Supporting Information: Development of Accurate Transferable Hydrofluorocarbon Refrigerant Force Fields Using a Machine Learning and Optimization Approach

Montana N. Carlozo, Ning Wang, Alexander W. Dowling,* and Edward J. Maginn*

Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame IN 46556, USA

E-mail: adowling@nd.edu; ed@nd.edu

S1 Atom Typing Schemes

This section details the atom typing schemes evaluated in this work, shows the preliminary results of AT-4 and AT-3 (atom types C, H, F), and also lists transformation matrices A_m.

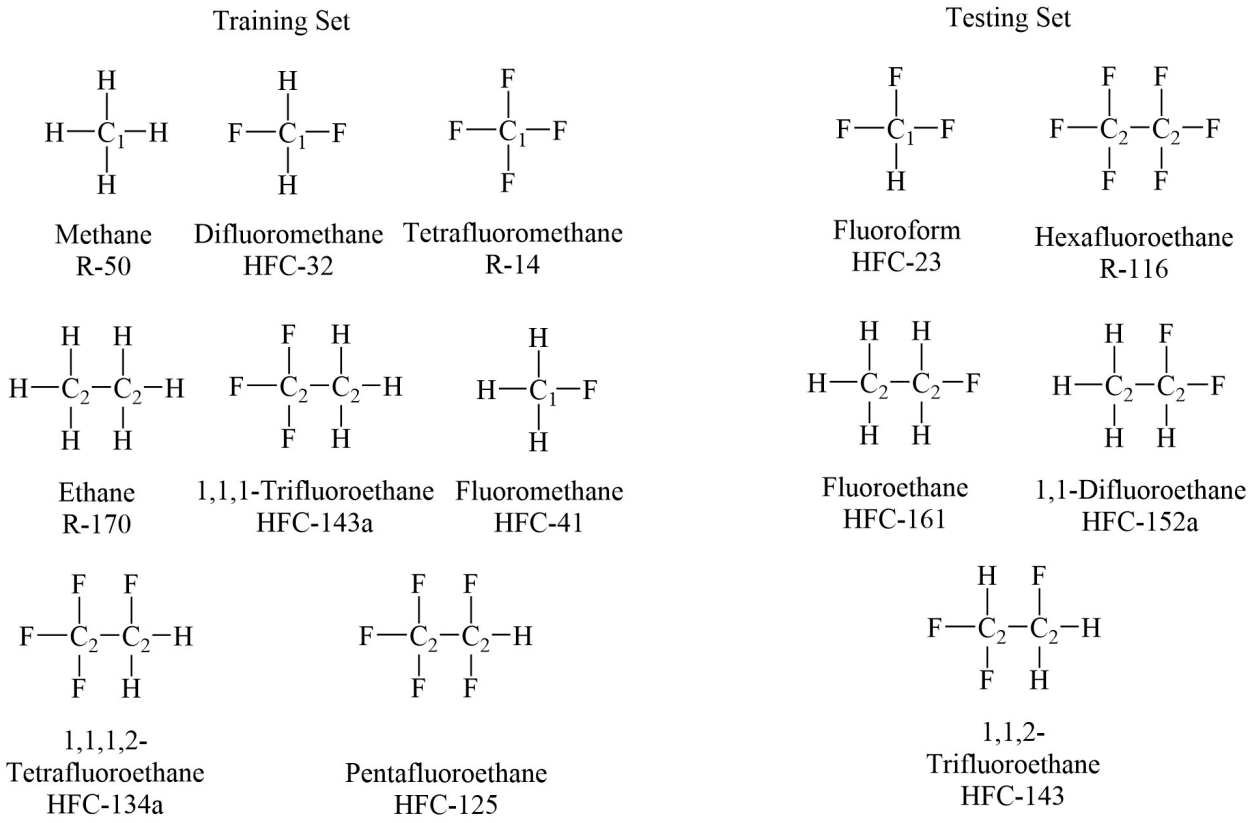

Figure S1: AT-4 uses the four most basic atom types (C_1, C_2, F, and H).

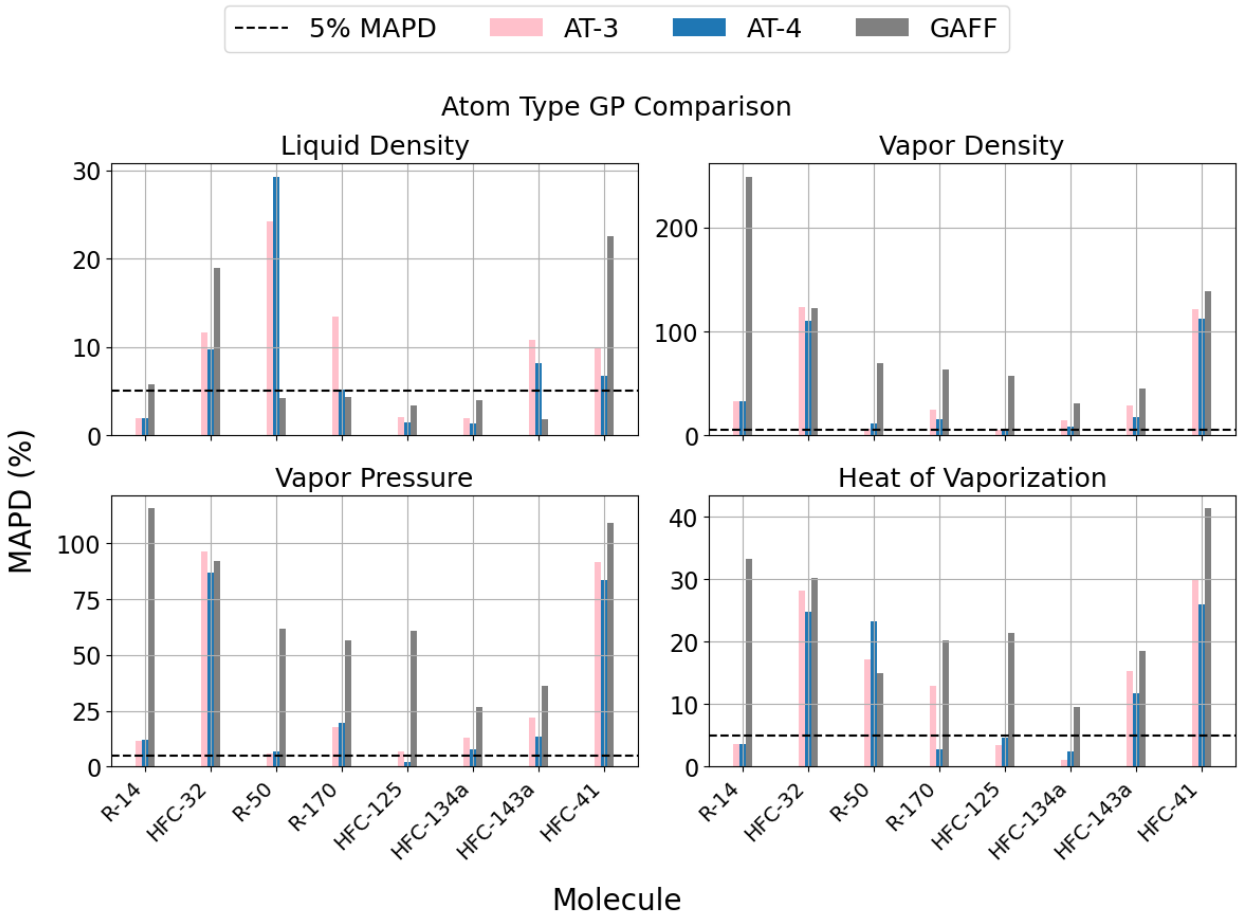

Figure S2: GP-predicted MAPD values for molecules in the training set for AT-3, AT-4, and GAFF.

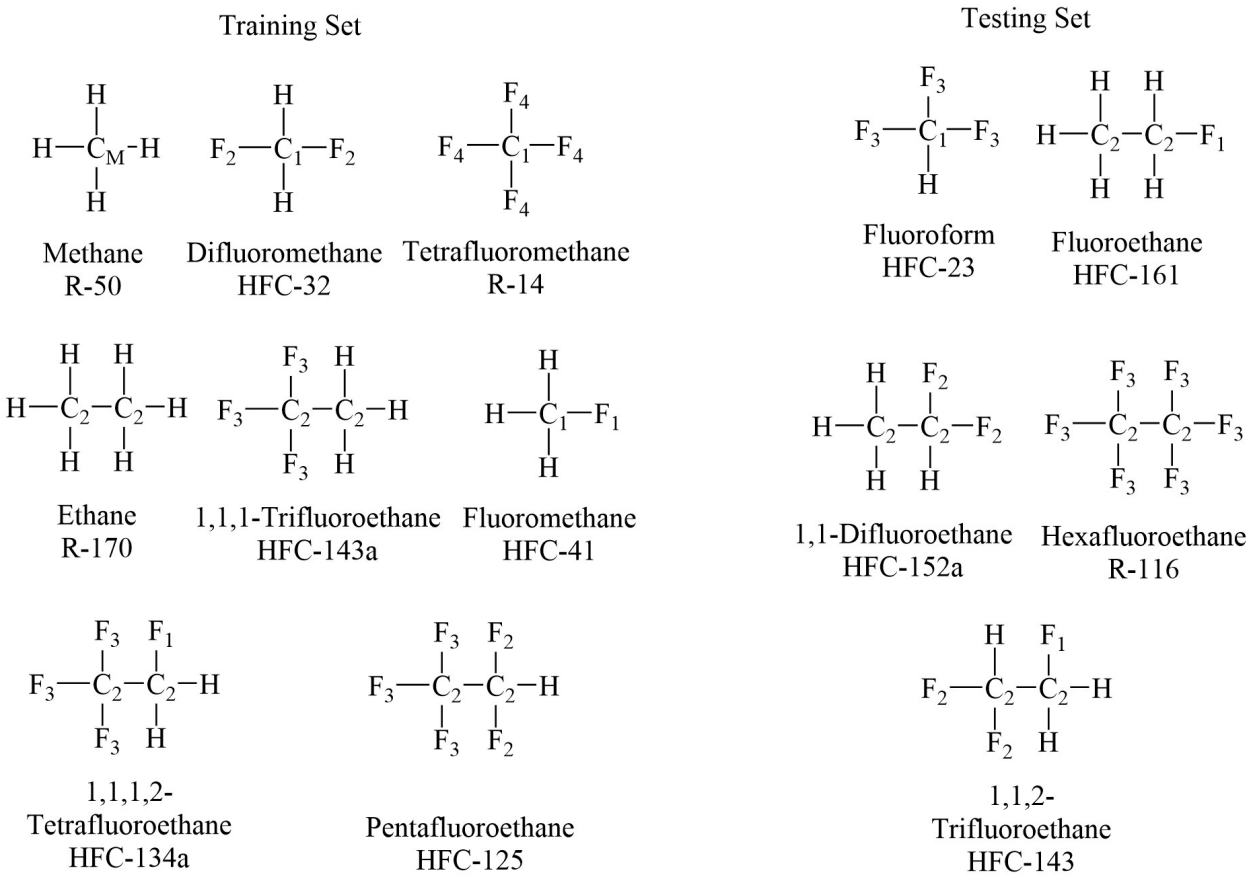

Figure S3: AT-8 uses eight atom types which were expected to fully represent all training molecules (C_M, C_1, C_2, F_1, F_2, F_3, F_4, and H).

