mu}(|\vec{r}_{1,7}|, 1, 1) r_{1,7} \\ & + f_{\mu}(|\vec{r}_{1,9}|, 1, 1) r_{1,9} + f_{\mu}(|\vec{r}_{1,10}|, 1, 1) r_{1,10} \\ & + f_{\mu}(|\vec{r}_{1,11}|, 1, 1) r_{1,11} + f_{\mu}(|\vec{r}_{1,13}|, 1, 1) r_{1,13} \end{aligned} \tag{8}$$

and for the green atom:

$$\begin{aligned} M_{\mu,1}(\eta_2) = {} & f_{\mu}(|\vec{r}_{2,1}|, 2, 1) r_{2,1} + f_{\mu}(|\vec{r}_{2,3}|, 2, 1) r_{2,3} \\ & + f_{\mu}(|\vec{r}_{2,4}|, 2, 1) r_{2,4} + f_{\mu}(|\vec{r}_{2,5}|, 2, 1) r_{2,5} \\ & + f_{\mu}(|\vec{r}_{2,6}|, 2, 2) r_{2,6} + f_{\mu}(|\vec{r}_{2,12}|, 2, 2) r_{2,12} \\ & + f_{\mu}(|\vec{r}_{2,17}|, 2, 2) r_{2,17} + f_{\mu}(|\vec{r}_{2,15}|, 2, 2) r_{2,15} \\ & + f_{\mu}(|\vec{r}_{2,14}|, 2, 2) r_{2,14} + f_{\mu}(|\vec{r}_{2,8}|, 2, 2) r_{2,8} \\ & + f_{\mu}(|\vec{r}_{2,18}|, 2, 2) r_{2,18} + f_{\mu}(|\vec{r}_{2,16}|, 2, 2) r_{1,16} \end{aligned}$$

Now let's do the same thing for the crystal in figure 3

$$\begin{aligned} M_{\mu,1}(\eta_1) = {} & f_{\mu}(|\vec{r}_{1,2}|, 1, 2) r_{1,2} + f_{\mu}(|\vec{r}_{1,12}|, 1, 2) r_{1,12} \\ & + f_{\mu}(|\vec{r}_{1,8}|, 1, 2) r_{1,8} + f_{\mu}(|\vec{r}_{1,6}|, 1, 2) r_{1,6} \\ & + f_{\mu}(|\vec{r}_{1,14}|, 1, 2) r_{1,14} + f_{\mu}(|\vec{r}_{1,15}|, 1, 2) r_{1,15} \\ & + f_{\mu}(|\vec{r}_{1,16}|, 1, 2) r_{1,16} + f_{\mu}(|\vec{r}_{1,17}|, 1, 2) r_{1,17} \\ & + f_{\mu}(|\vec{r}_{1,3}|, 1, 1) r_{1,3} + f_{\mu}(|\vec{r}_{1,4}|, 1, 1) r_{1,4} \\ & + f_{\mu}(|\vec{r}_{1,5}|, 1, 1) r_{1,5} + f_{\mu}(|\vec{r}_{1,7}|, 1, 1) r_{1,7} \\ & + f_{\mu}(|\vec{r}_{1,9}|, 1, 1) r_{1,9} + f_{\mu}(|\vec{r}_{1,10}|, 1, 1) r_{1,10} \\ & + f_{\mu}(|\vec{r}_{1,11}|, 1, 1) r_{1,11} + f_{\mu}(|\vec{r}_{1,13}|, 1, 1) r_{1,13} \end{aligned} \tag{9}$$

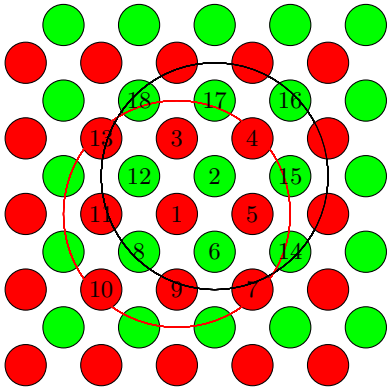

FIG. 2: Hypothetical binary, two-dimensional atomic configuration. The large red circle has radius $R_{\text{cut}} = 1.5$ and is centered at the red atom located at $(0, 0)$ and the large black circle is centered at the green atom located at $(0.5, 0.5)$.

