SUPPLEMENTARY INFORMATION: FITTING A SQUARE PEG IN A ROUND HOLE: PARAMETERISATION OF QUASI-SPHERICAL MOLECULES EMPLOYING THE MIE POTENTIAL

Gustavo Chaparro
Department of Chemical Engineering
Imperial College London
London, UK, SW7 2AZ
g.chaparro-maldonado21@imperial.ac.uk

Erich A. Müller*
Department of Chemical Engineering
Imperial College London
London, UK, SW7 2AZ
e.muller@imperial.ac.uk

Supplementary information for the parametrisation of quasi-spherical molecules through the Mie potential. The supplementary information covers the following:

1. Mie particle reduced units
2. Further modelling results

S.1 Mie particle reduced units

This work focuses on systems interacting through the Mie potential, Eq. (S.1).

$$\begin{aligned} \mathcal{U}^{\text{Mie}} &= \mathcal{C}^{\text{Mie}}\epsilon\left[\left(\frac{\sigma}{r}\right)^{\lambda_\text{r}} - \left(\frac{\sigma}{r}\right)^{\lambda_\text{a}}\right] \\ \mathcal{C}^{\text{Mie}} &= \frac{\lambda_\text{r}}{\lambda_\text{r}-\lambda_\text{a}}\left(\frac{\lambda_\text{r}}{\lambda_\text{a}}\right)^{\frac{\lambda_\text{a}}{\lambda_\text{r}-\lambda_\text{a}}} \end{aligned} \tag{S.1}$$

In Eq. (S.1), $\mathcal{U}^{\text{Mie}}$ refers to the interaction energy between two Mie particles, ϵ is the energy scale, σ is the shape parameter, which is loosely related to the diameter of a particle and r is the centre-to-centre distance between two particles. Finally, λ_r and λ_a are the repulsive and attractive exponents, respectively. Refs. [1–3] developed artificial neural network (ANN) models to predict the thermophysical properties of the Mie particle. These ANN-based models sue reduced units, which is equivalent to setting σ, ϵ, molar mass, M, Avogadro's constant, N_Av, and the Boltzmann constant, k_B, to 1. The definition of relevant thermophysical properties in reduced units are summarized in Table S.1.

*Correponding author

Table S.1: Definition of physical quantities in reduced units.

Quantity	Symbol	Units	Dimensionless Definition[a, b, c]
Time	τ	s	$\tau^* = \tau/\sigma\sqrt{M/N_{\mathrm{Av}}\epsilon}$
Distance	r	m	$r^* = r/\sigma$
Mass	m	$\mathrm{kg\,mol^{-1}}$	$m^* = m/M$
Density	ρ	$\mathrm{mol\,m^{-3}}$	$\rho^* = \rho\, N_{\mathrm{Av}}\sigma^3$
Temperature	T	K	$T^* = T\, k_{\mathrm{B}}/\epsilon$
Pressure	P	Pa	$P^* = P\, \sigma^3/\epsilon$
Entropy	S	$\mathrm{J\,K^{-1}\,mol^{-1}}$	$S^* = S/N_{\mathrm{Av}}k_{\mathrm{B}}$
Helmholtz free energy	A	$\mathrm{J\,mol^{-1}}$	$A^* = A/N_{\mathrm{Av}}\epsilon$
Internal energy	U	$\mathrm{J\,mol^{-1}}$	$U^* = U/N_{\mathrm{Av}}\epsilon$
Enthalpy	H	$\mathrm{J\,mol^{-1}}$	$H^* = H/N_a\epsilon$
Gibbs free energy	G	$\mathrm{J\,mol^{-1}}$	$G^* = G/N_{\mathrm{Av}}\epsilon$
Isochoric heat capacity	C_V	$\mathrm{J\,K^{-1}\,mol^{-1}}$	$C_V^* = C_V/N_{\mathrm{Av}}k_{\mathrm{B}}$
Thermal expansion coefficient	α_P	$\mathrm{K^{-1}}$	$\alpha_P^* = \alpha_P\, \epsilon/k_{\mathrm{B}}$
Isothermal compressibility	κ_T	$\mathrm{Pa^{-1}}$	$\kappa_T^* = \kappa_T\, \epsilon/\sigma^3$
Thermal pressure coefficient	γ_V	$\mathrm{Pa\,K^{-1}}$	$\gamma_V^* = \gamma_V\, \sigma^3/k_{\mathrm{B}}$
Isobaric heat capacity	C_P	$\mathrm{J\,K^{-1}\,mol^{-1}}$	$C_P^* = C_P/N_{\mathrm{Av}}k_{\mathrm{B}}$
Joule-Thomson coefficient	μ_{JT}	$\mathrm{K\,Pa^{-1}}$	$\mu_{\mathrm{JT}}^* = \mu_{\mathrm{JT}}\, k_{\mathrm{B}}/\sigma^3$
Second virial coefficient	B_2	$\mathrm{m^3\,mol^{-1}}$	$B_2^* = B_2/N_{\mathrm{Av}}\sigma^3$
Third virial coefficient	B_3	$\mathrm{m^6\,mol^{-2}}$	$B_3^* = B_3/N_{\mathrm{Av}}^2\sigma^6$
Self-diffusivity	D	$\mathrm{m^2\,s^{-1}}$	$D^* = D/\sigma\sqrt{N_{\mathrm{Av}}\epsilon/M}$
Shear Viscosity	η	$\mathrm{Pa\,s}$	$\eta^* = \eta\sigma^2/\sqrt{\epsilon M/N_{\mathrm{Av}}}$
Thermal Conductivity	κ	$\mathrm{Wm^{-1}K^{-1}}$	$\kappa^* = \kappa\sigma^2/k_{\mathrm{B}}\sqrt{N_{\mathrm{Av}}\epsilon/M}$

