On the many-body nature of intramolecular forces in FFLUX and its implications

Anton Konovalov1,2 | Benjamin C.B. Symons1,2 | Paul L.A. Popelier1,2

1Manchester Institute of Biotechnology (MIB), Manchester, United Kingdom
2Department of Chemistry, University of Manchester, Manchester, United Kingdom

Correspondence
Paul L.A. Popelier, Manchester Institute of Biotechnology (MIB), 131 Princess Street, Manchester M1 7DN, United Kingdom.
Email: pla@manchester.ac.uk

Funding information
Engineering and Physical Sciences Research Council, Grant/Award Number: EP/K005472

Abstract
FFLUX is a biomolecular force field under construction, based on Quantum Chemical Topology (QCT) and machine learning (kriging), with a minimalistic and physically motivated design. A detailed analysis of the forces within the kriging models as treated in FFLUX is presented, taking as a test example a liquid water model. The energies of topological atoms are modeled as $3N_{\text{atoms}}-6$ dimensional potential energy surfaces, using atomic local frames to represent the internal degrees of freedom. As a result, the forces within the kriging models in FFLUX are inherently N-body in nature where $N$ refers to $N_{\text{atoms}}$. This provides a fuller picture that is closer to a true quantum mechanical representation of interactions between atoms. The presented computational example quantitatively showcases the non-negligible (as much as 9%) three-body nature of bonded forces and angular forces in a water molecule. We discuss the practical impact on the pressure calculation with N-body forces and periodic boundary conditions (PBC) in molecular dynamics, as opposed to classical force fields with two-body forces. The equivalence between the PBC-related correction terms in the general virial equation is shown mathematically.

KEYWORDS
FFLUX, force field, kriging, machine learning, QTAIM, quantum chemical topology, water

1 | INTRODUCTION

Correctly modeling atomic interactions in order to predict the macroscopic behavior of a molecular system is the fundamental goal of atomistic force fields. Typically, the quantum mechanical description of these interactions is substituted for fully empirical or physically motivated approximate potential functions. This substitution is what gives force fields an enormous computational advantage compared to ab initio methods. However, this reduction in computational cost generally comes at a cost to accuracy. Nevertheless, force fields are the workhorse in biomolecular simulations at thermodynamically meaningful timescales.\(^1\) The deficiencies in accuracy of conventional pairwise additive force fields have been demonstrated several times.\(^2\)-\(^6\) Indeed, careful analysis of the results obtained for small molecules revealed the deficiencies from the early days of force field design.\(^7\)-\(^10\)

In response to these deficiencies, the re-introduction of many-body effects became a focus in both biomolecular force fields (see References 11 and 12 for a rich, critical and recent perspective) and water models, which are crucial components of biomolecular simulation. Typical improvements include the addition of multiple charge sites, inclusion of low order multipole moments and polarization, as a means of accounting for the effect of the surroundings.\(^11,13\) For completeness, we also mention the many-body effects on dispersion. As
discussed in a recent article, such effects may result in variations of atomic and molecular dispersion coefficients of leading order ($r^{-6}$) on the order of 10%.

In liquid water modeling, notable advances were made in using accurate ab initio data for the parameterization of analytical potential energy surfaces to capture the many-body effects, combined with polarization at longer range. Models such as the CC-pol family, WHBB and the state-of-the-art model MB-pol contain terms that explicitly model short-range two- and three-body interactions between molecules. Furthermore, the polarizable AMOEBA+ potential now also has geometry-dependent charge flux, a feature already present in FFLUX (see below) for a number of years.

Such advances in pure water modeling are in stark contrast with biomolecular simulations more generally, where the preference in practice is still given to simpler biomolecular force fields and water models. The focus in protein force fields is still on adjusting the existing potentials and inclusion of polarization. These apparent differences are due to a large extent due to higher computational cost in biomolecular modeling, shifting the cost-accuracy trade-off. Moreover, compatibility of sophisticated models with existing biomolecular force fields, or alternatively, applicability of these approaches to larger molecules remains an unsolved issue. A realistic, computationally efficient and widely applicable force field is yet to surface and many challenges remain.

Machine learning could assist in solving the aforementioned issues, boosted by its recent commercial success in areas such as computer vision and speech recognition. If computational chemistry adopts machine learning, it will take better advantage of the high-performance hardware that is increasingly available to it. Machine learning promises to be able to efficiently model multidimensional functions in many complex and useful scenarios. In the topological force field FFLUX, a machine learning technique called kriging (also known as Gaussian process regression [GPR]) is employed to predict atomic properties as functions of atomic coordinates.

FFLUX is a force field that incorporates Quantum Chemical Topology (QCT) in order to quantum mechanically define an atom in a molecule. An advantage of QCT is that atomic properties are obtained by a universal formula, that is, a volume integral of a property density over a (finite) atomic volume. Hence, an atom’s internal energy, as well as its charge, dipole moment and higher multipole moment, are all obtained in a uniform and consistent way. Machine learning is used to accurately capture these atomic properties and to make predictions efficiently, using trained models, during a molecular dynamics simulation.

The goal of the current article is to analyze the intramolecular forces in the FFLUX framework and clearly highlight the prominent features of our approach to these interactions that differentiate it from conventional approaches and latest developments. Note that the term “intramolecular” should be clarified because FFLUX handles both intra- and interatomic forces within a given molecule. Secondly, it should be understood that the whole approach applies not only to molecules but also to molecular clusters although this belongs to a future project.

2 | METHODS

2.1 | Quantum topological energy partitioning

The Quantum Theory of Atoms in Molecules (QTAIM) is a particular case (i.e., that of the electron density or its Laplacian) of the more general QCT approach. QTAIM partitions the electron density into unambiguously defined subspaces called topological atoms. The key concept to achieve this partitioning is the gradient vector, which naturally carves out these space-filling atoms: they do not overlap and leave no gaps between them. This space-filling property has consequences that have not yet been fully assessed (e.g., protein–ligand interaction in drug design) and also has conceptual advantages at short-range. Topological atoms are defined as subspaces with a vanishing Laplacian of the electron density over their volume. It is easy to prove that topological atoms are also quantum atoms, which are atoms that possess a well-defined kinetic energy.