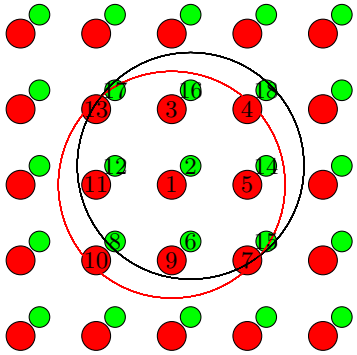

FIG. 3: Hypothetical binary, two-dimensional atomic configuration. The large red circle has radius $R_{\text{cut}} = 1.5$ and is centered at the red atom located at $(0, 0)$ and the large black circle is centered at the green atom located at $(0.25, 0.25)$. Notice that this structure has less symmetry than the structure shown in figure 2.

and on the green atom:

$$\begin{aligned} M_{\mu,1}(\eta_2) = {} & f_\mu(|\vec{r}_{2,1}|, 2, 1)r_{2,1} + f_\mu(|\vec{r}_{2,3}|, 2, 1)r_{2,3} \\ & + f_\mu(|\vec{r}_{2,4}|, 2, 1)r_{2,4} + f_\mu(|\vec{r}_{2,5}|, 2, 1)r_{2,5} \\ & + f_\mu(|\vec{r}_{2,9}|, 2, 1)r_{2,9} + f_\mu(|\vec{r}_{2,11}|, 2, 1)r_{2,11} \\ & + f_\mu(|\vec{r}_{2,13}|, 2, 1)r_{2,13} + f_\mu(|\vec{r}_{2,7}|, 2, 1)r_{2,7} \\ & + f_\mu(|\vec{r}_{2,14}|, 2, 2)r_{2,14} + f_\mu(|\vec{r}_{2,15}|, 2, 2)r_{2,15} \\ & + f_\mu(|\vec{r}_{2,6}|, 2, 2)r_{2,6} + f_\mu(|\vec{r}_{2,8}|, 2, 2)r_{2,8} \\ & + f_\mu(|\vec{r}_{2,12}|, 2, 2)r_{2,12} + f_\mu(|\vec{r}_{2,17}|, 2, 2)r_{2,17} \\ & + f_\mu(|\vec{r}_{2,16}|, 2, 2)r_{2,16} + f_\mu(|\vec{r}_{2,18}|, 2, 2)r_{2,18} \end{aligned}$$

Tables I and II summarize the calculation of $M_{\mu,1}$ and $M_{\mu,2}$ respectively. Each row in the table is a term in the sum shown in equation 5. Let's notice a few things about table I. One interpretation of the $M_{\mu,1}$ tensor is as the center of mass coordinates for this shell of atoms, with the weight not being mass but instead being the value of the scalar function $f_\mu(|\vec{r}_{ij}|, z_i, z_j)$. In the atomic configuration from figure 2, due to high symmetry, the center of mass is located at the origin (directly on the centering atom). Hence, $M_{\mu,1}$ is the zero vector. In contrast, the configuration in figure 3, which has lower symmetry, the "center of mass" is shifted off the central atom. Thus we can conclude that there are more atoms in the direction of $M_{\mu,1}$ than in the opposite direction.

Table II show the rank-2 tensors involved in calculating the sum of equation 5 just for the crystal in figure 3. One can interpret these tensors to indicate the degree to which the shell of atoms is "squeezed" in the direction of $M_{\mu,1}$. One can begin to see how these functions successfully distinguish between different crystal structures.

Notice that since $c^k_{\mu,1,2} = c^k_{\mu,2,1} = 20$ all of the $f_\mu(|\vec{r}_{ij}|, 1, 2)$ are equal. If we had chosen them to be different, the vectors being summed would have been different but would have still summed to zero. This is due to the high symmetry of the configuration.

D. Contractions

The tensors $M_{\mu,\nu}(\eta_i)$ themselves are not basis functions. To evaluate the basis function $B_\alpha(\eta)$ we must enumerate all possible contractions of any combination of tensors. When constructing these contraction, it's best to use Einstein summation convention. In Einstein summation convention, any repeated indices imply a sum. For example, the expression:

$$v_a v_a$$

implies

$$v_1 v_1 + v_2 v_2 + v_3 v_3$$

and

$$v_a v_b$$

implies

$$\begin{pmatrix} v_1 v_1 & v_1 v_2 \\ v_2 v_1 & v_2 v_2 \end{pmatrix}$$

In order to contract some combination of tensors down to a scalar, the total number of indices present in an expression must be even. If there were an odd number of indices, there would be an unrepeated index, implying that the result is a vector, not a scalar.