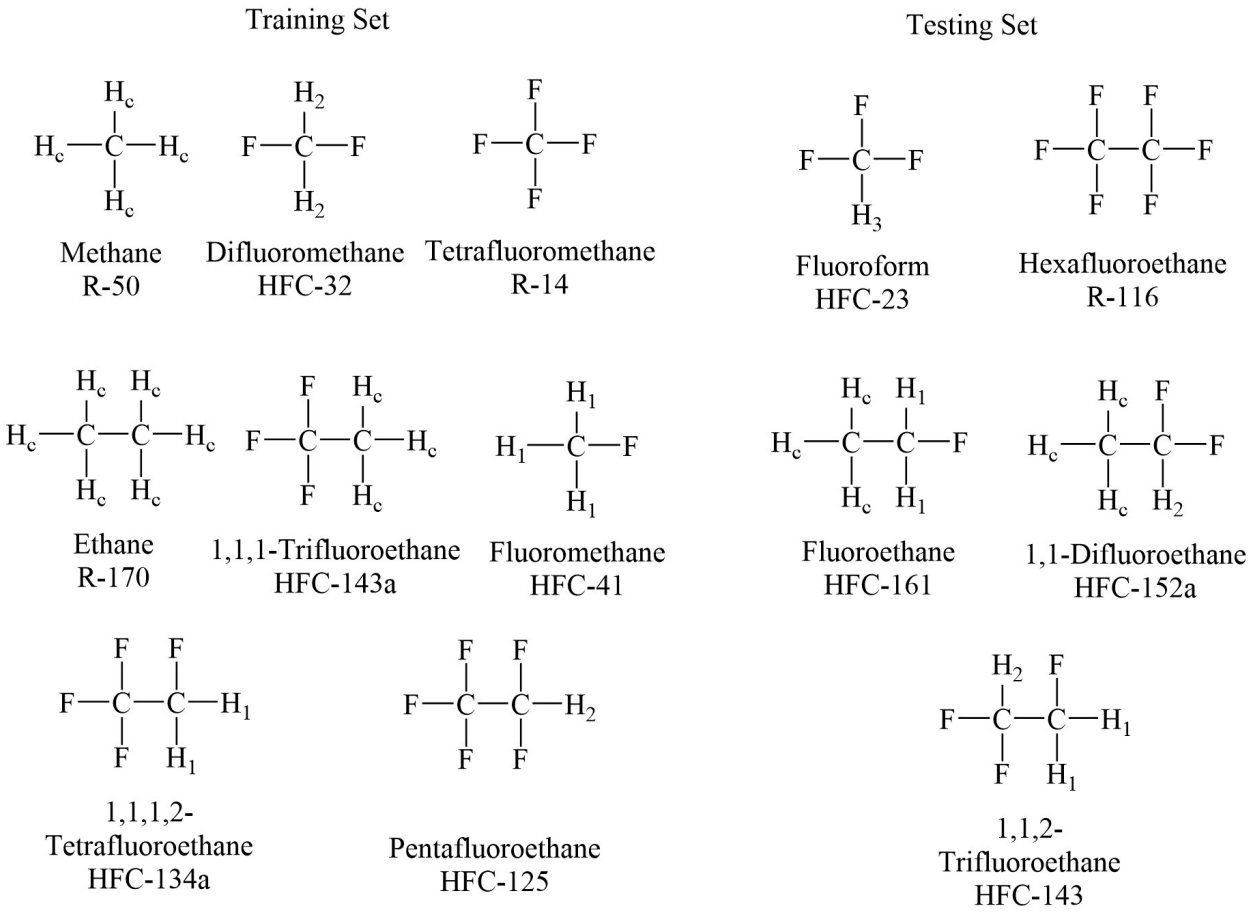

Figure S4: AT-6a uses six atom types from GAFF but re-optimizes the values of these parameters (C, H_c, H_1, H_2, H_3, and F).

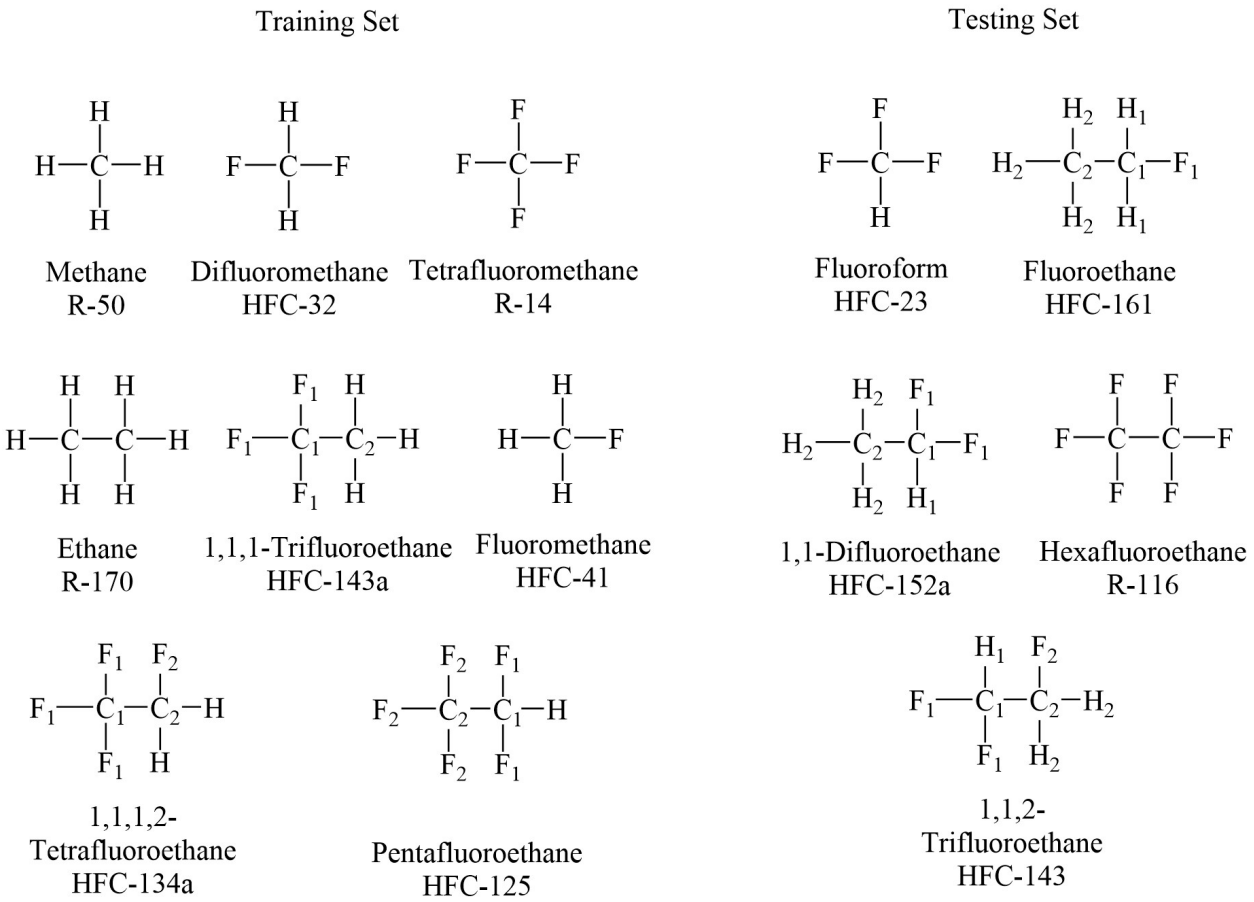

Figure S5: The distinct atom types of each molecule. The GPs for training molecules use these atom types as input features.

Table S1: Transformation matrices $\mathbf{A}_m$ for atom types listed in this work

Distinct Atom Types	AT-4	AT-6a	AT-6b	AT-8
R-14				
C	C_1	C	C_1	C_1
F	F	F	F	F_4
R-50				
C	C_1	C	C_m	C_m
H	H	H_c	H_c	H
HFC-32				

Distinct Atom Types	AT-4	AT-6a	AT-6b	AT-8
C	C_1	C	C_1	C_1
H	H	H_2	H	H
F	F	F	F	F_2
HFC-125				
C_1	C_2	C	C_2	C_2
C_2	C_2	C	C_2	C_2
H_1	H	H_2	H	H
F_1	F	F	F	F_2
F_2	F	F	F	F_3
HFC-134a				
C_1	C_2	C	C_2	C_2
C_2	C_2	C	C_2	C_2
H_1	H	H_1	H	H
F_1	F	F	F	F_3
F_2	F	F	F	F_1
HFC-143a				
C_1	C_2	C	C_2	C_2
C_2	C_2	C	C_2	C_2
H_1	H	H_c	H_c	H
F_1	F	F	F	F_3
R-170				
C	C_2	C	C_2	C_2
H	H	H_c	H_c	H
HFC-41				

Distinct Atom Types	AT-4	AT-6a	AT-6b	AT-8
C	C_1	C	C_1	C_1
H	H	H_1	H	H
F	F	F	F	F_1

S2 Justification of Heuristic q_j

Estimability exists on a spectrum and there multiple methods to quantify relative estimability. While graphical analysis is possible on small-scale models, numerical methods relying on the FIM and $\mathbf{F}$ are much more common.[1–5] For example, parameter correlation can be quantified by defining a correlation matrix[6,7] or collinearity index[8] which identifies linear dependencies in the columns of $\mathbf{F}$ instead of graphing it. However, estimability rankings drawn from $\mathbf{F}$ can be misleading. For example, correlation matrices[6,7] and collinearity indices[8,9] rely on arbitrary cutoffs to categorize parameters as estimable or inestimable and rely on heuristic-based scaling factors[6,9].[1,4]

Therefore, modelers will often investigate the eigen-decomposition of the FIM to determine if any specific linear combination of model parameters are consistently difficult to precisely estimate, i.e., are not practically identifiable. However, expert intuition is often required to inspect the eigen-decomposition and determine actionable modeling outcomes (e.g., fix parameters, adjust model equations). To help streamline this process, we introduce the following heuristic:

$$q_j = \sum_{i=1}^{P} \lambda_i \left| v_{i,j} \right| \tag{1}$$

Here, q_j approximates the contributions of parameter j to the information contained in the FIM. λ_i are the eigenvalues and $v_{i,j}$ are the corresponding eigenvectors.

From a linear algebra perspective, let the FIM be written

$$FIM = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}. \tag{2}$$

where λ_i, the diagonals of $\mathbf{\Lambda}$, are the eigenvalues of the FIM, $\mathbf{v}_i$ are the eigenvectors of the FIM, and $v_{i,j}\ \forall j \in J$ are the eigenvector components corresponding to each parameter. Each eigen-pair (λ_i, $\mathbf{v}_i$) represents an independent principal direction of information ($\mathbf{v}_i$) in the model with strength (λ_i). Note that larger eigenvalues are associated with more informative directions. It then follows that $v_{i,j}$ (and equally $|v_{i,j}|$) is a measure of how much a parameter j contributes to a principal information direction i.

Thus, the interpretation of q_j in eqn. 1 is an information-coordinate-wise participation score of parameter j which measures how much information parameter j contributes across all principal information directions. In this way, parameters consistently associated with larger eigenvalues are scored as more informative. This method is thus justified by its reasonable interpretation from a linear algebra perspective and lends itself nicely to the relative estimability ranking of parameters.