Topological atoms play the role of “information carriers” in FFLUX. Energies of molecular entities are calculated as the sum of energies of their constituent topological atoms, with forces arising from the interactions between all atoms in a system. A method called Interacting Quantum Atoms (IQA) is a topological energy partitioning scheme based on the QTAIM framework and is thus part of QCT. IQA partitions the total energy of a system into atomic energies, that is both intra- and interatomic, without ever invoking a(n) (artificial) reference state as is done in traditional energy decomposition analyses:

$$E_{\text{TOT}} = \sum_{A} E_{\text{QTA}}^{A} = \sum_{A} E_{\text{intra}}^{A} + \frac{1}{2} \sum_{A} V_{\text{inter}}^{AA} = \sum_{A} E_{\text{intra}}^{A} + \frac{1}{2} \sum_{B \neq A} \sum_{N} V_{\text{inter}}^{AB}$$

(1)

Equation (1) shows three levels of resolution (i.e., coarse versus fine) in terms of atomic localization and energy type. Firstly, the total energy is the sum of individual atomic energies $E_{\text{QTA}}^{A}$. This simple addition is a clean procedure as a consequence of the space-filling nature of topological atoms, which are well-defined at short-range. Secondly, these atomic energies can be decomposed into intra-atomic energies ($E_{\text{intra}}^{A}$) and interatomic energies ($V_{\text{inter}}^{AA}$), where $A'$ refers to all atoms in the system except atom $A$. At this resolution the partitioning is centred on any given atom $A$, and hence specific atom-atom interactions are not known. Of course it is possible to go beyond this resolution (i.e., towards a third level) and introduce $V_{\text{inter}}^{AB}$ as explicit interatomic potential energies. IQA also allows the contributions to be broken down even further, not shown in Equation (1), namely into kinetic, exchange-correlation (which can be split in exchange and correlation) and electrostatic components. This finer partitioning provides more chemical insight but at a computational cost. However, for the purpose of constructing a force field, the partitioning into individual atomic energies $E_{\text{QTA}}^{A}$ is sufficient. Indeed, any energy terms from a finer partitioning would be summed anyway by virtue of calculating the atomic energies, which is the only energy needed in geometry optimization or molecular dynamics simulations.
2.2 | Learning the energies and multipole moments with kriging

Kriging is a non-parametric machine learning technique at the heart \(^{[24]}\) of FFLUX. Kriging applies Gaussian processes to establish a mapping between a collection of inputs (called features) and a single output. In this case, kriging links the geometry of a collection of atoms to their atomic energies and multipole moments, for previously unseen geometries. Features are defined with respect to a so-called atomic local frame (ALF).\(^{[37]}\) As a result, features are independent of global translation and rotation of molecules, which prevents the kriging models unnecessarily learning the purely mathematical and well-known transformation between each ALF and the global frame. As such, the quantities of interest are modeled as functions of 3N-6 input features (for non-linear molecules), where N is the number of atoms in the modeled entity. A more detailed description of the ALF is presented in the next Section.

Kriging treats outputs (atomic energies or multipole moments) as Gaussian distributions around a predicted mean value. The interpolation between training samples is carried out by conditioning a prior probability matrix, given by the prior mean and form of the covariance matrix, on the observed output values of the training samples. The approach exploits the idea of spatial correlation between output values as dictated by the covariance matrix. The covariance matrix is the central object in kriging; it expresses the prior uncertainty of the distribution as well the correlation between output values at different inputs. The entries of the covariance matrix are determined by the so-called kernel function, which sets the general shape of the fitted curves. Equation (2) is used in FFLUX to make predictions:

\[
\hat{Y}(\mathbf{f}') = \mu + \sum_{j=1}^{N_u} w_j \exp \left( -\sum_{k=1}^{N_u} \theta_k | f_{k,j} - f_{k,j}' |^{\rho_k} \right) \tag{2}
\]

where \(\hat{Y}(\mathbf{f}')\) is the predicted output of the function (atomic energy or multipole moment), \(\mathbf{f}'\) is the feature vector of interest (not present in the training set), \(\mu\) is the mean output of the training samples, \(w_j\) is the "weight" of a given training point \(j\), the number of training points is \(N_u\), the number of features is \(N_{\text{feat}}\), while \(f_{k,j}\) and \(f_{k,j}'\) are the features of a \(N_{\text{feat}}\)-dimensional training sample vector \(\mathbf{f}\) and the vector \(\mathbf{f}'\), respectively. In Equation (2), the kernel function is the exponential term that multiplies the weight \(w_j\). The symbols \(\theta_k\) and \(\rho_k\) refer to the hyperparameters of the kernel being optimized during training. The hyperparameter \(\theta_k\) determines the importance of the dimension \(k\), while \(\rho_k\) determines the smoothness of the function variation along the given dimension. In the current version of FFLUX we set \(\rho_k = 2\), turning the kernel into a widely-used squared exponential kernel.\(^{[24]}\) It should be noted that the weights \(w_j\) are actually functions of the covariance matrix and consequently of \(\theta_k\) and \(\rho_k\), and are also calculated during the construction of a model.\(^{[23]}\) This is why the term "weights" was put in quotes the first time it appeared, to take away the false impression that they were somehow adjustable.

A general framework for production of kriging models\(^{[38]}\) consists of a number of steps, starting with the preparation of the training and test sets. The wavefunction for the training geometries are then calculated (GAUSSIAN99\(^{[39]}\)), and the atomic IQA energies and multipole moments are then produced from the wavefunctions (using AIMALL\(^{[40]}\)). The kriging models are then trained, after which they are assessed by a geometry optimization and by so-called S-curves, which are sigmoidal curves that fully display the accuracy of a model as a cumulative distribution of the prediction errors for all training geometries.

2.3 | Training features

In its spirit, the ALF is similar to a Z-matrix, where the internal coordinate representation consists of bond lengths, bond angles and dihedral angles, thereby fully specifying the 3N-6 (or 3N-5 in the rare case of linear molecules) internal degrees of freedom. The differences lie in the details of how the axes and angles are defined.