Let's look at a few examples of contractions of the moment tensors. For example, you could take two rank-1 tensors and contract them using the dot product:

$$M_a M_a = M_a \cdot M_a + M_b \cdot M_b + M_c \cdot M_c$$

or you could combine a rank-2 tensor with a rank-1 tensor and contract as:

		Centered on Red Atom					Centered on Green Atom			
		coordinates	norm	$f_\mu(\lvert\vec{r}_{ij}\rvert, z_i, z_j)$	$f_\mu(\lvert\vec{r}_{ij}\rvert, z_i, z_j)\vec{r}_{ij}$		coordinates	norm	$f_\mu(\lvert\vec{r}_{ij}\rvert, z_i, z_j)$	$f_\mu(\lvert\vec{r}_{ij}\rvert, z_i, z_j)\vec{r}_{ij}$
Crystal shown in figure 2	$r_{1,2}$	(0.5,0.5)	0.707	13.756	(6.8783,6.8783)	$r_{2,4}$	(0.5,0.5)	0.707	13.756	(6.8783,6.8783)
	$r_{1,3}$	(0,1.0)	1.0	-0.216	(0,-0.216)	$r_{2,17}$	(0,1.0)	1.0	-1.08	(0,-1.08)
	$r_{1,4}$	(1.0,1.0)	1.414	0.1635	(0.1635,0.1635)	$r_{2,16}$	(1.0,1.0)	1.414	0.818	(0.818,0.818)
	$r_{1,5}$	(1.0,0.0)	1.0	-0.216	(-0.216,0.0)	$r_{2,15}$	(1.0,0.0)	1.0	-1.08	(-1.08,0.0)
	$r_{1,6}$	(0.5,-0.5)	0.707	13.7566	(6.8783,-6.8783)	$r_{2,5}$	(0.5,-0.5)	0.707	13.7566	(6.8783,-6.8783)
	$r_{1,7}$	(1.0,-1.0)	1.414	0.16356	(0.1635,-0.1635)	$r_{2,14}$	(1.0,-1.0)	1.414	0.818	(0.818,-0.818)
	$r_{1,8}$	(-0.5,-0.5)	0.707	13.7566	(-6.878,-6.878)	$r_{2,1}$	(-0.5,-0.5)	0.707	13.7566	(-6.878,-6.878)
	$r_{1,9}$	(0.0,-1.0)	1.0	-0.216	(0,0.216)	$r_{2,6}$	(0.0,-1.0)	1.0	-1.08	(0,1.08)
	$r_{1,10}$	(-1.0,-1.0)	1.414	0.16356	(-0.1635,-0.1635)	$r_{2,8}$	(-1.0,-1.0)	1.414	0.8178	(-0.818,-0.818)
	$r_{1,11}$	(-1.0,0.0)	1.0	-0.216	(0.216,0.0)	$r_{2,12}$	(-1.0,0.0)	1.0	-1.08	(1.08,0.0)
	$r_{1,12}$	(-0.5,0.5)	0.707	13.7566	(-6.878,6.878)	$r_{2,3}$	(-0.5,0.5)	0.707	13.7566	(-6.878,6.878)
	$r_{1,13}$	(-1.0,1.0)	1.414	0.16356	(-0.1635,0.1635)	$r_{2,18}$	(-1.0,1.0)	1.414	0.818	(-0.818,0.818)
				Sum =	(0,0)				Sum =	(0,0)
Crystal shown in figure 3	$r_{1,2}$	(0.25,0.25)	0.3535	10.8958	(2.723,2.723)	$r_{2,1}$	(-0.25,-0.25)	0.3535	10.8958	(-2.723,-2.723)
	$r_{1,3}$	(0,1.0)	1.0	-0.216	(0,-0.216)	$r_{2,16}$	(0,1.0)	1.0	-1.08	(0,-1.08)
	$r_{1,4}$	(1.0,1.0)	1.414	0.1635	(0.1635,0.1635)	$r_{2,18}$	(1.0,1.0)	1.414	0.818	(0.818,0.818)
	$r_{1,5}$	(1.0,0.0)	1.0	-0.216	(-0.216,0.0)	$r_{2,14}$	(1.0,0.0)	1.0	-1.08	(-1.08,0.0)
	$r_{1,6}$	(0.25,-0.75)	0.7905	8.807	(2.20,-6.60525)	$r_{2,3}$	(-0.25,0.75)	0.7905	8.807	(-2.20,6.60525)
	$r_{1,7}$	(1.0,-1.0)	1.414	0.16356	(0.1635,-0.1635)	$r_{2,15}$	(1.0,-1.0)	1.414	0.818	(0.818,-0.818)
	$r_{1,8}$	(-0.75,-0.75)	1.06	-1.305	(0.97875,0.97875)	$r_{2,4}$	(0.75,0.75)	1.06	-1.305	(-0.97875,-0.97875)
	$r_{1,9}$	(0.0,-1.0)	1.0	-0.216	(0,0.216)	$r_{2,6}$	(0.0,-1.0)	1.0	-1.08	(0,1.08)
	$r_{1,10}$	(-1.0,-1.0)	1.414	0.16356	(-0.1635,-0.1635)	$r_{2,8}$	(-1.0,-1.0)	1.414	0.818	(-0.818,-0.818)
	$r_{1,11}$	(-1.0,0.0)	1.0	-0.216	(0.216,0.0)	$r_{2,12}$	(-1.0,0.0)	1.0	-1.08	(1.08,0.0)
	$r_{1,12}$	(-0.75,0.25)	0.7905	8.807	(-6.605,2.20175)	$r_{2,5}$	(0.75,-0.25)	0.7905	8.807	(6.605,-2.20175)
	$r_{1,13}$	(-1.0,1.0)	1.414	0.16356	(-0.1635,0.1635)	$r_{2,17}$	(-1.0,1.0)	1.414	0.818	(-0.818,0.818)
	$r_{1,14}$	(1.25,0.25)	1.2747	-0.05003	(-0.0625,-0.0125075)	$r_{2,11}$	(-1.25,-0.25)	1.2747	-0.05003	(0.0625,0.0125075)
	$r_{1,15}$	(1.25,-0.75)	1.457	0.12411	(0.15513,-0.093)	$r_{2,13}$	(-1.25,0.75)	1.457	0.12411	(-0.15513,0.093)
	$r_{1,16}$	(0.25,1.25)	1.2747	-0.05003	(-0.0125,-0.0625)	$r_{2,9}$	(-0.25,-1.25)	1.2747	-0.05003	(0.0125,0.0625)
	$r_{1,17}$	(-0.75,1.25)	1.457	0.12411	(-0.0930,0.15513)	$r_{2,7}$	(0.75,-1.25)	1.457	0.12411	(0.0930,-0.15513)
				Sum =	(-0.7137,-0.7137)				Sum =	(0.7137,0.7137)