S3 Derivation of the $\mathbb{E}$[SSE] Objective Function

Consider a random vector for the error between our GP model and simulated data as

$$\mathbf{r}(\mathbf{T}, \boldsymbol{\theta}) = \left\{\mathbf{r}(\mathbf{T}, \boldsymbol{\theta}A_m^{\top})\right\}_{m=1}^{M} \sim \mathcal{N}\big(\boldsymbol{\mu}_r(\mathbf{T}, \boldsymbol{\theta}), \boldsymbol{\Sigma}(\mathbf{T}, \boldsymbol{\theta})\big) \tag{3}$$

where

$$\boldsymbol{\mu}_r(\mathbf{T},\boldsymbol{\theta}) = \begin{bmatrix} y^{\text{exp}}_{1,p,T} - \mu_{\text{GP},p}\big(\mathbf{T},\boldsymbol{\theta}A_1^\top\big) \\ y^{\text{exp}}_{2,p,T} - \mu_{\text{GP},p}\big(\mathbf{T},\boldsymbol{\theta}A_2^\top\big) \\ \vdots \\ y^{\text{exp}}_{M,p,T} - \mu_{\text{GP},p}\big(\mathbf{T},\boldsymbol{\theta}A_M^\top\big) \end{bmatrix} \tag{4}$$

and

$$\boldsymbol{\Sigma}(\mathbf{T},\boldsymbol{\theta})_{mm'} = \boldsymbol{\Sigma}\Big(\mathbf{T},\boldsymbol{\theta}A_m^\top\,;\;\mathbf{T},\boldsymbol{\theta}A_{m'}^\top\Big), \qquad m,m' = 1,\ldots,M. \tag{5}$$

We then define the full model of SSE loss as $\tilde{\mathscr{L}}(\mathbf{T},\boldsymbol{\theta}) = \mathbf{r}^\top(\mathbf{T},\boldsymbol{\theta})\mathbf{W}\mathbf{r}(\mathbf{T},\boldsymbol{\theta})$ where matrix $\mathbf{W}$ is a matrix of weights. We derive an analytical solution for $\mathbb{E}[\tilde{\mathscr{L}}(\mathbf{T},\boldsymbol{\theta})]$ ($\mathbb{E}$[SSE]) following the derivation by Rencher and Schaalje.[10] We begin by noting that since $\mathbf{r}(\mathbf{T},\boldsymbol{\theta})$ is a $1 \times N$ vector, that $\tilde{\mathscr{L}}(\mathbf{T},\boldsymbol{\theta})$ is a scalar. This allows us to compute $\mathbb{E}[\tilde{\mathscr{L}}(\mathbf{T},\boldsymbol{\theta})]$ by analytically computing the mean and covariance components of $\mathbf{r}(\mathbf{T},\boldsymbol{\theta})$ separately and noting that $\mathbb{E}[\mathbf{r}\mathbf{r}^\top] = \boldsymbol{\Sigma}(\mathbf{T},\boldsymbol{\theta}) + \boldsymbol{\mu}_r\boldsymbol{\mu}_r^\top$ where $\boldsymbol{\mu}_r$ holds regardless of whether predictions are correlated. This derivation also takes advantage of the properties of the trace function which allows us to factor out $\mathbf{W}$ and calculate the expected value of the unscaled loss function ($\mathbb{E}[\mathbf{r}\mathbf{r}^\top]$) analytically.

$$\mathbb{E}[\tilde{\mathscr{L}}(\mathbf{X}^{exp}, \boldsymbol{\theta})] = \mathbb{E}[\mathbf{r}^\top \mathbf{W}\mathbf{r}] = \mathbb{E}[\mathrm{Tr}(\mathbf{r}^\top \mathbf{W}\mathbf{r})] \tag{6a}$$

$$= \mathbb{E}[\mathrm{Tr}(\mathbf{W}\mathbf{r}\mathbf{r}^\top)] \tag{6b}$$

$$= \mathrm{Tr}(\mathbb{E}[\mathbf{W}\mathbf{r}\mathbf{r}^\top]) \tag{6c}$$

$$= \mathrm{Tr}(\mathbf{W}\mathbb{E}[\mathbf{r}\mathbf{r}^\top]) \tag{6d}$$

$$= \mathrm{Tr}(\mathbf{W}(\boldsymbol{\Sigma}(\mathbf{T}, \boldsymbol{\theta}) + \boldsymbol{\mu}_r \boldsymbol{\mu}_r^\top)) \tag{6e}$$

$$= \mathrm{Tr}(\mathbf{W}\boldsymbol{\Sigma}(\mathbf{T}, \boldsymbol{\theta}) + \mathbf{W}\boldsymbol{\mu}_r \boldsymbol{\mu}_r^\top) \tag{6f}$$

$$= \mathrm{Tr}(\mathbf{W}\boldsymbol{\Sigma}(\mathbf{T}, \boldsymbol{\theta})) + \mathrm{Tr}(\boldsymbol{\mu}_r^\top \mathbf{W}\boldsymbol{\mu}_r) \tag{6g}$$

$$= \mathrm{Tr}(\mathbf{W}\boldsymbol{\Sigma}(\mathbf{T}, \boldsymbol{\theta})) + \boldsymbol{\mu}_r^\top \mathbf{W}\boldsymbol{\mu}_r \tag{6h}$$

In this work we use the covariance matrix of the experimental measurements as weights such that $\mathbf{W} = \boldsymbol{\Sigma}_{\text{exp}}^{-1}$. Further, we assume that experimental measurements are independent across molecules and temperatures i.e., $\boldsymbol{\Sigma}_{\text{exp}}^{-1}$ is a diagonal matrix of the variances of each experimental measurement:

$$\boldsymbol{\Sigma}_{\text{exp}}^{-1} = \begin{pmatrix} \frac{1}{\sigma^2_{\text{exp},1,1,1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma^2_{\text{exp},2,2,2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma^2_{\text{exp},m,p,T}} \end{pmatrix} \tag{7}$$

Thus eqn. (6h) simplifies as follows through the properties of the trace of a matrix:

$$\mathbb{E}[\tilde{\mathscr{L}}(\mathbf{T}, \boldsymbol{\theta})] = \sum_{m \in \mathcal{M}} \sum_{p \in \mathcal{P}} \sum_{T \in \mathcal{T}} \frac{1}{\sigma^2_{\text{exp},m,p,T}} \left(y^{\text{exp}}_{m,p,T} - \mu_{\text{GP},p}(T, \boldsymbol{\theta}\mathbf{A}_m^\top)\right)^2 + \frac{\sigma^2_{\text{GP},p}(T, \boldsymbol{\theta}\mathbf{A}_m^\top)}{\sigma^2_{\text{exp},m,p,T}}. \tag{8}$$

S4 Derivation of the Approximation of the FIM Using the Hessian of the $\mathbb{E}$[SSE] Objective Function

From SI section S3 the $\mathbb{E}$[SSE] objective function in eqn. (27), $g(\boldsymbol{\theta})$, with respect to LJ parameters, $\boldsymbol{\theta}$, is:

$$g(\boldsymbol{\theta}) = \mathrm{Tr}({\boldsymbol{\Sigma}_{\mathrm{exp}}}^{-1}\boldsymbol{\Sigma}) + \boldsymbol{\mu}_r^\top \boldsymbol{\Sigma}_{\mathrm{exp}}^{-1}\boldsymbol{\mu}_r. \tag{9}$$

Note that the arguments of $\boldsymbol{\Sigma}(\mathbf{T}, \boldsymbol{\theta})$ and $\boldsymbol{\mu}_r(\mathbf{T}, \boldsymbol{\theta})$ are omitted for convenience.

Let us define, $\boldsymbol{\mu}_{\mathrm{GP}} = \mu_{\mathrm{GP};m,p,T}$ and $\boldsymbol{\Sigma}_{\mathrm{GP}} = \sigma^2_{\mathrm{GP};m,p,T}\ \forall\ n \coloneqq \{m, p, T\} \in N$ where $N \coloneqq \{\mathcal{M}, \mathcal{P}, \mathcal{T}\}$. If we note that because $\mathbf{y}^{\mathrm{exp}}$ and the elements of ${\boldsymbol{\Sigma}_{\mathrm{exp}}}^{-1}$ are constants, the first derivative of the residuals with respect to the LJ parameters are equal to those of the GP for a given molecule and property such that $\nabla\mathbf{r}^\top = \nabla\boldsymbol{\mu}_{\mathrm{GP}}^\top$. Given the definition of $\boldsymbol{\mu}_{\mathrm{GP}}$, each element of the first derivative (ignoring a factor of two in the first term) of eqn. (9) with respect to the LJ parameters only can be approximated as:

$$\nabla g(\boldsymbol{\theta})_j = \mathrm{Tr}\left({\boldsymbol{\Sigma}_{\mathrm{exp}}}^{-1}\frac{\partial\boldsymbol{\Sigma}_{\mathrm{GP}}}{\partial\theta_j}\right) + \boldsymbol{\mu}_r^\top {\boldsymbol{\Sigma}_{\mathrm{exp}}}^{-1}\frac{\partial\boldsymbol{\mu}_{\mathrm{GP}}}{\partial\theta_j}. \tag{10}$$

Given this definition, we can also define each element of the second derivative as:

$$\nabla^2 g(\boldsymbol{\theta})_{j,j\prime} = \mathrm{Tr}\left({\boldsymbol{\Sigma}_{\mathrm{exp}}}^{-1}\frac{\partial^2\boldsymbol{\Sigma}_{\mathrm{GP}}}{\partial\theta_j\partial\theta j\prime}\right) + \frac{\partial\boldsymbol{\mu}_{\mathrm{GP}}^\top}{\partial\theta_j}{\boldsymbol{\Sigma}_{\mathrm{exp}}}^{-1}\frac{\partial\boldsymbol{\mu}_{\mathrm{GP}}}{\partial\theta_{j\prime}} + \boldsymbol{\mu}_r^\top\frac{\partial^2\boldsymbol{\mu}_{\mathrm{GP}}}{\partial\theta_j\partial\theta j\prime}. \tag{11}$$

However, we can expect that at the maximum likelihood estimate (MLE), all values of $\boldsymbol{\mu}_r$ are near-zero. Therefore, eqn. (11) simplifies to:

$$\nabla^2 g(\boldsymbol{\theta})_{j,j\prime} = \mathrm{Tr}\left({\boldsymbol{\Sigma}_{\mathrm{exp}}}^{-1}\frac{\partial^2\boldsymbol{\Sigma}_{\mathrm{GP}}}{\partial\theta_j\partial\theta j\prime}\right) + \frac{\partial\boldsymbol{\mu}_{\mathrm{GP}}^\top}{\partial\theta_j}{\boldsymbol{\Sigma}_{\mathrm{exp}}}^{-1}\frac{\partial\boldsymbol{\mu}_{\mathrm{GP}}}{\partial\theta_{j\prime}}. \tag{12}$$

Recall that for this work, we assume that experimental measurements are independent across molecules and temperatures such that $\boldsymbol{\Sigma}_{\mathrm{exp}}^{-1}$ is a diagonal matrix defined by eqn. (7). Thus,

eqn. (12) can also be simplified:

$$\mathbf{H}_{j,j\prime} = \nabla^2 g(\boldsymbol{\theta})_{j,j\prime} \approx \sum_{n\in N} \frac{1}{\sigma^2_{\text{exp},n}} \left(\frac{\partial \boldsymbol{\mu}_{GP}}{\partial \theta_j}\right)^\top \left(\frac{\partial \boldsymbol{\mu}_{GP}}{\partial \theta_j}\right) + \sum_{n\in N} \frac{1}{\sigma^2_{\text{exp},n}} \frac{\partial^2 \boldsymbol{\Sigma}_{\text{GP}}}{\partial \theta_j^2}. \tag{13}$$

S5 Comparison Between Rigid Bond and Flexible Bond Implementations.

We computed liquid densities as a function of temperature for all of the molecules in this study using the LJ parameters from AT-6b and either rigid bonds, or flexible bonds modeled with the GAFF harmonic bond potential. This section shows the numerical results for this analysis.