Each atom’s nucleus serves as the origin of its own ALF. The corresponding input features are used to model the energy and multipole moments of the central atom. Figure 1 illustrates the idea by depicting the ALF of a carbonyl carbon atom in N-methylacetamide. This approach has the added benefit of ensuring atomic multipole moments are defined at the origin of the coordinate frame.

The first ALF feature, representing an internal degree of freedom, is the distance \(R_{\text{ALF}}^2\) (Equation [4]) from the central atom \(\Omega_0\) to the bonded atom with the highest priority \(\Omega_x\), according to Cahn–Ingold–Prelog rules, which form the basis of prioritization for all atoms. This bond also fixes the first local coordinate axis, \(x_{\text{ALF}}\). A third atom \(\Omega_{xy}\) (which is nitrogen in the example of Figure 1) is now assigned by the rules, in order to fix the ALF completely. The second feature is the distance \(R_{\text{ALF}}^2\) from the central atom \(\Omega_0\) to the next-highest-priority bonded atom \(\Omega_{xy}\). The third feature is the angle \(\chi_{\text{ALF}}\), which is subtended by \(\Omega_{xy}\) and \(\Omega_x\) from \(\Omega_0\). The three aforementioned atoms form

FIGURE 1 The atomic local frame (ALF) of the carbonyl carbon atom in N-methylacetamide [Color figure can be viewed at wileyonlinelibrary.com]
a xy-plane. The ALF’s z-axis can then be defined as the axis through the origin and perpendicular to xy-plane, forming a right-handed coordinate system. Positions of other atoms are described by spherical polar coordinates \((R_{\Omega}^{\alpha}, \phi_{\Omega}^{\alpha}, \theta_{\Omega}^{\alpha})\) defined with respect to the ALF. The list of the feature definitions\(^{[38]}\) is as follows:

\[
\begin{align*}
R_{\Omega}^{\alpha} &= R^{\alpha} - R^{\Omega} \\
R_{\Omega}^{\alpha} &= \sqrt{(R_{\Omega}^{\alpha})^2} \\
\chi^{\alpha} &= \arccos \left( \frac{R_{\Omega}^{\alpha} R_{\Omega}^{\Omega} - R^{\Omega}}{R_{\Omega}^{\Omega} R^{\Omega}} \right) \\
R_{\Omega}^{\alpha} &= \sqrt{(R_{\Omega}^{\alpha})^2} \\
\phi^{\alpha} &= \frac{\Delta_{\Omega}}{R_{\Omega}^{\alpha}} \\
\theta^{\alpha} &= \frac{\Delta_{\Omega}}{R_{\Omega}^{\alpha}}
\end{align*}
\]

where \(R^{\alpha}\) is the global position vector of atom \(n\), and \(\chi^{\alpha}\) \((i = 1, 2, 3)\) are local Cartesian coordinates \((x, y, z\) respectively) of atom \(n\) expressed in the ALF.

### 2.4 Analytic forces

The intramolecular forces are calculated as negative gradients of predicted \(E_{\text{IQA}}^A\) energy values with respect to displacement. As shown in previous work,\(^{[38]}\) since the energy is parameterized by the ALF features, the practical definition is the rightmost expression in Equation (10),

\[
F^{\Omega}_{l} = -\frac{\partial E_{\text{TOT}}}{\partial q^{\Omega}_{l}} = -\sum_{A=1}^{N_{\text{atom}}} \frac{\partial E_{\text{IQA}}^{A}}{\partial q^{\Omega}_{l}} - \sum_{A=1}^{N_{\text{atom}}} \sum_{k=1}^{N_{\text{feat}}} \frac{\partial E_{\text{IQA}}^{A}}{\partial q^{k}} \frac{\partial q^{k}}{\partial q^{\Omega}_{l}}
\]

where \(F^{\Omega}_{l}\) is the \(l\)th \((x, y, z\) force component force on the atom \(\Omega\) while \(q^{k}\) is the global \(k\)th coordinate of that atom. The quantity \(E_{\text{TOT}}\) is the molecular energy, which is the sum over \(N_{\text{atom}}\) terms of \(E_{\text{IQA}}^{A}\), that is, the predicted (marked by the hat) atomic energies. The quantity \(F^{\Omega}_{l}\) is the \(l\)th feature of atom \(A\), exerting a force on atom \(\Omega\). Both partial derivatives were previously obtained analytically.\(^{[38,41]}\)

\[
\frac{\partial E_{\text{IQA}}^{A}}{\partial q^{k}} = \sum_{j=1}^{N_{\text{atom}}} w_{j} (\text{sign} \gamma) \left( \frac{\partial_{k}^{j}}{\partial_{k}^{A}} - R_{\Omega}^{\alpha} \Delta_{\Omega} \right) \exp \left( -\sum_{h=1}^{N_{\text{atom}}} \frac{\partial_{h}^{j}}{\partial_{h}^{A}} \left| r_{h}^{j} - r_{h}^{A} \right|^{\gamma} \right)
\]

### 3 Interpretation of the forces

The focus of the present article is on the nature of the forces resulting from the kriged atomic energy surfaces, since this approach differs in three significant ways from classical force fields: (a) the use of QCT, (b) features defined in the ALF, and (c) N-body forces. We now discuss these points in turn and at length. First, the QCT picture of a molecule is that of a union of well-defined, space-filling quantum mechanical subspaces. These subspaces are topological atoms and stand in sharp contrast to the classical definition of atoms in typical force fields. Instead of being seen as a point mass, a topological atom has an internal structure and its own rigorously defined internal energy. Consequently, the origin of each force contribution in FFLUX can be traced back to energy gradients of individual atoms. This is not the case in classical force fields where the energies are typically viewed as belonging to "interactions," for example, energies "stored" in bonds, without the possibility for a full and true atomistic view.