TABLE I: Calculation of $M_{\mu,1}$ for the 2D crystals in figures 2 and 3. Each row in the table represents a term in equation 5. The top section corresponds to the crystal

$$M_{aa}M_{bb} = \sum_a M_{aa} \cdot \sum_b M_{bb}$$

Here is another way to contract two rank-2 tensors:

$$M_{ab}M_{ab} = \sum_a M_{ab}^2$$

In other words, you square each element of matrix M_{ab} and add them all up.

Here is an example of contracting a combination of three tensors: one rank-2 and two rank-1 tensors:

$$\begin{aligned} M_{ab}M_aM_b = {} & M_{11}M_1M_1 + M_{12}M_1M_2 + M_{13}M_1M_3 \\ & + M_{21}M_2M_1 + M_{22}M_2M_2 + M_{23}M_2M_3 + \ldots \end{aligned}$$

A hierarchy of basis functions can be defined by calculating $\sum_i(2\mu_i + \nu_i)$ where i runs over all tensors used in the contraction. Truncation of the basis can be done by choosing all contractions below a chosen threshold ($\text{lev}_{\max}$).

	Centered on Red Atom					Centered on Green Atom			
	coordinates	norm	$f_\mu(\lvert\vec{r}_{ij}\rvert, z_i, z_j)$	$f_\mu(\lvert\vec{r}_{ij}\rvert, z_i, z_j)\vec{r}_{ij} \otimes \vec{r}_{ij}$		coordinates	norm	$f_\mu(\lvert\vec{r}_{ij}\rvert, z_i, z_j)$	$f_\mu(\lvert\vec{r}_{ij}\rvert, z_i, z_j)\vec{r}_{ij} \otimes \vec{r}_{ij}$
$r_{1,2}$	(0.25,0.25)	0.3535	10.8958	$\begin{pmatrix} 0.68098 & 0.68098 \\ 0.68098 & 0.68098 \end{pmatrix}$	$r_{2,1}$	(-0.25,-0.25)	0.3535	10.8958	$\begin{pmatrix} 0.68098 & 0.68098 \\ 0.68098 & 0.68098 \end{pmatrix}$
$r_{1,3}$	(0,1.0)	1.0	-0.216	$\begin{pmatrix} 0 & 0 \\ 0 & -0.216 \end{pmatrix}$	$r_{2,16}$	(0,1.0)	1.0	-1.08	$\begin{pmatrix} 0 & 0 \\ 0 & -1.08 \end{pmatrix}$
$r_{1,4}$	(1.0,1.0)	1.414	0.1635	$\begin{pmatrix} 0.1635 & 0.1635 \\ 0.1635 & 0.1635 \end{pmatrix}$	$r_{2,18}$	(1.0,1.0)	1.414	0.8178	$\begin{pmatrix} 0.818 & 0.818 \\ 0.818 & 0.818 \end{pmatrix}$
$r_{1,5}$	(1.0,0.0)	1.0	-0.216	$\begin{pmatrix} -0.216 & 0 \\ 0 & 0 \end{pmatrix}$	$r_{2,14}$	(1.0,0.0)	1.0	-1.08	$\begin{pmatrix} -1.08 & 0 \\ 0 & 0 \end{pmatrix}$
$r_{1,6}$	(0.25,-0.75)	0.7905	8.807	$\begin{pmatrix} 0.5504 & -1.6513125 \\ -1.6513125 & 4.9539 \end{pmatrix}$	$r_{2,3}$	(-0.25,0.75)	0.7905	8.807	$\begin{pmatrix} 0.5504 & -1.651312 \\ -1.6513125 & 4.9539 \end{pmatrix}$
$r_{1,7}$	(1.0,-1.0)	1.414	0.16356	$\begin{pmatrix} 0.16356 & -0.16356 \\ -0.16356 & 0.16356 \end{pmatrix}$	$r_{2,15}$	(1.0,-1.0)	1.414	0.8178	$\begin{pmatrix} 0.818 & -0.818 \\ -0.818 & 0.818 \end{pmatrix}$