Table S2: Average percent deviation of liquid density between the rigid bond and flexible bond implementations using the AT-6b LJ parameters

Molecule	**APD %**
R-116	2.46
HFC-125	0.93
HFC-134a	0.20
R-14	30.30
HFC-143	0.59
HFC-143a	1.51
HFC-152a	0.93
HFC-161	0.57
R-170	0.27
HFC-23	0.92
HFC-32	0.86
HFC-41	1.00
R-50	2.12

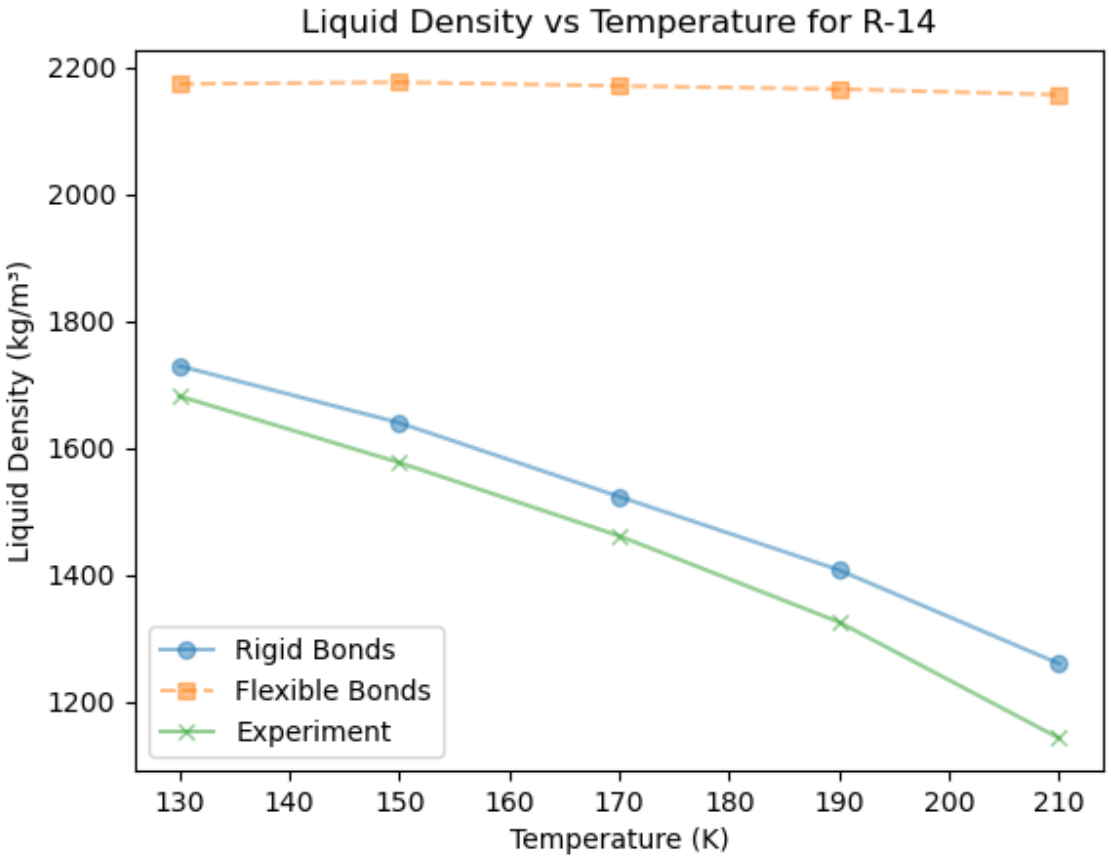

Figure S6: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for R-14.

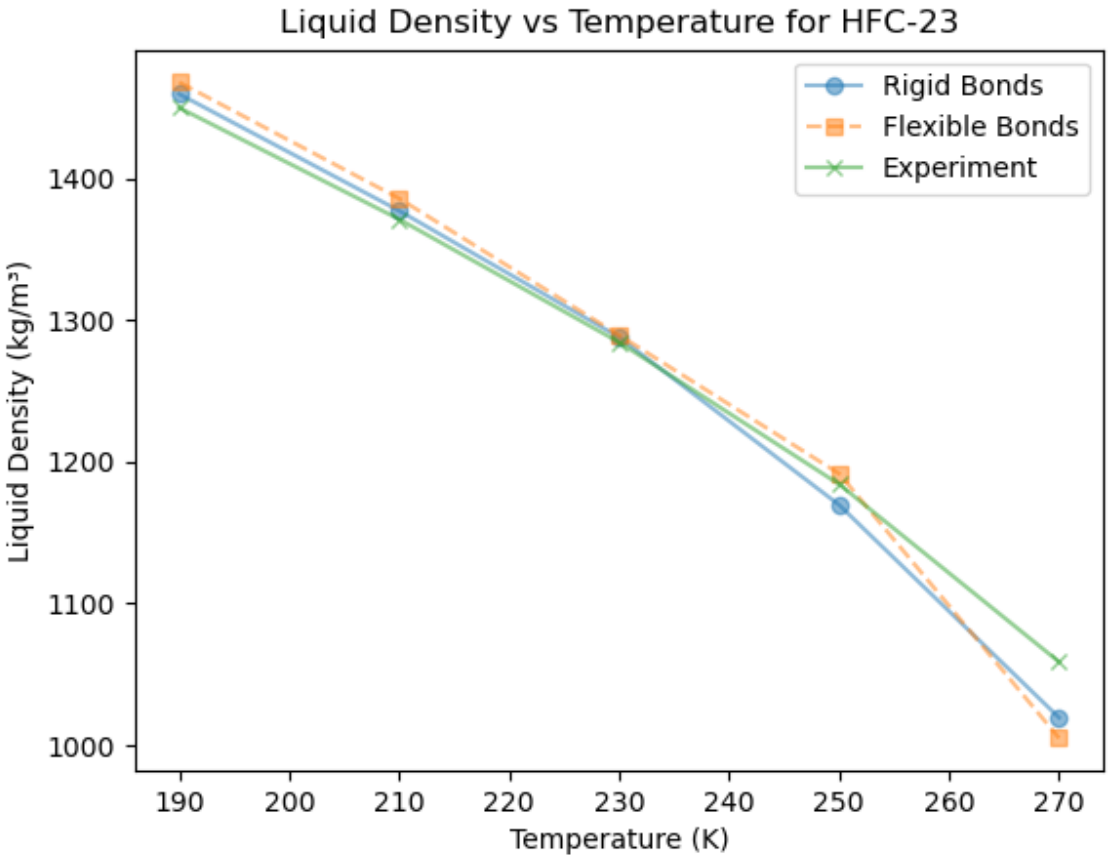

Figure S7: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for HFC-23.

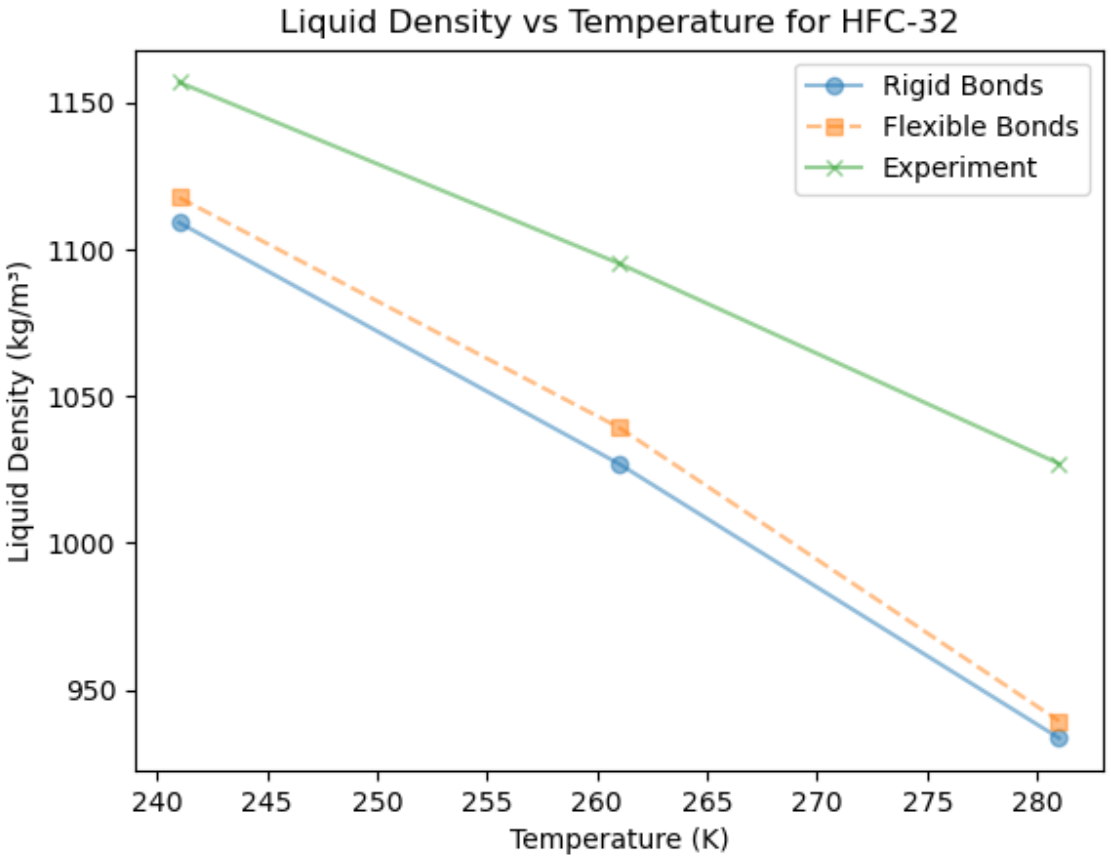

Figure S8: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for HFC-32.

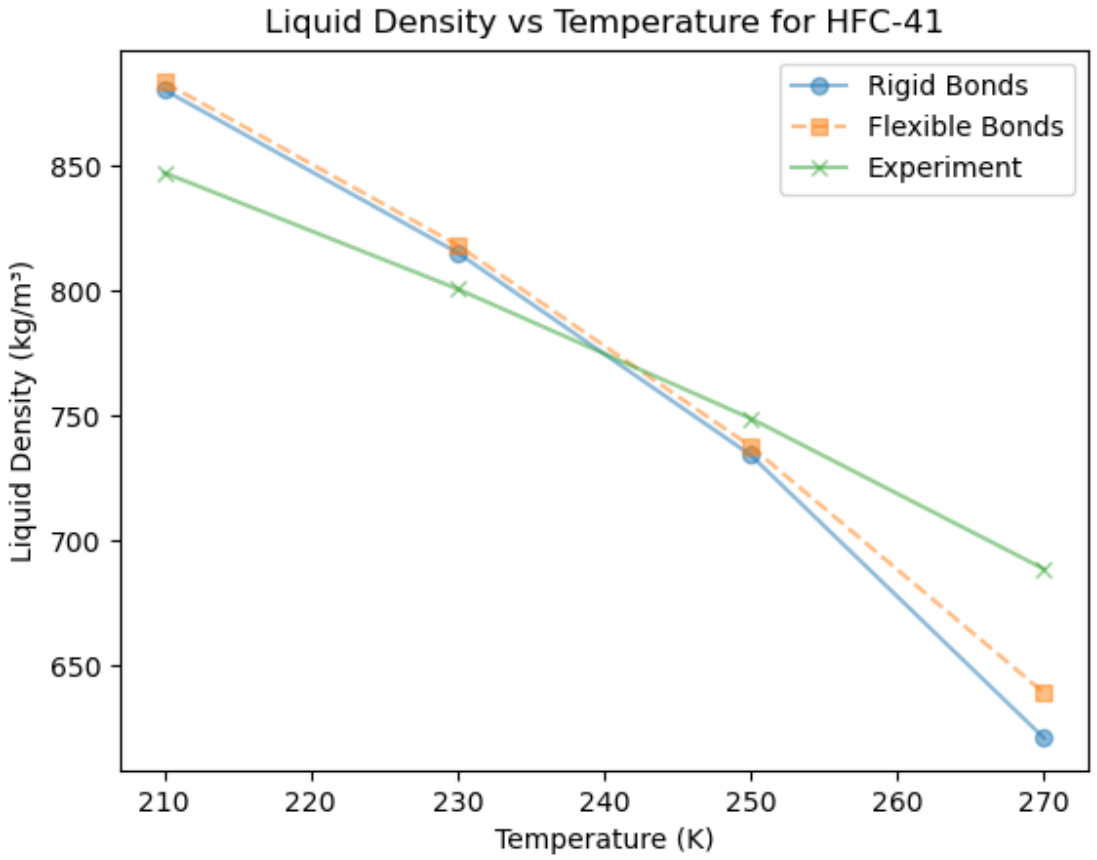

Figure S9: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for HFC-41.