Let us take the water molecule as an example and write out the terms on the right-hand side of Equation (10) while omitting the hat for brevity from now on and focusing on the x-component only. We restrict ourselves to two features only, \(R^{01H2}\) and \(R^{H2O1}\), one of which is associated with \(O1\) and the other with \(H2\), respectively. As a result, only a handful of derivative terms from Equation (10) will remain: only the terms containing the aforementioned features. In other words, we isolate the partial forces, on each of the three atoms in water, purely due to changes in these two features (i.e., the \(O1H2\) distance), resembling the bond length-related forces in classical force fields:

\[
\begin{align*}
F_{01}^{O1 \ x} &= \left( \frac{\partial E_{\text{IQA}}^{A}}{\partial R^{01H2}} \frac{\partial R^{01H2}}{\partial \Omega_{x}} \right) - \left( \frac{\partial E_{\text{IQA}}^{A}}{\partial R^{H2O1}} \frac{\partial R^{H2O1}}{\partial \Omega_{x}} \right) \\
F_{H2}^{O1 \ x} &= \left( \frac{\partial E_{\text{IQA}}^{A}}{\partial R^{01H2}} \frac{\partial R^{01H2}}{\partial \Omega_{x}} \right) - \left( \frac{\partial E_{\text{IQA}}^{A}}{\partial R^{H2O1}} \frac{\partial R^{H2O1}}{\partial \Omega_{x}} \right)
\end{align*}
\]
where \( \mathbf{F}^{\text{O1}}_{i,x} \) is the x-component of force exerted on O1, but only due to a change in the features \( \mathbf{R}^{\text{O1H2}} \) and \( \mathbf{R}^{\text{O2H2}} \), and consequently only due to energy changes of O1 and H2 because the energy of H3 does not depend on either of these two features. \( \mathbf{R}^{12} \) was used in the subscript of the force to refer to both features collectively. One should also notice that the partial force on H3 (\( \mathbf{F}^{\text{H3}}_{i,x} \) in Equation (18)), caused by a change in \( \mathbf{R}^{\text{O1H2}} \) and \( \mathbf{R}^{\text{O2H2}} \), is zero because \( \partial \mathbf{R}_{\text{O1H2}} / \partial \mathbf{x} = \partial \mathbf{R}_{\text{O2H2}} / \partial \mathbf{x} = 0 \). The latter equation follows from the definitions of the features because \( \mathbf{R}^{\text{O1H2}} = \mathbf{R}^{\text{O2H2}} = [\mathbf{R}^{\text{O1}} - \mathbf{R}^{\text{O2}}] \), that is, H3 does not participate in either of these features although formally written out in Equation (18). A novel consequence of FFLUX’s force treatment is that the partial forces \( \mathbf{F}^{\text{O1}}_{i,x} \) and \( \mathbf{F}^{\text{H2}}_{i,x} \) clearly consist of two contributions: an energy gradient of O1 and one of H2. The “self” force term, representing the energy change of an atom under its own translation, is completely natural from the point of view of a topological atom, as explained above. Yet this idea is usually obscured by thinking of energies as belonging to interactions, not atoms.

The second significant way FFLUX differs from classical force fields is the definition of internal degrees of freedom with respect to an ALF centered on the modeled atom makes the force description origin-atom-centric. Local molecular features such as bond lengths and bond angles are replaced by the ALF’s features centered on the origin-atom, relative to which the molecular geometry is described. This framework discards the top-down imposition of the Lewis diagram, in favor of a bottom-up, complete geometric picture that is properly consistent with an N-body system. This framework is minimal and sufficient as it describes completely the geometrical relationship between the central atom and every other atom without imposing any prior beliefs about these relationships. In this framework, the modeled atom has features relating it to all other atoms and it actively participates in all of them. The derivatives of its features with respect to its coordinates are not zero

\[
\frac{\partial \mathbf{f}_k}{\partial \mathbf{x}} \neq 0 \text{ for all } k, 0 < k < N_{\text{feat}} \text{ and for all } \Omega, 0 < \Omega < N_{\text{atoms}} \tag{19}
\]

If we had Lewis-diagram-like features instead, then the atom under consideration would have some vanishing \( \partial \mathbf{f}_k / \partial \mathbf{x} \) values. However, an ALF avoids this, and thus all features are dependent on ALF’s central atom.

Finally, the third significant way FFLUX differs from classical force fields is that kriging successfully captures the influence of each and every intramolecular degree of freedom on a single potential energy surface (PES) that has a quantum mechanical origin (i.e., kinetic, electrostatic, exchange-correlation energy). The energies of topological atoms are functions of all internal degrees of freedom, defined with respect to an ALF. In principle, a well-sampled kriging model allows reconstruction of the PES in the chemically meaningful part of the configuration space at ab initio accuracy (at the level of theory of the training samples). This goal has been realized in a number of previous publications.\[30,38,42,43\] Also, the number of training samples that kriging requires in order to achieve a certain level of accuracy was shown to be small relative to neural networks.\[44\] The main consequence of having a single PES is that there are no separate, uncoupled forces: all forces are N-body in nature.

This fact is worth emphasizing, as it is an appealing feature of the FFLUX force field that overhauls the idea of parametrizing separate interactions independently (e.g., bond, angular, dihedral and non-bonded between the distant atoms) as classical force fields do. For example, “bond forces” typically depend only on bond lengths and hence only on the coordinates of the two bonded atoms. This interaction forms its own partial and thus incomplete PES with a single internal degree of freedom as an input, and the interaction energy as an output (Figure 2). In this case, the total energy of the molecule is viewed as a collection of independent energy surfaces grouped by interaction type (e.g., bond, angle ...). This top-down, Lewis-diagram-centric view of a molecule may (or may not) capture the main interactions. However, it is clear that the positions of the surrounding atoms also have a non-zero effect on the magnitudes of the forces, consistent with the quantum mechanical view of a single wavefunction describing the whole system.