$r_{1,8}$	(-0.75,-0.75)	1.06	-1.305	$\begin{pmatrix} -0.7340625 & -0.7340625 \\ -0.7340625 & -0.7340625 \end{pmatrix}$	$r_{2,4}$	(0.75,0.75)	1.06	-1.305	$\begin{pmatrix} -0.7340625 & -0.734062 \\ -0.7340625 & -0.734062 \end{pmatrix}$
$r_{1,9}$	(0.0,-1.0)	1.0	-0.216	$\begin{pmatrix} 0 & 0 \\ 0 & -0.216 \end{pmatrix}$	$r_{2,6}$	(0.0,-1.0)	1.0	-1.08	$\begin{pmatrix} 0 & 0 \\ 0 & 1.08 \end{pmatrix}$
$r_{1,10}$	(-1.0,-1.0)	1.414	0.16356	$\begin{pmatrix} 0.16356 & 0.16356 \\ 0.16356 & 0.16356 \end{pmatrix}$	$r_{2,8}$	(-1.0,-1.0)	1.414	0.8178	$\begin{pmatrix} 0.818 & 0.818 \\ 0.818 & 0.818 \end{pmatrix}$
$r_{1,11}$	(-1.0,0.0)	1.0	-0.216	$\begin{pmatrix} -0.216 & 0 \\ 0 & 0 \end{pmatrix}$	$r_{2,12}$	(-1.0,0.0)	1.0	-1.08	$\begin{pmatrix} -1.08 & 0 \\ 0 & 0 \end{pmatrix}$
$r_{1,12}$	(-0.75,0.25)	0.7905	8.807	$\begin{pmatrix} 4.9539 & -1.6513 \\ -1.6513 & 0.5504 \end{pmatrix}$	$r_{2,5}$	(0.75,-0.25)	0.7905	8.807	$\begin{pmatrix} 4.9539 & -1.6513 \\ -1.6513 & 0.5504 \end{pmatrix}$
$r_{1,13}$	(-1.0,1.0)	1.414	0.16356	$\begin{pmatrix} 0.16356 & -0.16356 \\ -0.16356 & 0.16356 \end{pmatrix}$	$r_{2,17}$	(-1.0,1.0)	1.414	0.8178	$\begin{pmatrix} 0.818 & -0.818 \\ -0.818 & 0.818 \end{pmatrix}$
$r_{1,14}$	(1.25,0.25)	1.2747	-0.05003	$\begin{pmatrix} -0.0781 & -0.015625 \\ -0.015625 & -0.003125 \end{pmatrix}$	$r_{2,11}$	(-1.25,-0.25)	1.2747	-0.05003	$\begin{pmatrix} -0.0781 & -0.015625 \\ -0.015625 & -0.003125 \end{pmatrix}$
$r_{1,15}$	(1.25,-0.75)	1.457	0.12411	$\begin{pmatrix} 0.1939 & -0.11634375 \\ -0.1163 & 0.0698 \end{pmatrix}$	$r_{2,13}$	(-1.25,0.75)	1.457	0.12411	$\begin{pmatrix} 0.1939 & -0.1163437 \\ -0.1163 & 0.0698 \end{pmatrix}$
$r_{1,16}$	(0.25,1.25)	1.2747	-0.05003	$\begin{pmatrix} -0.0031 & -0.0156 \\ -0.0156 & -0.0782 \end{pmatrix}$	$r_{2,9}$	(-0.25,-1.25)	1.2747	-0.05003	$\begin{pmatrix} -0.0031 & -0.0156 \\ -0.0156 & -0.0782 \end{pmatrix}$
$r_{1,17}$	(-0.75,1.25)	1.457	0.12411	$\begin{pmatrix} 0.069 & -0.1163 \\ -0.1163 & 0.1939 \end{pmatrix}$	$r_{2,7}$	(0.75,-1.25)	1.457	0.12411	$\begin{pmatrix} 0.069 & -0.1163 \\ -0.1163 & 0.1939 \end{pmatrix}$
			Sum =	$\begin{pmatrix} 5.855 & -3.619 \\ -3.619 & 5.855 \end{pmatrix}$				Sum =	$\begin{pmatrix} 6.744 & -3.619 \\ -3.619 & 6.744 \end{pmatrix}$