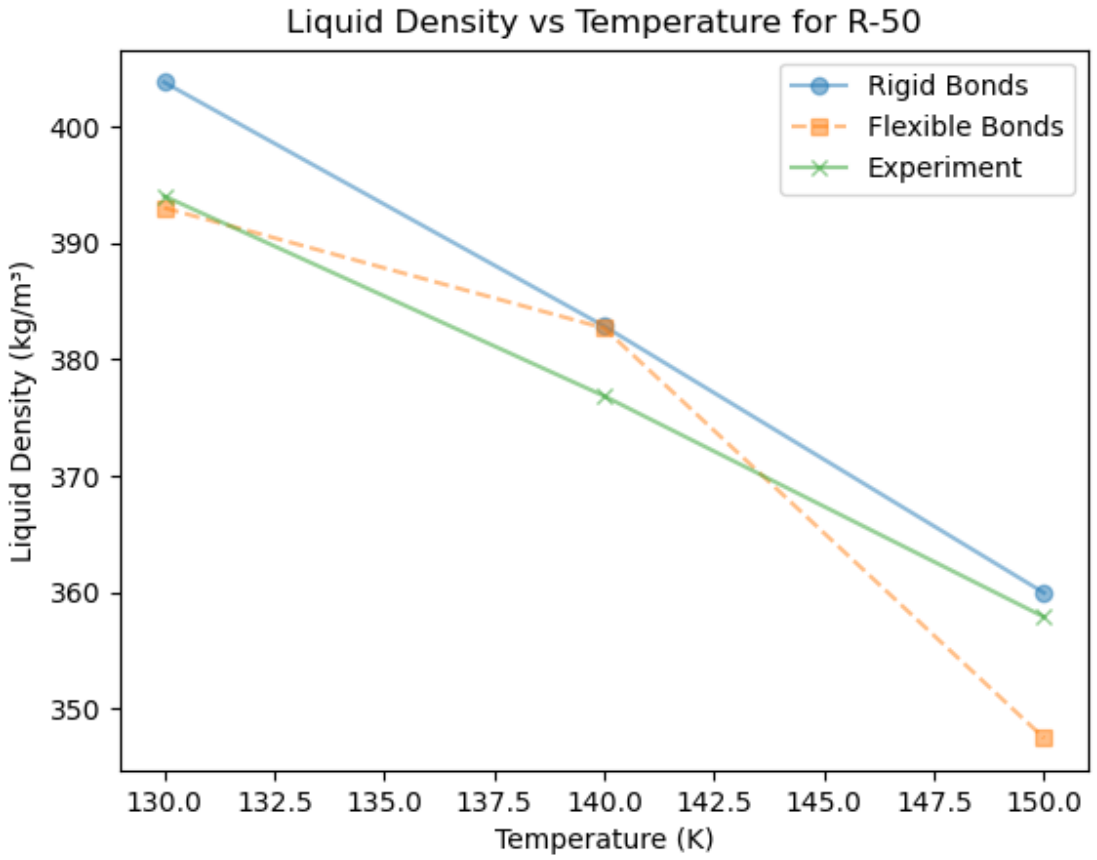

Figure S10: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for R-50.

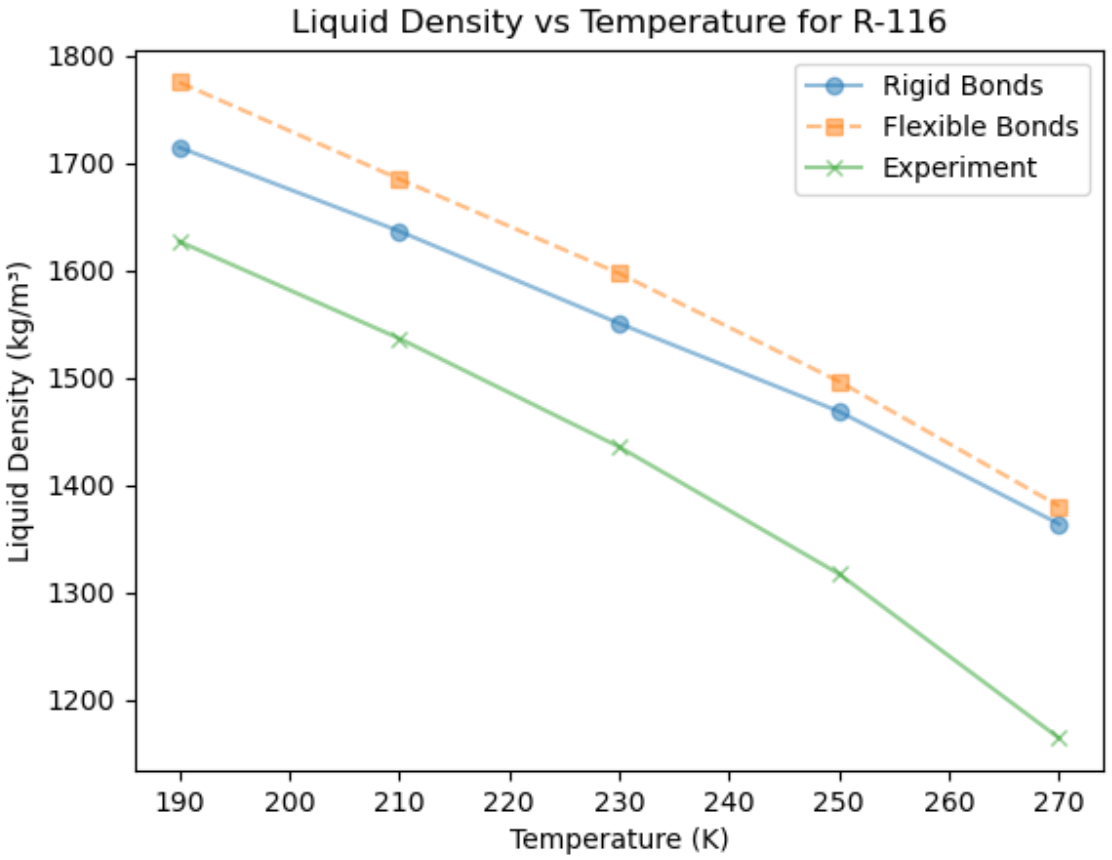

Figure S11: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for R-116.

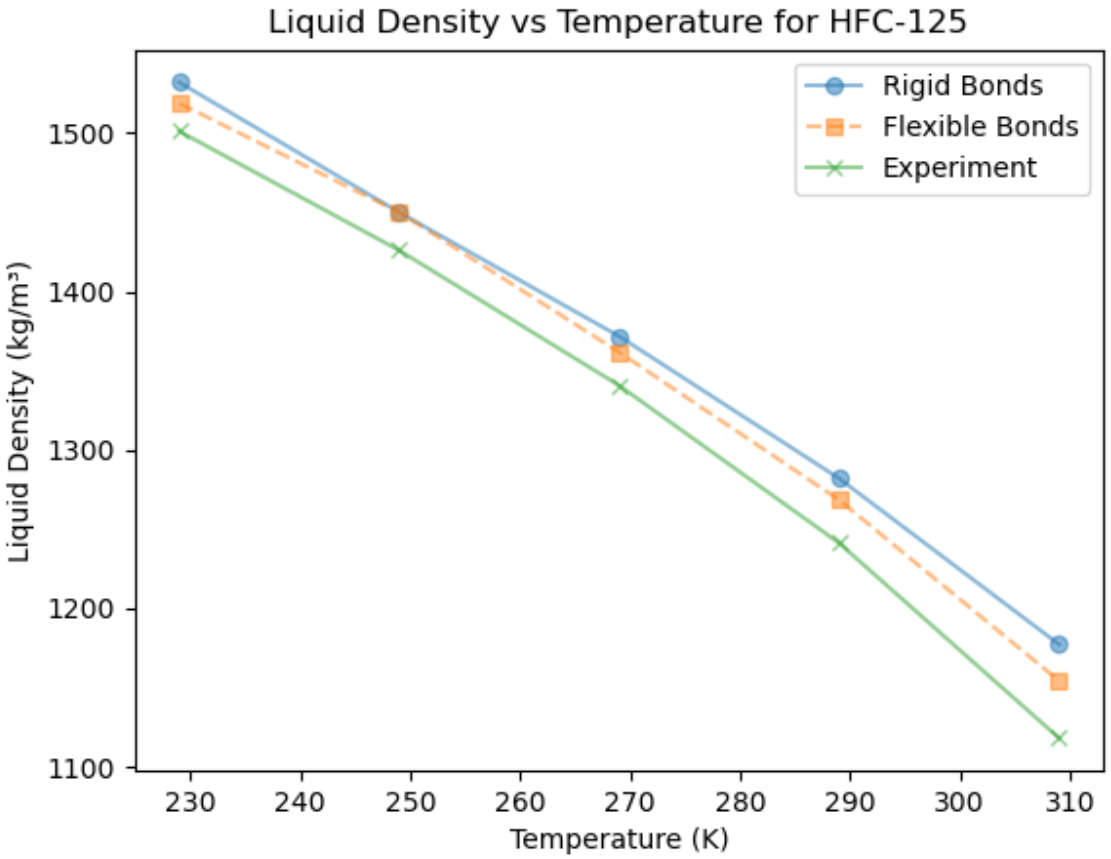

Figure S12: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for HFC-125.

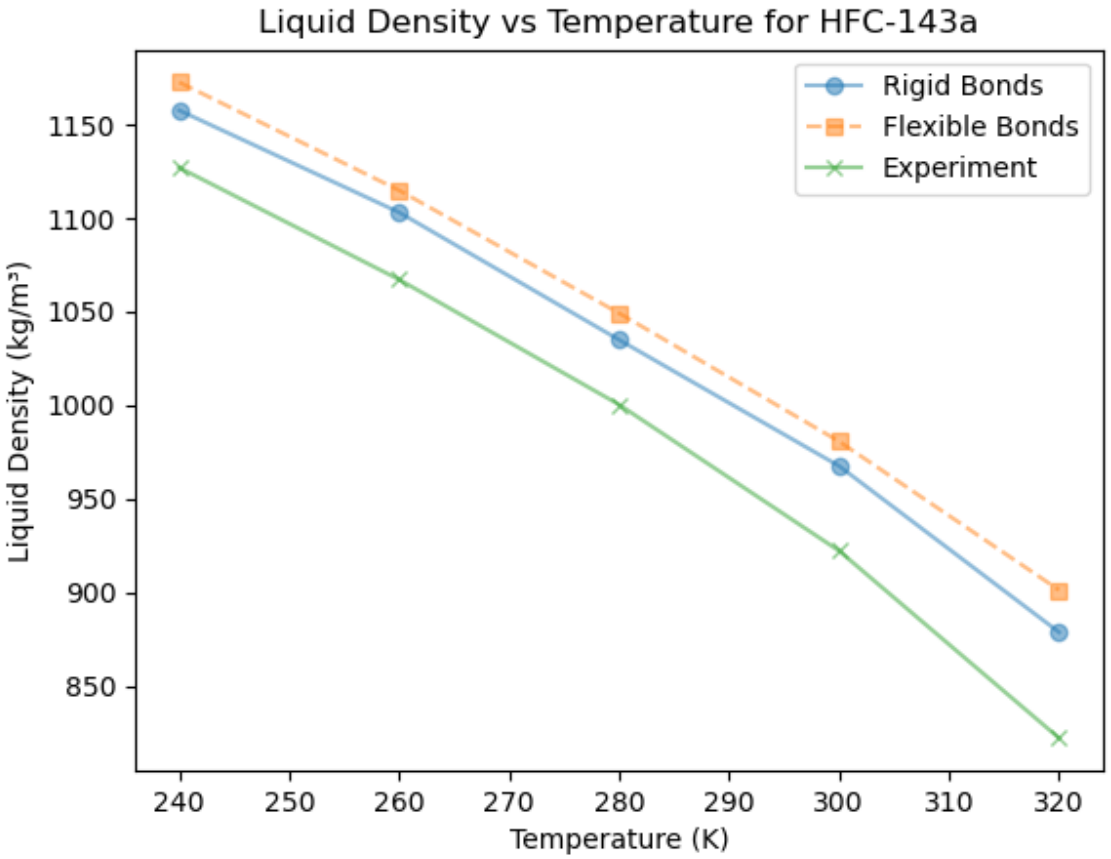

Figure S13: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for HFC-143a.