This non-zero effect was already expressed in the first classical force fields\[45,46\] as coupling terms (e.g., between bond length and bond angle energies), partially because they were originally derived to account for vibrational frequencies.\[47,48\] Many popular force fields later removed interactions in order to simplify the force field so they could be used on early-generation computers, for large molecules such as proteins. It has long been known that these approximations are not the complete picture\[49-51\] and were initially made out of necessity. However, advances in computational capabilities have allowed for the re-introduction of complexity in classical force fields in the form of cross-terms (e.g., describing a coupling between bond length and bond angle energies). However, FFLUX has no need for the addition of extra terms to make up for the insufficiency of the initial description. The full complexity of the system is inherent in

![FIGURE 2](https://wileyonlinelibrary.com)
FFLUX’s underlying architecture, which already provides an N-body description of the potential energy function. The forces in FFLUX can be collated with respect to given features, representing the internal degrees of freedom such as distances or angles, as shown in Equations (16)–(18). However, these forces retain an imprint of the remaining degrees of freedom, as will be explained below. Using the previous example, one can isolate the x-component of the force contribution related to $R_{O1H2}$ and $R_{O2H1}$ on each atom, see Equations (16)–(18). It was previously explained that $F_{E}^{x,x} = 0$ because $H3$ does not participate in either of these features. A different aspect of the same definition (i.e., $R_{O1H2} = R_{O2H1} = \chi_{O1}$) that we emphasize here is that $\partial E_{O1} / \partial x = -\partial E_{O2} / \partial x$ and also $\partial E_{H2} / \partial x = -\partial E_{H3} / \partial x$. This is mathematically trivial and can be understood by realizing that $O1$ and $H2$ are located on the opposite ends of the bond. This connection allows one to relate $F_{E}^{x,x}$ and $F_{H2}^{x,x}$ by recognizing similar terms. In fact, $F_{E}^{O1} = -F_{E}^{H2}$ or if expanded, using Equations (16) and (17), we obtain

$$
\left[ \frac{\partial E_{O1}}{\partial R_{O1H2}} \right] \frac{\partial R_{O1H2}}{\partial x} + \left[ \frac{\partial E_{H2}}{\partial R_{O1H2}} \right] \frac{\partial R_{O1H2}}{\partial x} = 0
$$

which is true at any point on the PES. The equality in Equation (20) is a manifestation of the familiar $F_{E} = -F_{E}$ (Newton’s third law). Now, while the forces related to $O1H2$ bond are exerted only on two atoms ($F_{E}^{x,x} = 0$) and behave like a two-body force, the atomic energies ($E_{O1}$ and $E_{H2}$), and thus $\frac{\partial E_{O1}}{\partial x}$ and $\frac{\partial E_{H2}}{\partial x}$, are functions of all features (and consequently all atoms), see the summation in Equation (11). In the example above, the force exerted on $O1$ through the $R_{O1H2}$ feature will implicitly depend on the position of $H3$ because the point on the PES at which the derivative is taken will change as a function of $H3$’s coordinates (see Equation (11)). The same arguments are true for all other features (including angular features).

Thus the features in FFLUX have either two-body (all radial features), three-body (angle $\chi_{Alk}$) or four-body character (polar angles) according to the number of atoms the forces are exerted on. However, all forces are implicitly N-body in nature, with the magnitudes varying due to displacements of other bodies (i.e. the remaining N-2; N-3; N-4 bodies). The Computational Example Section below gives a quantitative example of these effects.

## 4 | Computational Example

Here we present a computational example, showcasing the N-body nature of forces in a water molecule, as discussed earlier. The internal force contributions originating from the $E_{O1}$ energy gradient with respect to the $\chi_{O1}$ angular feature and $E_{H2}$ or $E_{H3}$ energy gradients with respect to $R_{O1H2}$ or $R_{O2H1}$ distance features were examined at two different O1H3 distances. While the values of the features under consideration do not change, the magnitudes of the forces are shown to be dependent on the position of $H3$. The forces were calculated in DL_FFLUX (modified DL_POLY 4.08) with a preliminary kriging model trained on 32 training points (molecular geometries), which was enough to produce a good quality PES. The reference energies were calculated at B3LYP/6-31+G(dp) level.

A single water molecule positioned in the xy-plane was examined using a distorted geometry with $R_{O1H2} = 1.00$ Å and $\chi_{O1} = 90\degree$. It was compared to a water molecule with $R_{O1H2} = 1.00$ Å, $R_{O1H3} = 0.95$ Å, and $\chi_{O1} = 90\degree$ were compared. The particular geometry was chosen for illustrative convenience only, and also indicates the ability of FFLUX to handle geometries far from the minimum energy. The comparison involved only the atoms participating in the examined features ($\chi_{O1}$ and $R_{O1H2}/R_{O2H1}$). The results are illustrated in Figure 3 and presented in full in Table 1.

The bond length of 1.00 Å is longer than the equilibrium distance at the level of theory of the training set (B3LYP/6-31+G(dp)), resulting in attraction between $O1$ and $H2$ through the corresponding feature ($R_{O1H2}/R_{O2H1}$). The opposite is true for the angle, which is smaller than that of the equilibrium geometry. From Table 1, the angular force magnitudes vary with $O1H3$ distance $\frac{\partial E_{O1}}{\partial x}$ decreased by $1.7\%$ at smaller $RO1H3$, thereby reducing the magnitudes of the forces contributions. However, $\frac{\partial E_{H2}}{\partial x}$ inversely proportional to the bond length, counteracted this effect on $H3$, overall increasing the force contribution on that atom by $3.5\%$ at shorter $O1H3$ bond length. The three-body effect played a more noticeable role in the $O1H2$ bonded forces. All the differences originate from the different positions on the three-dimensional PES of the two geometries. The $O1H2$ bond forces overall became $3.4\%$ more attractive due to the displacement of $H3$ (i.e. shorter $O1H3$ bond), which is non-negligible.

The QCT view allows us to look further into the atomic energy gradients. The energy gradient of $H2$ with respect to $R_{O2H1}$ increased by $9.4\%$ accompanied with a $0.4\%$ decrease in the $O1$ gradient. We hypothesize that such a large effect on $H2$ atomic energy gradient is related to its low electron density (absence of the core electron density), making it more sensitive to the surrounding.$^{53}$ An identical comparison of the force magnitudes at a less constrained angle ($\chi_{O1} = 100\degree$) showed analogous patterns and variations in magnitudes (Table 2).