TABLE II: Calculation of $M_{\mu,2}$ for the 2D crystals in figure 3

E. Active Learning

In addition to fitting a training set to the MTP basis, the MLP (machine learning potential) package provides a framework for iteratively refining and improving the moment tensor potential. The steps for this process are outlined below:

1. A set of crystals is first defined. This set defines the search space and ought to include all atomic configurations that will be considered in the study. If the study involves an exhaustive search for ground state phases, this set of structures should include all possible candidates. If the study involves a single crystal structure (as with UO_2) this set may only include slight variations on that crystal structure.
2. Using the D-optimality criteria, an initial set of configurations is chosen to be trained on. The D-optimality criteria chooses the "most orthogonal" set of structures, where orthogonality is defined with respect to the MTP basis.
3. DFT is used to generate data for the configurations chosen.
4. The model is trained.

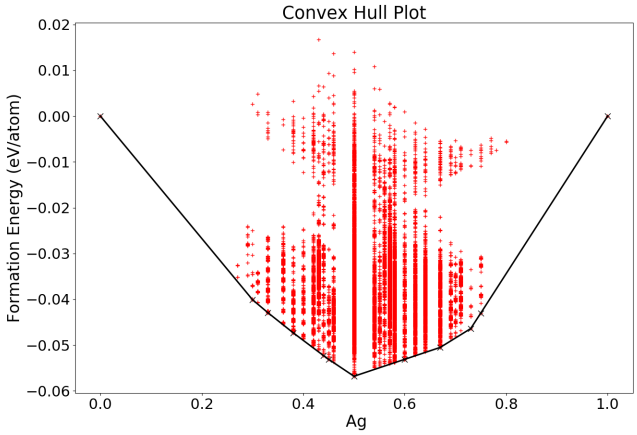

FIG. 4: Convex hull for Ag-Au as predicted by MTP. Black "X" indicates predicted ground states.

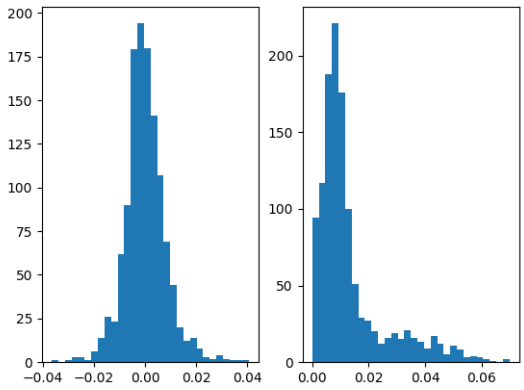

FIG. 5: Left:

5. The model attempts to relax all of the atomic configurations in the search set. If MLP encounters an extrapolation during the relaxation, it flags the structure for consideration to be included in the next training set. Extrapolation in this setting is meant to indicate that the model is calculating in a region of the space that it has not been trained on. Hence, the calculation is expected to be of questionable accuracy.

6. MLP then selects which of the extrapolation structures should be added to the training set. Once again, this is done using the D-optimality criteria used in step 1.