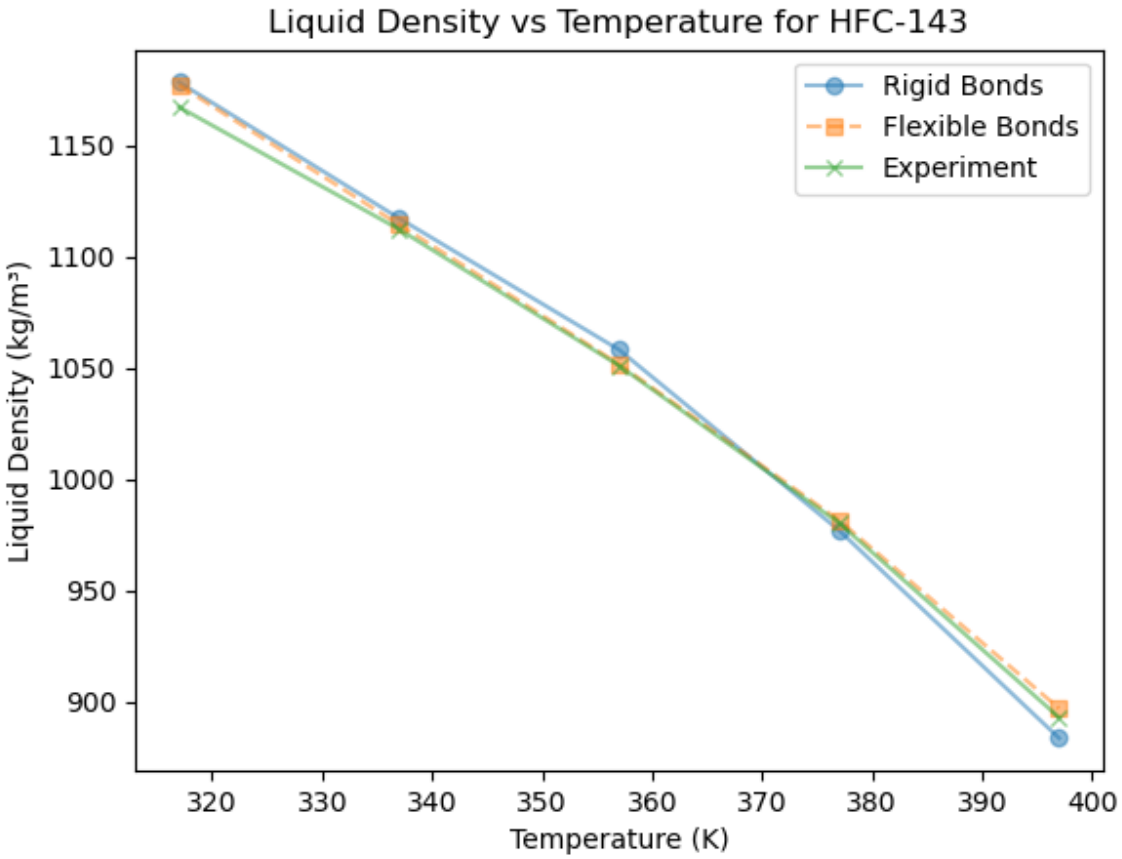

Figure S14: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for HFC-143.

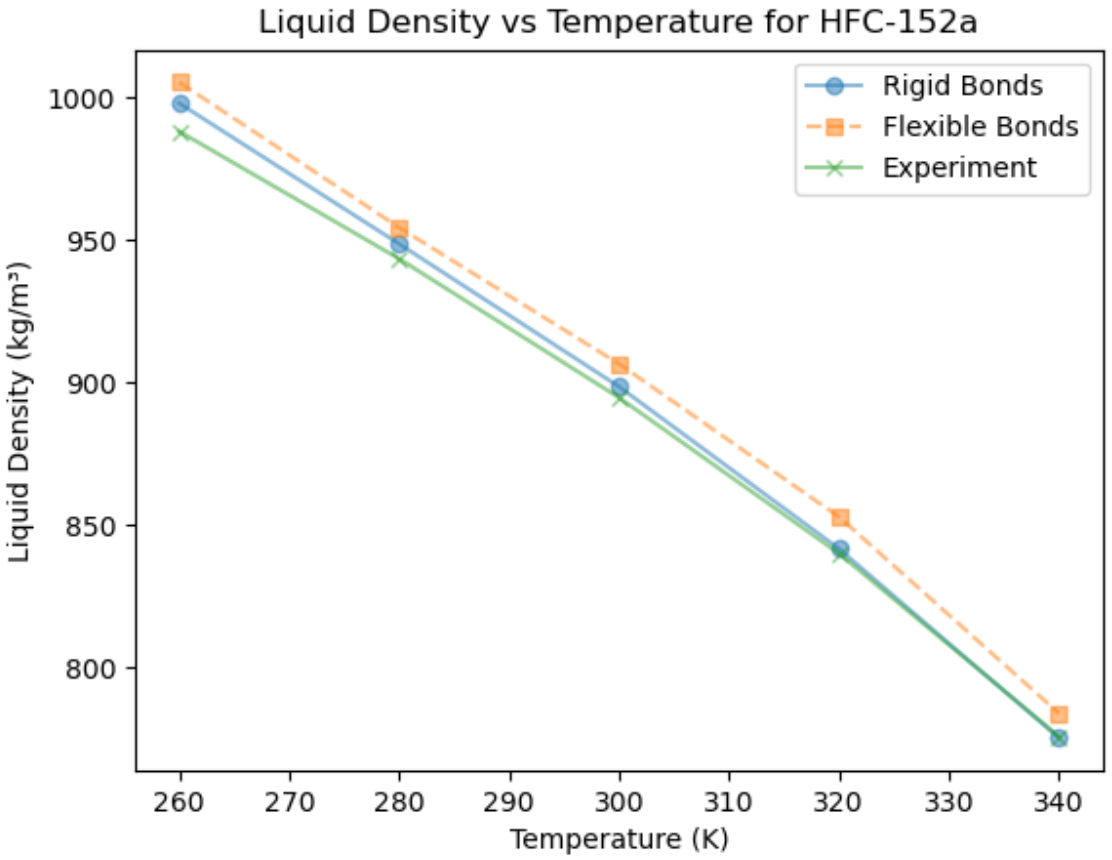

Figure S15: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for HFC-152a.

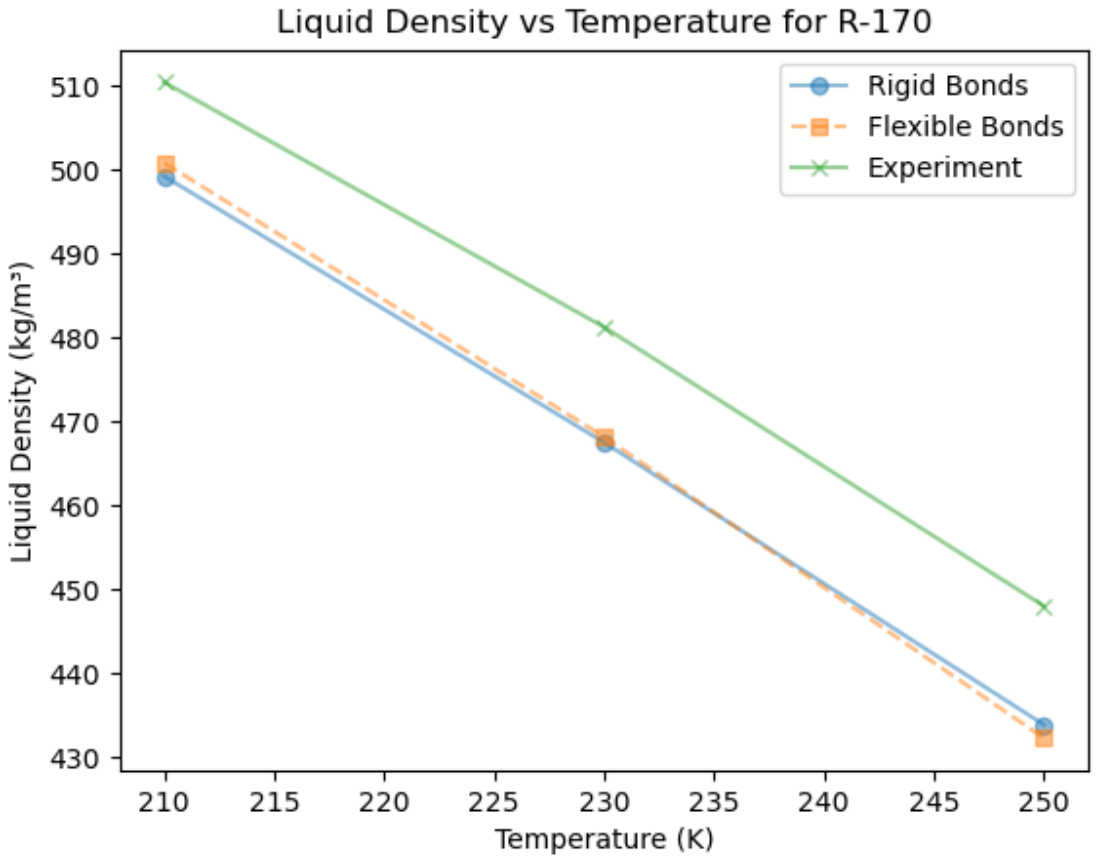

Figure S16: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for R-170.