Overall, the illustrated three-body effects identified from this comparison tended to be small, which is consistent with the intuition about N-body effects being “correction terms.” This observation may explain the reasonable success of force fields that isolate interactions as well as that of force fields accounting for many-body effects by the post hoc addition of approximate terms. In the latter case an approximate description of many-body effects may possibly suffice to improve results considerably. However, the rigorous theoretical grounding of the N-body effects in FFLUX puts it in better stead to expand to larger systems. It is expected that N-body effects, and therefore any errors in their description, will accumulate in more complex molecules. In fact, the latter revealed less than the careful work on small model compounds because the flaws in the force field were easier to tease out and correct in small model compound systems compared to in protein or oligopeptide simulations. This accumulation
may contribute to the lack of success of classical force fields when modeling large systems such as proteins.

5 | IMPLICATIONS OF MANY-BODY INTERACTIONS

While coming at increased computational cost compared to simpler force fields, our approach leverages a powerful machine learning technique for effective interpolation of multidimensional functions, and applies long-range multipolar electrostatics. The MPI parallelization of the code means that the in-house DL_FFLUX is closely comparable to the original DL_POLY (4.08) code in terms of performance, with any extra cost located in the prediction routines, as one would expect. In light of the ever increasing availability of computational resources, the onus is now on more accurate, physically motivated potentials. Water was chosen as a visually tractable and relevant example for this paper but, our approach is equally applicable to larger molecules in exactly the same way. Furthermore, the technique is not limited to single molecules but can and will be extended to molecular complexes. While not being fundamentally different from the QCT point of view—both molecules and molecular complexes are collections of atoms—it is more challenging in practice, due to the need to sample much larger regions of input space. An important example of such a system would be the water dimer (or trimer, etc.). Applying kriging to such a system would model both intra- and intermolecular degrees of freedom, successfully capturing the intermolecular N-body effects where N is now the number of molecules.

The example of MB-pol water model, which explicitly captures two- and three-body interactions at short-range based on

| $F_1^\Omega = -\frac{\partial^2 E}{\partial \alpha_\Omega \partial f_\Omega}$ | $R^{O1H3} = 1.00 \AA$ | $R^{O1H3} = 0.95 \AA$
|---|---|---|
| $\Omega$ | $A$ | $f_\Omega$ | $a_\Omega^I$ | $F_1^\Omega$ | $a_\Omega^I$ | $F_1^\Omega$ |
| O1 | O1 | $x$ | $x^{O1}$ | $-1.8024$ | $0.52918$ | $0.95378$ | $-1.7726$ | $0.55703$ | $0.98736$ |
| O1 | O1 | $y$ | $x^{O1}$ | $-1.8024$ | $0.52918$ | $0.95378$ | $-1.7726$ | $0.52918$ | $0.93800$ |
| H2 | O1 | $y$ | $x^{O1}$ | $-1.8024$ | $-0.52918$ | $-0.95378$ | $-1.7726$ | $-0.52918$ | $-0.93800$ |
| H3 | O1 | $y$ | $x^{O1}$ | $-1.8024$ | $-0.52918$ | $-0.95378$ | $-1.7726$ | $-0.55703$ | $-0.98736$ |
| O1 | H2 | $x$ | $R^{O1H2}$ | $1.9596$ | $1.0$ | $1.9596$ | $1.9516$ | $-1.0$ | $1.9516$ |
| O1 | H2 | $x$ | $R^{O1H2}$ | $1.2312$ | $-1.0$ | $1.2312$ | $1.3473$ | $-1.0$ | $1.3473$ |
| H2 | O1 | $x$ | $R^{O1H2}$ | $1.9596$ | $1.0$ | $-1.9596$ | $1.9516$ | $1.0$ | $-1.9516$ |
| H2 | O1 | $x$ | $R^{O1H2}$ | $1.2312$ | $1.0$ | $-1.2312$ | $1.3473$ | $1.0$ | $-1.3473$ |

Note: The entries are built up according to equation given in this table (see also Equation (10)). The complete list of eight non-zero components is presented, as determined by the alignment of the molecule with respect to the axis (e.g., all z-components are 0).
TABLE 2  Comparison of non-zero force contributions $F_i^j$ (in $10^{-2}$ atomic force units; to obtain the values in atomic units one needs to multiply them by 0.01) related to features $\chi^{O1}$, $R^{O1H2}$, and $R^{O1H3}$ at two different O1H3 distances and $\chi^{O1} = 100^\circ$

| $\Omega$ | $A$ | $A^i$ | $d^i$ | $R^{O1H3} = 1.00\text{Å}$ | $R^{O1H3} = 0.95\text{Å}$ |
|---------|-----|------|------|--------------------------|--------------------------|
| O1      | O1  | $\chi^{O1}$ | x    | $-1.7522$ 0.52114 0.91313 | $-1.7337$ 0.54856 0.95105 |
| O1      | O1  | $\chi^{O1}$ | y    | $-1.7522$ 0.6210 1.0882 | $-1.7337$ 0.62591 1.08514 |
| H2      | O1  | $\chi^{O1}$ | y    | $-1.7522$ -0.52918 -0.92722 | $-1.7337$ -0.52918 -0.91744 |
| H3      | O1  | $\chi^{O1}$ | x    | $-1.7522$ -0.52114 -0.91313 | $-1.7337$ -0.54856 -0.95105 |
| H3      | O1  | $\chi^{O1}$ | y    | $-1.7522$ -0.09189 -0.16101 | $-1.7337$ -0.09673 -0.16770 |
| O1      | O1  | $R^{O1H2}$ | x    | 1.9122 0.95 1.0 | 1.8691 1.0 1.0 |
| O1      | H2  | $R^{O1H2}$ | x    | 1.0138 1.0 | 1.0770 1.0 |
| H2      | O1  | $R^{O1H2}$ | x    | 1.9122 1.0 | 1.8691 1.0 |
| H2      | H2  | $R^{O1H2}$ | x    | 1.0138 1.0 | 1.0770 1.0 |

Note: The entries are built up according to the equation given in this table (see also Equation (10)). The complete list of the 9 non-zero components is presented, as determined by the alignment of the molecule with respect to the axis (e.g., all z-components are 0).