7. Return to step 3.

III. RESULTS

A. A test system: Ag-Au

To demonstrate the utility and accuracy of MTP, we first study a well-known binary system in which we expect the first-principles calculations to be straightforward and the relaxations to be small. Ag-Au, with lattice parameters of 4.085 and 4.078 respectively, is a good choice, and we will attempt to find the ground state crystal structures for this system. Although we have strong reason to believe that the ground states will also be fcc-derived crystal structures, we'll include fcc-, bcc-, and hcp-derived superstructures in our search set. This will demonstrate MTP's ability to perform calculations across a diverse set of atomic configurations. The total number of configurations in this set is $60,000$.

After six iterations of the active learning process, 1500 configurations had been selected for addition into the training set. The average energy error over this set was 0.65 meV/atom and the average force error was 0.0104 eV/Å. The distributions of these values over the entire training set are given in figure 5. The MTP-predicted convex hull for Ag-Au is shown in figure 4. The topology of the hull agrees well with experimental results and other computational studies of this system.

B. U-O_2

Uranium dioxide presents a much bigger challenge compared to Ag-Au, mostly because DFT calculations for this system are challenging. The f electrons present in this system become erroneously de-localized in normal DFT calculations due to self-interaction Coulomb interaction. This is typically corrected by adding a parameter (U) that accounts for the orbital dependence on the Hamiltonian. [8] Furthermore, it has been shown that even with the +U correction to DFT, the final electronic configuration may settle into an erroneous occupancy, with energies that are several electron volts above the ground state energy. Meredig et. al[9] proposed a U-ramping technique for finding the electronic minima, in which the U paramater is slowly "turned on" and ramped up to the correct value.

The corrections discussed above can be used reliably for calculations of UO_2 on the fluorite crystal structure (shown in figure 6), but are not as reliable once the atomic configuration strays from ideal fluorite. Hence, any DFT-generated training data must be very similar to fluorite in nature.

The DFT challenges associated with UO_2 seem to be to great to overcome. However, we anticipate that a molecular dynamics simulation of UO_2 will not stray too far from the ideal fluorite structure. Therefore, if we can generate a MTP that predicts well for the fluorite crystal structure and small deviations from that ideal structure, it will suffice for our purposes. Thus, our space of training configurations is much smaller than for Ag-Au.

Atom #	$\lvert F_{\mathrm{MTP}}\rvert$	$\lvert F_{\mathrm{DFT}}\rvert$	$\lvert F_{\mathrm{MTP}}\rvert - \lvert F_{\mathrm{DFT}}\rvert$
1	24.5635	24.5635	0
2	0.417038	0.416797	0.000241
3	24.5635	24.5635	0
4	0.417038	0.416797	0.000241
5	11.3386	11.3386	0
6	11.3386	11.3386	0.000241
7	11.3386	11.3386	0
8	11.3386	11.3386	0
9	2.79092	2.79084	0
10	0.297829	0.297796	0
11	2.79092	2.79084	0.000079
12	0.297829	0.297796	0.000033
13	1.5971	1.59709	0.000012
14	1.74836	1.74835	0.000012
15	2.54467	2.54467	0
16	1.59709	1.59709	0
17	1.74836	1.74835	0.000017
18	2.54468	2.54467	0.000011
19	1.59709	1.59709	0
20	1.74836	1.74835	0.000017
21	2.54468	2.54467	0.000011
22	1.5971	1.59709	0.000012
23	1.74836	1.74835	0.000012
24	1.74836	2.54467	0
		mean	0.000034

TABLE III: Predicted and calculated forces for one UO_2 structure. The mean error was 0.000034 for this structure. The large forces are due to several uranium atoms being very close to each other.

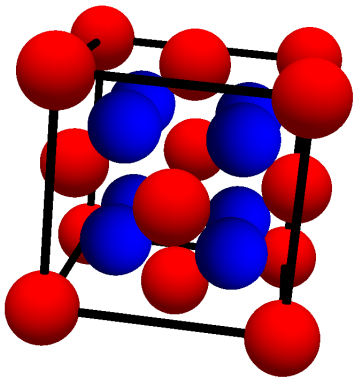

FIG. 6: Fluorite crystal structure which for uranium-dioxide(UO_2)

Total Energy (eV) predicted	**Total Energy (eV) calculated**	|diff|**(eV)**
-117.250882889086	-117.25088298	9.1e-8
-140.126355786625	-140.12635566	1.3e-7
-150.328035440859	-150.32803526	1.8e-7
-181.549228746869	-181.54922873	1.7e-8
-168.618518091217	-168.61851806	3.1e-8
-137.376277272381	-137.37627747	2.0e-7
-124.495047858794	-124.49504792	6.1e-8
	mean	1.0e-7

TABLE IV: Predicted and calculated energies for several UO_2 "fluorite-like" structures with (100) AFM ordering. The mean absolute error was $1.0e-7$eV for all structures, with a mean absolute error per atom of $6.45887 * 10^{-10}$eV/atom.