[a]Superscript "*" refers to a reduced property. [b]$N_{\mathrm{Av}} = 6.02214076 \cdot 10^{23}\ \mathrm{mol^{-1}}$ is the Avogadro constant [4], $k_{\mathrm{B}} = 1.380649 \cdot 10^{-23}\ \mathrm{J\,K^{-1}}$ is the Boltzmann constant [4] and M is the molar mass of the Mie particle. [c]σ is the shape parameter in meters, and ϵ is the potential well depth of the Mie potential in Joules. See Eq. (S.1).

S.2 Further modelling results

This section presents further parameterisation results for quasi-spherical molecules using the Mie potential. These results consider the modelling results for Krypton (Figure S.2), Xenon (Figure S.3), and Carbon Monoxide (Figure S.4). These figures present two possible modelling approaches that use different molecular parameter sets. Figures (a), (b), and (c) show vapour-liquid equilibria (VLE) results predicted by the FE-ANN EoS [1]. Figures (d) to (g) show vapour-liquid (VLE), solid-liquid (SLE) and solid-vapour (SVE) equilibria obtained using the FE-ANN(s) EoS [3]. This EoS is also used to predict the speed of sound in Figure (g). Then, the shear viscosity and thermal conductivity results obtained using the ANN-based models from Ref. [2] are shown in Figures (h) and (i). Additionally, the different molecular parameters are assessed against experimental self-diffusivity data collected by Allers *et al.* [5] in Figure S.5.

The following parametrisation approaches considered minimising the following objective function.

$$\mathrm{OF}(\sigma, \epsilon, \lambda_{\mathrm{r}}) = \sum_k \frac{w_k}{N_{p,k}} \sum_{i=1}^{N_{p,k}} \left| \frac{y_{k,i}^{\mathrm{pred}}}{y_{k,i}^{\mathrm{exp}}} - 1 \right| \tag{S.2}$$

Details of the different approaches are given in Table S.2. The reader is referred to the main article for a discussion of the impact of different thermophysical properties on the resulting molecular parameters.

These objective functions rely on the FE-ANN [1] or FE-ANN(s) [3] EoSs for modelling the Helmholtz free energy.

$$\begin{aligned} A^{\mathrm{res}} &= N_{\mathrm{Av}}\epsilon A^{*,\mathrm{res}} \\ A^{*,\mathrm{res}} &= \mathrm{ANN}\left(\alpha_{\mathrm{vdw}}, \rho N_{\mathrm{Av}}\sigma^3, \frac{\epsilon}{T k_{\mathrm{B}}}\right) - \mathrm{ANN}\left(\alpha_{\mathrm{vdw}}, \rho = 0, \frac{\epsilon}{T k_{\mathrm{B}}}\right) \end{aligned} \tag{S.3}$$

Table S.2: Parameterisation approaches to obtain the molecular parameters of the Mie potential.[a]

	P_{VLE}[b]	$\rho_{\mathrm{l,VLE}}$[b]	ΔH_{VLE}[b]	P_{SLE}[c]	P_{SVE}[d]	ΔH_{SVE}[d]	η[e]
OF_1	✓	✓	✓	×	×	×	×
OF_2	✓	✓	✓	×	×	×	✓
OF_3	✓	✓	✓	✓	✓	✓	×