accurate ab initio reference data, has shown that such approaches lead to remarkable results.\[^{[18]}\] The advantage of kriging can be seen in its generality as a mathematical technique and effective interpolation. The PES of a single water molecule was previously shown to be nearly perfectly reproduced at the theory level of the training set with 50 training samples.\[^{[43]}\]

On a more practical note, while many-body forces for the most part do not result in significant algorithmic changes for MD simulations (apart from introducing the interaction-related code), the calculation of pressure is something that needs to be implemented correctly. The well-known virial equation for pressure for pairwise additive forces (Equation (21)) clearly does not apply if distance-dependent N-body interactions are present,

$$p = \frac{2(E_{\text{kin}})}{3V} + \frac{1}{6V} \left( \sum_{i<j}^{N} F_i \cdot r_j \right)$$ \tag{21}

Note also that forces dependent purely on angles and dihedrals can be shown to contribute nothing to the virial.\[^{[56]}\] In FFLUX, where the forces are related to features representing different types of degrees of freedom, it is convenient to have a general equation that accounts for everything in the same way. Alternatively, the pairwise virial with the minimal image convention could have been applied exclusively to radial features because within a single simulation frame the forces related to radial degrees of freedom act as two-body forces. Angular features could be then treated as purely angular within each simulation frame. However, following the route of the most general virial equation (Equation (22)) is a natural way to calculate the pressure in such a situation, and this equation is thus applicable with many-body interactions

$$W = \sum_{i}^{N} F_i \cdot r_i$$ \tag{22}

where $F_i$ is the resultant force on atom $i$ and $r_i$ is its global coordinate. However, a fact that appears to be somewhat overlooked in the literature is that this equation does not apply in the case of periodic boundary conditions (PBC).\[^{[57,58]}\] The difference arises from the need to correctly account for the interactions of the main cell with periodic images. Interestingly, Equation (21) still produces correct results even in the case of periodic boundary conditions if used with the minimum-image convention as already mentioned above. This observation and the prevalence of force fields with only pairwise additive and purely angular interactions is most likely the cause of the seemingly low awareness of the literature showcasing the problem with Equation (22), since it is rarely used in that form.

Louwerse et al. described\[^{[57]}\] the problem with Equation (22) and expressed the required pressure correction term in the case of a cubic simulation box, resulting in

$$p = \frac{2(E_{\text{kin}})}{3V} + \left( \frac{1}{6V} \sum_{i<j}^{N} F_i \cdot r_j - \frac{1}{3V^2} \frac{\partial U}{\partial L} \right)$$ \tag{23}

where $U$ is the potential energy per unit cell (equivalently of the main cell) and $L$ is the cubic box length. The correction term $-\frac{1}{3V^2} \frac{\partial U}{\partial L}$ expresses the dependence of $U$ on the box length, while keeping the global coordinates of the main cell atoms $r_i$ constant. On the other hand, a somewhat different route led Thompson et al.\[^{[58]}\] to several equivalent expressions for the virial in the presence of arbitrary many-body potentials and periodic boundary conditions. One of their expressions, Equation (24), closely resembles the corrected virial term in Equation (23), enabling them to conclude the equivalence of the expressions by straightforward comparison,

$$W = \sum_{i}^{N} F_i \cdot r_i + \sum_{n \in \Omega} H_n \sum_{i}^{N} F_i^w$$ \tag{24}

In the above equation, $H$ is the matrix of the cell vectors and $n$ represents the offset vector of the periodic cell and can be any vector with integer components. As explained by Thompson et al.\[^{[58]}\] $F_i^w$ can...
be thought of as a partial force on the periodic image of $i$ in the cell $n$ related to the main cell, as opposed to the total force on the periodic image atom $i$ in the cell $n$. As indicated by the authors, Equation (24) also has the form of the equation earlier derived by Bekker et al.\textsuperscript{59} We complement the presented ideas by providing a simple mathematical proof by explicit differentiation of $\left( \frac{\partial U}{\partial L} \right)_{\{r\}}$, that the correction terms in Equations (23) and (24) are indeed identical for cubic cells, leveraging the physical meaning of $\left( \frac{\partial U}{\partial L} \right)_{\{r\}}$ (see below). Application of the ideas as presented by Thompson et al.\textsuperscript{58} resulted in a well-behaved pressure of the yet unpublished FFLUX liquid water model with PBC.

Starting with the view presented in the paper by Louwerse et al.,\textsuperscript{57} the factor $\left( \frac{\partial U}{\partial r} \right)_{\{r\}}$ in Equation (23) is the change in potential energy per unit cell (e.g., the main cell) associated with the increase of the box size without changing the absolute coordinates of atoms in the main cell $\{r\}$ ($n = 0$, the zero vector). Physically, such expansion/contraction affects the potential energy of the main cell via the changes in interactions between the atoms inside the main cell and the image atoms. The insight, not explicitly presented in the papers,\textsuperscript{57,58} is to incorporate this physical meaning (change in $U$ due to change in $r_n$) mathematically into the equation:

\[
\left( \frac{\partial U}{\partial L} \right)_{\{r\}} = \sum_{n \in Z^3} \sum_{i \in n} \frac{\partial U}{\partial r_n} \frac{\partial r_n}{\partial L} - \sum_{n \in Z^3} \sum_{i \in n} F_n \cdot \frac{\partial X_n}{\partial L} \quad (25)
\]