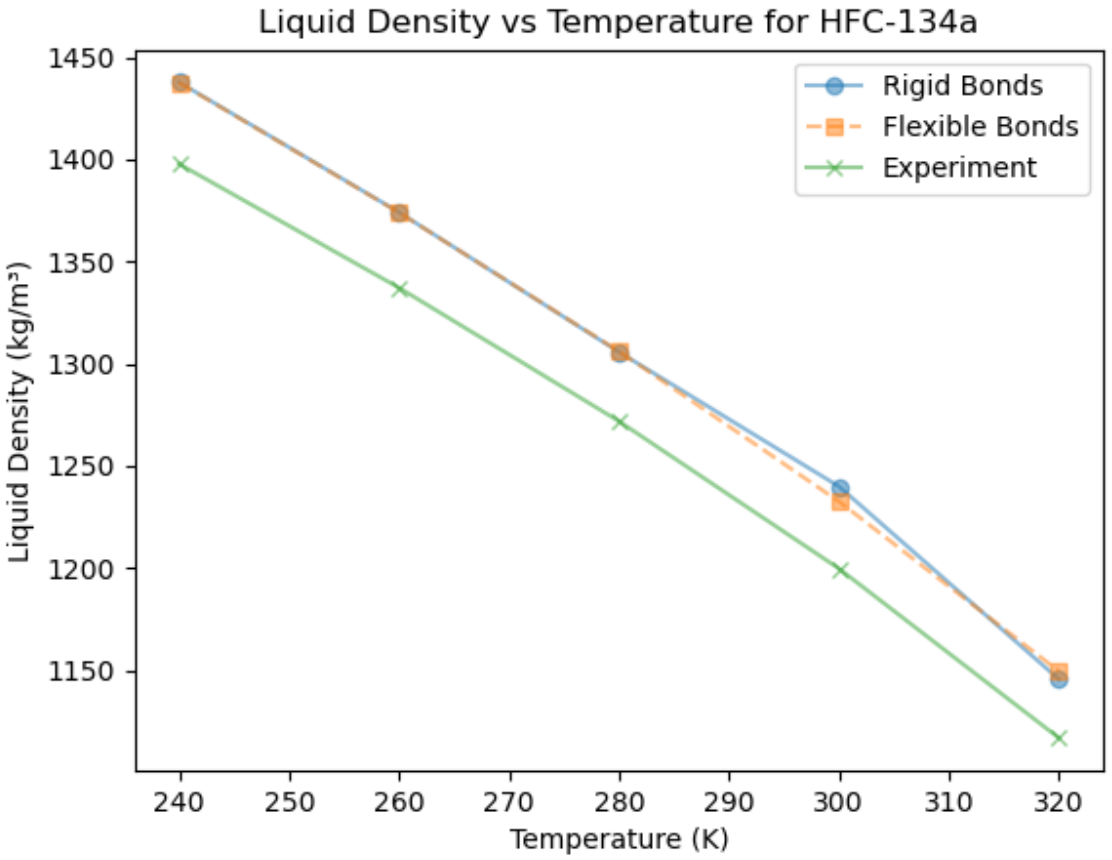

Figure S17: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for HFC-134a.

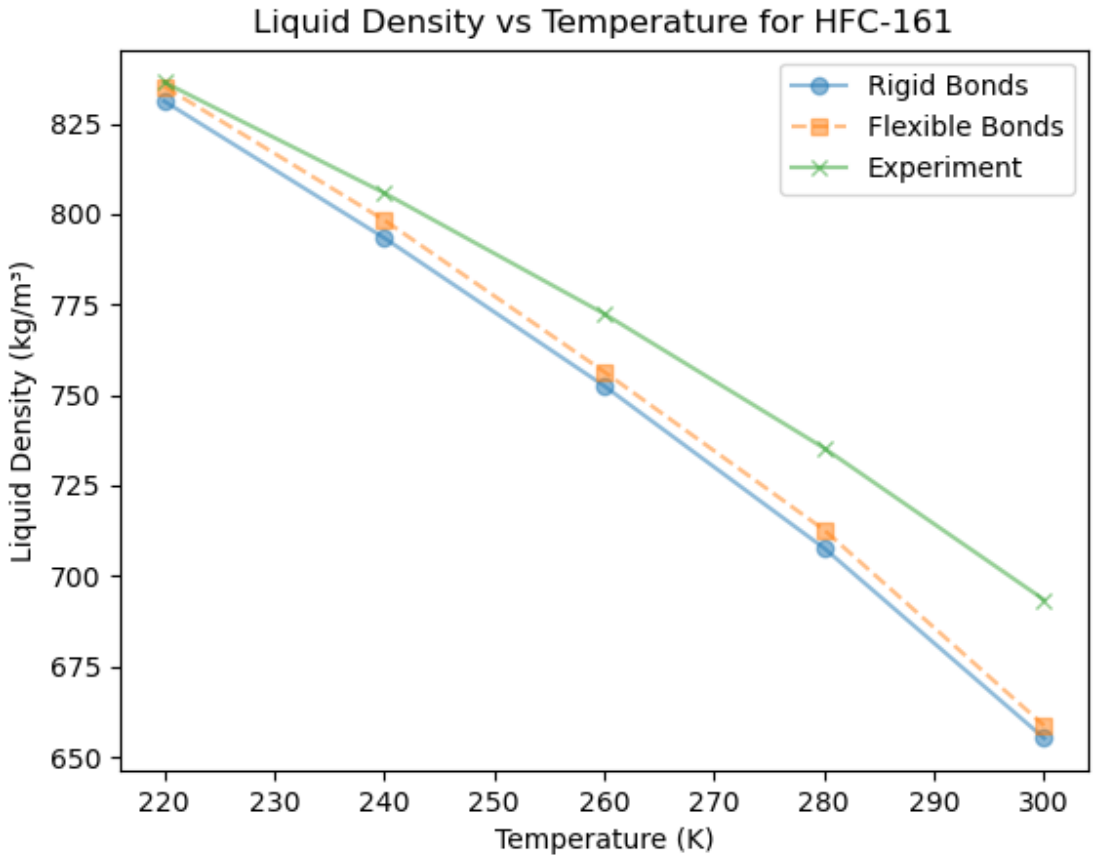

Figure S18: Liquid density predictions as a function of temperature for the rigid bond and flexible bond implementations for HFC-161.

S6 Full Model Validation Results

This section contains the figure which breaks down the validation results of each FF studied in this work by molecule and property.

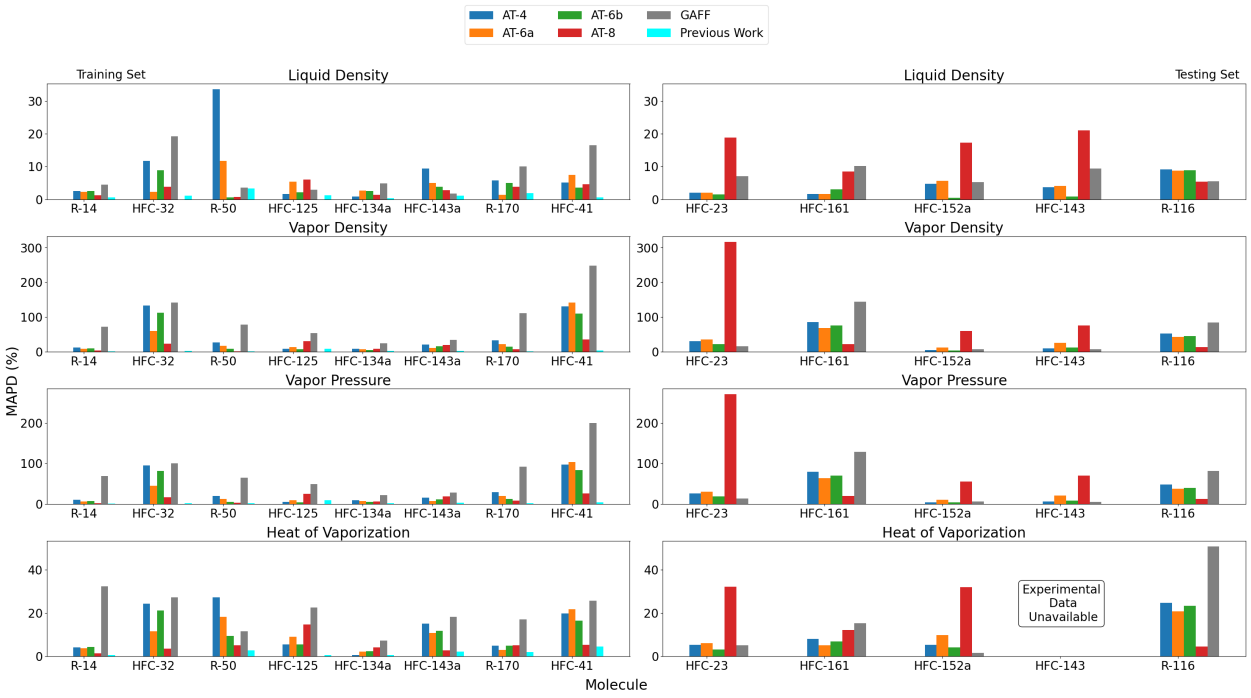

Figure S19: Average MAPD over all temperature values for each molecule and property of interest. Training set molecules appear on the left and testing set molecules appear on the right. We are unaware of enthalpy of vaporization data for HFC-143 and as such, MAPD values could not be computed. Previous work was completed by Befort et al. [11] (HFC-32 and HFC-125) and Wang et al. [12] (R-14, R-50, HFC-134a, HFC-143a, and R-170).

S7 Estimability Analysis Tables

This section displays the eigen-decomposition and orthoganlization algorithm analysis of the relevant atom typing schemes in this work. Note that in this section $\mathbf{H}$ is the Hessian of the objective function with respect to the LJ parameters. λ is an eigenvalue. Note that larger eigenvalues correspond to more important parameters.

Table S3: AT-6a eigen-decomposition

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
λ **of** H	1.98×10^{8}	8.78×10^{7}	9.54×10^{6}	1.76×10^{6}	1.18×10^{6}	9.51×10^{5}	6.36×10^{5}	3.50×10^{5}	7.08×10^{4}	1.62×10^{5}	0.00	0.00
σ_C	-2.20×10^{-1}	7.60×10^{-1}	-3.10×10^{-1}	-4.66×10^{-1}	-2.86×10^{-2}	2.11×10^{-1}	1.12×10^{-1}	3.70×10^{-2}	-4.11×10^{-2}	-1.37×10^{-2}	0.00	0.00
σ_{H_c}	4.22×10^{-3}	-2.50×10^{-2}	1.24×10^{-1}	4.23×10^{-2}	2.04×10^{-1}	4.65×10^{-1}	1.48×10^{-1}	-8.38×10^{-1}	2.88×10^{-4}	1.26×10^{-4}	0.00	0.00
σ_{H_1}	1.48×10^{-3}	-4.70×10^{-3}	1.39×10^{-2}	1.41×10^{-1}	-1.36×10^{-1}	-1.91×10^{-1}	9.61×10^{-1}	4.03×10^{-2}	-2.05×10^{-2}	-1.12×10^{-2}	0.00	0.00
σ_{H_2}	1.47×10^{-3}	-3.02×10^{-3}	6.14×10^{-3}	2.42×10^{-2}	-5.38×10^{-3}	-5.01×10^{-2}	-3.32×10^{-2}	-3.31×10^{-2}	-4.66×10^{-1}	-8.82×10^{-1}	0.00	0.00
σ_{H_3}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
σ_F	9.41×10^{-1}	3.29×10^{-1}	5.89×10^{-2}	4.36×10^{-2}	4.99×10^{-3}	-1.42×10^{-2}	-9.04×10^{-3}	-2.41×10^{-3}	3.31×10^{-3}	1.51×10^{-3}	0.00	0.00
ϵ_C	-1.45×10^{-1}	3.59×10^{-1}	-3.37×10^{-1}	8.31×10^{-1}	1.85×10^{-1}	-2.02×10^{-2}	-9.19×10^{-2}	-1.74×10^{-3}	5.23×10^{-2}	-5.10×10^{-3}	0.00	0.00
ϵ_{H_c}	-7.69×10^{-3}	5.23×10^{-2}	-2.81×10^{-1}	-9.69×10^{-2}	-3.43×10^{-1}	-6.90×10^{-1}	-1.43×10^{-1}	-5.40×10^{-1}	2.97×10^{-2}	4.64×10^{-2}	0.00	0.00
ϵ_{H_1}	-1.82×10^{-3}	5.41×10^{-3}	-1.16×10^{-2}	-2.07×10^{-1}	8.84×10^{-1}	-4.11×10^{-1}	7.53×10^{-2}	-1.18×10^{-2}	8.76×10^{-3}	5.17×10^{-3}	0.00	0.00
ϵ_{H_2}	-9.42×10^{-4}	8.24×10^{-4}	1.33×10^{-3}	5.25×10^{-2}	1.40×10^{-2}	-1.29×10^{-2}	-2.14×10^{-2}	-4.96×10^{-3}	-8.81×10^{-1}	4.69×10^{-1}	0.00	0.00
ϵ_{H_3}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
ϵ_F	-2.11×10^{-1}	4.27×10^{-1}	8.32×10^{-1}	1.15×10^{-1}	-6.78×10^{-2}	-2.33×10^{-1}	-8.00×10^{-2}	-4.50×10^{-2}	2.10×10^{-2}	1.44×10^{-2}	0.00	0.00