The training set was constructed with ideal (though not primitive) fluorite as the starting point. (see figure) By applying random, normally-distributed displacement vectors to all 12 atoms inside of the unit cell, we can generate a "fluorite-like" structure. The training set was chosen by MLP from a set of $10,000$ of these crystals. The hope is that this set is diverse enough to capture everything that may be encountered in a molecular dynamics simulation. This system will be a real test for the MTP since it is being asked to distinguish between atomic configurations that are very similar to one another.

As of the today's date, this work is still incomplete. 8 DFT calculations have been performed and the MTP has been trained on those 8 "fluorite- like" crystals. The accuracy of the model is depicted in figures III and IV. Clearly, the model is predicting accurately. What remains is a more complete training of the model, wherein a more complete set of training data is used for training. However, it is clear that this model shows promising results and may provide an improvement in model accuracy and speed.

IV. FUTURE WORK

The goal of the work is a better molecular dynamics simulation of UO_2. The construction of a model for UO_2 is nearing completion and the accuracy of this model promises to be high quality. The next step is to use the model in LAMMPS to perform molecular dynamics for a database of grain boundaries. Comparison of these results with those obtain previously using previos potentials will reveal whether our model represents an improvement.

V. CONCLUSION

A new model (the Moment Tensor Potential) has been thoroughly studied and applied to a model system and later to the more lab-mission-related system of UO_2. Two undergraduate students have gained valuable experience in a research environment and gained real skills in a computational science environment.

Appendix A: Participants

Both participants have been hard-working and diligent in their projects. Admittedly, they were given challenging projects and I am proud to report that they rose to the challenge. They will be great professionals some day.

1. Lydia Harris - Brigham Young University-Idaho

- Built MTP for AgAu test system
- Built MTP for anti-ferromagnetic (AFM) UO_2
- Built python code to search for AFM-compatible crystals
- Learn to use git
- Extensive experience with object-oriented python.

2. Eli Harris - Brigham Young University-Idaho

- Built MTP for AuCu test system
- Built MTP for ferromagnetic (FM) UO_2
- Built python code to automate the execution and error handling of VASP calculations.
- Learn to use git
- Extensive experience with object-oriented python.

Appendix B: Acknowledgements

We gratefully acknowledge the DOE VFP for providing funding for this experience. The skils and knowledge gained will serve those involved well and we are confident that our contribution made an impact in the realm of materials modeling. We also express gratitude to the INL high performance computing center for granting access to computing resources. We thank Yongfeng Zhang, Larry Aagesen, and Benjamin Beeler for fruitful discussions that helped guide the project.

Appendix C: Further Collaboration

This work will continue at our home institution in collaboration with INL scientists. During our time here we also discovered further overlaps in our specialties that could be leveraged to initiate future studies together. These include studies of high-entropy alloys and metallic nuclear fuels.

[1] S. B. Maisel, M. Höfler, and S. Müller, Nature **491**, 740 (2012).
[2] S. Wang, Z. Wang, W. Setyawan, N. Mingo, and S. Curtarolo, Physical Review X **1**, 021012 (2011).
[3] J. M. Sanchez, F. Ducastelle, and D. Gratias, Physica A: Statistical Mechanics and its Applications **128**, 334 (1984).
[4] D. De Fontaine, in *Solid state physics* (Elsevier, 1994), vol. 47, pp. 33–176.
[5] C. Wolverton and A. Zunger, Physical Review B **50**, 10548 (1994).
[6] J. Connolly and A. Williams, Physical Review B **27**, 5169 (1983).
[7] K. Gubaev, E. V. Podryabinkin, G. L. Hart, and A. V. Shapeev, Computational Materials Science **156**, 148 (2019).
[8] V. I. Anisimov, F. Aryasetiawan, and A. Lichtenstein, Journal of Physics: Condensed Matter **9**, 767 (1997).
[9] B. Meredig, A. Thompson, H. Hansen, C. Wolverton, and A. Van de Walle, Physical Review B **82**, 195128 (2010).