The notation of Thompson et al.,\textsuperscript{58} Equation (24), is helpful in this situation: the double sum iterates over all atoms in all cells, $F_n$ is the partial force on the $n$-th atom associated with the main cell potential energy $U$, not the total force on that image atom. The vector $r_n$ can be expressed as a function of $r_i$ and $L$ if the cell is cubic, Equation (27), and the derivative of $r_i$ with respect to $L$ is then given by Equation (28),

\[
r_n = r_i + Hn \quad \text{(in general)}
\]

\[
r_n = r_i + Ln \quad \text{(for a cubic cell)}
\]

\[
\left( \frac{\partial r_n}{\partial L} \right)_{\{r\}} = Ln \quad \text{(for a cubic cell)}
\]

where $L$ is the identity matrix. One can clearly see that the atoms in the main cell ($n = 0$) are unaffected. After substituting Equation (28) into Equation (25) and using the fact that $V = L^3$ (for a cubic cell), the equivalence of the equations of interest becomes obvious, Equation (31):

\[
\left( \frac{\partial U}{\partial L} \right)_{\{r_i\}} = -\sum_{n \in Z^3} \sum_{i \in n} F_n \cdot Hn = -Hn \sum_{n \in Z^3} \sum_{i \in n} F_n 
\]

\[
L \left( \frac{\partial U}{\partial L} \right)_{\{r_i\}} = -\sum_{n \in Z^3} \sum_{i \in n} F_n \cdot Hn = -Hn \sum_{n \in Z^3} \sum_{i \in n} F_n
\]

\[
p = \frac{2(E_{\text{kin}})}{3V} + \frac{1}{3V} \sum_{i} F_i \cdot r_i - \frac{1}{2\epsilon^2} \left( \frac{\partial U}{\partial L} \right)_{\{r_i\}} \quad (31)
\]

6 | CONCLUSIONS

We have presented the analysis of the intramolecular interactions as viewed within the FFLUX framework. Its main characteristics, stemming from decisions on FFLUX's principal design, are the atomic view on molecules and the full N-body treatment of atoms in the model.

The IQA partitioning scheme allows for determination of reference-free atomic energies extracted from the molecular wavefunction. The resulting forces are fundamentally viewed as the changes of the energies of separate topological atoms. The internal degrees of freedom in FFLUX are defined with respect to an ALF and are origin-centric compared to typical force-field parameterization through bond lengths, bond angles and dihedral angles. This procedure expresses the dependence of atomic energies on the positions of every other atom (in the model) through features explicitly involving the origin-atom. These energies are learned as functions of 3N-6 internal coordinates represented within an atomic local frame (ALF). The consequence of having a single PES describing the energy is that the forces are also functions of all internal degrees of freedom. This description is fuller and much closer to the quantum mechanical picture of interactions.

The computational example showed numerically the presence of the three-body effects in water. In the particular example, the magnitudes of the energy gradients varied from 0.4 to 9.4% upon translation along internal degrees of freedom on the PES, other than the one studied. These small changes are non-negligible and we believe that these interactions are important for achieving higher accuracy and capturing finer details in molecular dynamics.

Application of these ideas might have even more significant effect on larger molecules and molecular complexes, enabling to accurately describe intra-and intermolecular interaction with the same approach. Research on such systems is currently work in progress.

In addition, calculation of pressure is discussed, advising caution when the pairwise virial is not applicable. Periodic boundary conditions require one to correctly account for the interactions of the atoms in the main cell with the image atoms. They effectively require the introduction of a correction term to the most general equation for the virial, as described in the literature. A short mathematical proof of equivalence between the PBC-related correction terms (suggested by Louwerse et al. and Thompson et al.) in the general virial equation is presented.

ACKNOWLEDGMENTS

We thank Matthew J. Burn for providing the kriging model. P.L.A.P. thanks EPRSC for the award of an Established Career Fellowship, grant EP/K005472.
REFERENCES

[1] P. S. Nerenberg, T. Head-Gordon, Curr. Opin. Struct. Biol. 2018, 49, 129.
[2] R. B. Best, J. Mittal, Proteins: Struct. Funct. Bioinformat. 2011, 79, 1318.
[3] K. Lindorff-Larsen, P. Maragakis, S. Plana, M. P. Eastwood, R. O. Dror, D. E. Shaw, Proc Natl Acad Sci USA 2012, 7, e32131.
[4] J. Henriques, C. Cragnell, M. Skepò, J. Chem. Theory Comput. 2015, 11, 3420.
[5] S. Rauscher, V. Gapsys, M. J. Gajda, M. Zweckstetter, B. L. de Groot, H. Grubmüller, J. Phys. Chem. B 2018, 122, 116.
[6] A. T. Hagler, J. Comp-Aided Mol. Des. 2019, 33, 133.
[7] J. Chem. Theory Comput. 2010, 6, 271.
[8] J. Chem. Theory Comput. 2013, 9, 154107.
[9] J. Chem. Theory Comput. 2015, 11, 5513.
[10] J. Chem. Theory Comput. 2011, 7, 455.
[11] J. Chem. Theory Comput. 2014, 10, 23.
[12] J. Chem. Theory Comput. 2016, 12, 1541.
[13] J. Chem. Theory Comput. 2017, 13, 6365.
[14] J. Comput. Chem. 2016, 37, 3830.
[15] J. Chem. Theor. Comput. 2015, 11, 5326.
[16] J. Chem. Theor. Comput. 2014, 10, 228.
[17] J. Chem. Theor. Comput. 2013, 9, 5395.
[18] J. Chem. Theor. Comput. 2014, 10, 1599.
[19] J. Comput. Chem. 2003, 24, 843.
[20] J. Comput. Chem. 2016, 37, e1347.
[21] J. Comput. Chem. 2015, 36, 5513.
[22] J. Comput. Chem. 2019, 40, 6119.
[23] J. Comput. Chem. 2017, 38, 419.
[24] J. Comput. Chem. 2015, 36, 3840.
[25] J. Comput. Chem. 2014, 35, 228.
[26] J. Comput. Chem. 2012, 33, 106.
[27] J. Comput. Chem. 2011, 32, 205.
[28] J. Comput. Chem. 2016, 37, 50.
[29] J. Comput. Chem. 2014, 35, 505.
[30] J. Comput. Chem. 2013, 34, 1850.
[31] J. Comput. Chem. 2012, 33, 1137.
[32] J. Comput. Chem. 2011, 32, 86.
[33] J. Comput. Chem. 2010, 31, 507.
[34] J. Comput. Chem. 2009, 30, 1234–1241.
[35] J. Comput. Chem. 2017, 38, 3553.
[36] J. Comput. Chem. 2013, 34, 419.
[37] J. Comput. Chem. 2018, 49, 9127.
[38] J. Comput. Chem. 2015, 36, 5327.
[39] J. Comput. Chem. 2019, 40, 6119.