Table S4: AT-6b eigen-decomposition

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
λ **of** H	1.96×10^{8}	6.39×10^{7}	2.48×10^{7}	5.01×10^{6}	3.51×10^{6}	1.60×10^{6}	1.39×10^{6}	8.07×10^{5}	6.10×10^{5}	4.31×10^{4}	8.09×10^{4}	1.79×10^{5}
σ_{Cm}	-1.78×10^{-6}	-5.24×10^{-5}	-3.55×10^{-3}	6.51×10^{-2}	5.11×10^{-1}	1.84×10^{-2}	6.06×10^{-2}	-1.10×10^{-1}	1.85×10^{-1}	-7.86×10^{-1}	2.58×10^{-1}	1.99×10^{-3}
σ_{C1}	-1.53×10^{-1}	-7.09×10^{-1}	1.96×10^{-1}	3.49×10^{-1}	-4.03×10^{-2}	-3.28×10^{-1}	-3.67×10^{-1}	1.34×10^{-1}	2.27×10^{-1}	1.13×10^{-3}	1.54×10^{-3}	1.80×10^{-2}
σ_{C2}	-1.77×10^{-2}	-6.57×10^{-2}	-7.93×10^{-1}	1.73×10^{-1}	-8.77×10^{-2}	-7.24×10^{-3}	6.05×10^{-3}	-9.65×10^{-2}	2.48×10^{-1}	-1.28×10^{-1}	-4.90×10^{-1}	-3.60×10^{-2}
σ_{Hc}	6.17×10^{-4}	2.62×10^{-3}	5.29×10^{-2}	-3.64×10^{-2}	1.72×10^{-1}	2.55×10^{-2}	1.03×10^{-1}	-2.38×10^{-1}	4.51×10^{-1}	2.96×10^{-1}	1.26×10^{-1}	-7.70×10^{-1}
σ_H	3.25×10^{-3}	7.82×10^{-3}	3.63×10^{-3}	-3.48×10^{-2}	1.08×10^{-2}	4.72×10^{-1}	5.83×10^{-2}	7.89×10^{-1}	3.86×10^{-1}	3.68×10^{-3}	1.08×10^{-2}	1.36×10^{-2}
σ_F	9.63×10^{-1}	-2.62×10^{-1}	-1.15×10^{-2}	-2.92×10^{-2}	3.17×10^{-3}	2.69×10^{-2}	3.11×10^{-2}	-1.16×10^{-2}	-1.91×10^{-2}	8.21×10^{-5}	6.73×10^{-4}	-1.25×10^{-3}
ϵ_{Cm}	-2.70×10^{-6}	-7.49×10^{-5}	-5.02×10^{-3}	8.93×10^{-2}	6.90×10^{-1}	2.34×10^{-2}	7.52×10^{-2}	-1.20×10^{-1}	1.77×10^{-1}	4.80×10^{-1}	-1.29×10^{-1}	4.65×10^{-1}
ϵ_{C1}	-1.07×10^{-1}	-3.15×10^{-1}	1.09×10^{-1}	4.89×10^{-1}	-1.15×10^{-1}	5.07×10^{-1}	5.28×10^{-1}	-2.15×10^{-1}	-2.06×10^{-1}	-3.38×10^{-4}	5.92×10^{-4}	-9.54×10^{-3}
ϵ_{C2}	-1.21×10^{-2}	-4.14×10^{-2}	-5.22×10^{-1}	1.45×10^{-1}	-4.70×10^{-2}	-1.98×10^{-2}	-5.46×10^{-2}	5.12×10^{-2}	-7.95×10^{-2}	2.15×10^{-1}	7.97×10^{-1}	8.91×10^{-2}
ϵ_{Hc}	-7.63×10^{-4}	-5.56×10^{-3}	-1.30×10^{-1}	2.13×10^{-1}	4.39×10^{-1}	-3.83×10^{-2}	-1.49×10^{-1}	3.53×10^{-1}	-6.24×10^{-1}	4.23×10^{-2}	-1.58×10^{-1}	-4.26×10^{-1}
ϵ_H	-2.77×10^{-3}	-6.98×10^{-3}	-1.19×10^{-2}	1.72×10^{-2}	-1.44×10^{-2}	-6.25×10^{-1}	7.17×10^{-1}	3.04×10^{-1}	3.64×10^{-2}	6.62×10^{-4}	5.48×10^{-3}	-9.50×10^{-4}
ϵ_F	-1.91×10^{-1}	-5.68×10^{-1}	-1.68×10^{-1}	-7.26×10^{-1}	1.23×10^{-1}	1.32×10^{-1}	1.48×10^{-1}	-3.25×10^{-2}	-1.71×10^{-1}	-1.16×10^{-3}	-1.85×10^{-3}	-1.82×10^{-2}

Table S5: Importance ranking of AT-4 LJ parameters

Ranking	Eigen Decomposition	Algorithm 1[4]
1	σ_H	σ_{C_2}
2	ϵ_H	σ_F
3	σ_{C_1}	σ_{C_1}
4	ϵ_{C_1}	ϵ_F
5	σ_{C_2}	ϵ_{C_1}
6	ϵ_{C_2}	ϵ_{C_2}
7	ϵ_F	σ_H
8	σ_F	ϵ_H

Table S6: Importance ranking of AT-6a LJ parameters

Ranking	Eigen Decomposition	Algorithm 1^4
1	ϵ_{H_2}	σ_C
2	σ_{H_2}	σ_F
3	ϵ_C	ϵ_F
4	σ_C	σ_{H_c}
5	ϵ_{H_c}	ϵ_{H_c}
6	ϵ_F	ϵ_C
7	σ_{H_1}	σ_{H_1}
8	ϵ_{H_1}	ϵ_{H_1}
9	σ_F	ϵ_{H_2}
10	σ_{H_c}	σ_{H_2}
11	σ_{H_3}	σ_{H_3}
12	ϵ_{H_3}	ϵ_{H_3}

Table S7: Importance ranking of AT-6b LJ parameters

Ranking	Eigen Decomposition	Algorithm 1[4]
1	ϵ_{H_c}	σ_{C_2}
2	σ_{H_c}	σ_F
3	σ_H	ϵ_{C_m}
4	σ_{C_2}	σ_{C_1}
5	σ_{C_1}	ϵ_F
6	ϵ_{C_1}	σ_{H_c}
7	σ_{C_m}	σ_{C_m}
8	ϵ_{C_m}	ϵ_{C_1}
9	ϵ_F	ϵ_{H_c}
10	ϵ_{C_2}	σ_H
11	ϵ_H	ϵ_H
12	σ_F	ϵ_{C_2}

Table S8: Importance ranking of AT-8 LJ parameters

Ranking	Eigen Decomposition	Algorithm 1^4
1	ϵ_{F_2}	σ_{C_2}
2	σ_{F_2}	ϵ_{C_m}
3	σ_{F_3}	σ_{C_1}
4	ϵ_{F_3}	σ_{F_3}
5	ϵ_{C_1}	σ_{F_4}
6	ϵ_{C_2}	σ_H
7	σ_H	σ_{F_2}
8	ϵ_{C_m}	ϵ_{F_4}
9	σ_{F_1}	σ_{C_m}
10	σ_{C_2}	ϵ_H
11	σ_{C_1}	ϵ_{C_2}
12	ϵ_{F_1}	ϵ_{F_2}
13	σ_{C_m}	ϵ_{F_3}
14	ϵ_{F_4}	σ_{F_1}
15	ϵ_H	ϵ_{C_1}
16	σ_{F_4}	ϵ_{F_1}

S8 GP Error Metrics

This section details the GP MAPD for each molecule and property for the testing set.

Table S9: GP testing set MAPD values

Property	R-14	HFC-32	R-50	R-170	HFC-125	HFC-134a	HFC-143a
ΔH_{vap}	0.21	0.26	0.19	0.27	0.27	0.27	0.53
ρ_l	0.22	0.19	0.17	0.24	0.20	0.18	0.29
P_{vap}	6.93	1.80	2.35	1.67	4.50	5.99	3.11
ρ_v	14.45	2.34	3.47	3.88	5.58	6.80	4.79

References

(1) McLean, K. A.; McAuley, K. B. Mathematical Modelling of Chemical Processes - Obtaining the Best Model Predictions and Parameter Estimates Using Identifiability and Estimability Procedures. *Can. J. Chem. Eng.* **2012**, *90*, 351–366.

(2) Chis, O. T.; Villaverde, A. F.; Banga, J. R.; Balsa-Canto, E. On the Relationship Between Sloppiness and Identifiability. *Math. Biosci.* **2016**, *282*, 147–161.

(3) Wang, J.; Alexander, .; Dowling, W.; Dowling, A. W. Pyomo.DOE: An Open-Source Package for Model-Based Design of Experiments in Python. *AIChE Journal* **2022**, *68*, 1–24.

(4) Yao, K. Z.; Shaw, B. M.; Kou, B.; McAuley, K. B.; Bacon, D. W. Modeling Ethylene/Butene Copolymerization with Multi-Site Catalysts: Parameter Estimability and Experimental Design. *Polym. React. Eng.* **2003**, *11*, 563–588.

(5) Wu, S.; Mclean, K. A. P.; Harris, T. J.; Mcauley, K. B.; Wu, S.; Mclean, K. A. P.;

Harris, T. J.; Mcauley, K. B. Selection of Optimal Parameter Set Using Estimability Analysis and MSE-Based Model-Selection Criterion. *Int. J. Advanced Mechatronic Systems* **2011**, *3*, 188–197.

(6) Jacquez, J. A.; Greif, P. Numerical Parameter Identifiability and Estimability: Integrating Identifiability, Estimability, and Optimal Sampling Design. *Math. Biosci.* **1985**, *77*, 201–227.

(7) Rodriguez-Fernandez, M.; Egea, J. A.; Banga, J. R. Novel Metaheuristic for Parameter Estimation in Nonlinear Dynamic Biological Systems. *BMC Bioinf.* **2006**, *7*, 483.

(8) Brun, R.; Reichert, P.; Künsch, H. R. Practical Identifiability Analysis of Large Environmental Simulation Models. *Water Resour. Res.* **2001**, *37*, 1015–1030.

(9) Brun, R.; Kühni, M.; Siegrist, H.; Gujer, W.; Reichert, P. Practical Identifiability of ASM2d Parameters — Systematic Selection and Tuning of Parameter Subsets. *Water Res.* **2002**, *36*, 4113–4127.

(10) Rencher, A. C.; Schaalje, G. B. *Linear Models in Statistics*, 2nd ed.; John Wiley & Sons, Inc., 2008; pp 105–115.

(11) Befort, B. J.; Defever, R. S.; Tow, G. M.; Dowling, A. W.; Maginn, E. J. Machine Learning Directed Optimization of Classical Molecular Modeling Force Fields. *J. Chem. Inf. Model* **2021**, *19*, 28.

(12) Wang, N.; Carlozo, M. N.; Marin-Rimoldi, E.; Befort, B. J.; Dowling, A. W.; Maginn, E. J. Machine Learning-Enabled Development of Accurate Force Fields for Refrigerants. *J. Chem. Theory Comput.* **2023**, *19*, 4546–4558.