[a] Thermophysical properties considered in the objective function (OF) are indicated by a "✓" in a grey-coloured cell. [b] Vapour-liquid Equilibria (VLE) properties are also referred to as "vaporisation" properties. These include the vaporization pressure (P_{VLE}), saturated liquid density ($\rho_{\mathrm{l,VLE}}$), and vaporization enthalpy (ΔH_{VLE}). VLE properties can be obtained either with the FE-ANN EoS [1] or FE-ANN(s) EoS [3], based on Eq. (S.3). [c] Solid-liquid equilibria (SLE) properties are also referred to as "melting" properties. Only the melting pressure (P_{SLE}) is considered here. [d] Solid-vapour equilibria (SVE) properties are also referred to as "sublimation" properties. These include the sublimation pressure (P_{SVE}) and sublimation enthalpy (ΔH_{SVE}). SLE and SVE properties are obtained using the FE-ANN(s) EoS. [e] The shear viscosity (η) is modelled using the ANN-based models developed in Ref. [2], based on Eq. (S.6).

These EoSs rely on the α_{vdw} parameter, which is defined as follows.

$$\alpha_{\mathrm{vdw}}(\lambda_{\mathrm{r}}, \lambda_{\mathrm{a}} = 6) = \mathcal{C}^{\mathrm{Mie}} \left[\frac{1}{3} - \frac{1}{\lambda_{\mathrm{r}} - 3} \right] \tag{S.4}$$

The FE-ANN/FE-ANN(s) EoS explicitly model the Helmholtz free energy. The reader is referred to the original publications [1, 3] or the book of Michelsen & Mollerup [6] for details about Helmholtz free energy thermodynamics.

Additionally, the self-diffusivity (D), shear viscosity (η) and thermal conductivity (κ) are modelled using the ANN-based models developed in Ref. [2].

- Self-diffusivity

$$\begin{aligned} \rho D &= \frac{1}{\sigma^2} \sqrt{\frac{\epsilon}{M N_{\mathrm{Av}}}} \rho^* D^* \\ \rho^* D^* &= \mathrm{ANN} \left(\alpha_{\mathrm{vdw}}, \rho N_{\mathrm{Av}} \sigma^3, \frac{\epsilon}{T k_{\mathrm{B}}} \right) \end{aligned} \tag{S.5}$$

- Shear viscosity

$$\begin{aligned} \eta &= \frac{\sqrt{\epsilon M / N_{\mathrm{Av}}}}{\sigma^2} \exp[\ln \eta^*] \\ \ln \eta^* &= \mathrm{ANN} \left(\alpha_{\mathrm{vdw}}, \rho N_{\mathrm{Av}} \sigma^3, \frac{\epsilon}{T k_{\mathrm{B}}} \right) \end{aligned} \tag{S.6}$$

- Thermal conductivity

$$\begin{aligned} \kappa^{\mathrm{ANN}} &= \frac{k_{\mathrm{B}} \sqrt{N_{\mathrm{Av}} \epsilon / M}}{\sigma^2} \exp[\ln \kappa^*] \\ \ln \kappa^* &= \mathrm{ANN} \left(\alpha_{\mathrm{vdw}}, \rho N_{\mathrm{Av}} \sigma^3, \frac{\epsilon}{T k_{\mathrm{B}}} \right) \end{aligned} \tag{S.7}$$

The thermal conductivity is corrected to account for vibrational contributions [7] not considered in the ANN-based model. This correction uses the self-diffusivity obtained from Eq. (S.5).

$$\kappa = \kappa^{\mathrm{ANN}} + \rho D \left(C_V^{\mathrm{id}} - \#\text{ non-vibrational degrees of freedom} \cdot \frac{R}{2} \right) \tag{S.8}$$

The thermal conductivity obtained from Eq. (S.7) has been trained over a wide range of density and temperature conditions. It has been formulated to fulfil physics-inspired constraints. This formulation ensures correct behaviour within the trained phase space. As shown in the original publication [2], the performance of the model degrades in the vicinity of the phase envelope and at low temperatures. In Figure S.1, it is shown the data distribution and liquid thermal conductivity across the saturation line of different $(\lambda_{\mathrm{r}}, 6)$ Mie fluids. From the top row of this figure, it is noted that the data distribution in the vicinity of the VLE is not even in the database, where "lower" to "moderate" repulsive exponents (i.e., $\lambda_{\mathrm{r}} = 8$ and 12) have fewer data points approaching the triple temperature. This data scarcity affects the prediction across the saturation line. As observed in the bottom row of Figure S.1, the ANN model predicts a monotonic behaviour close to the triple point for these fluids ($\lambda_{\mathrm{r}} = 8$ and 12). This unexpected behaviour is predicted beyond the training region

and should not be considered correct. This incorrect behaviour is corrected for Mie fluid where more data is available (i.e., $\lambda_r = 16$, 24 and 30). In this case, the liquid thermal conductivity across the saturation line follows a monotonic behaviour.

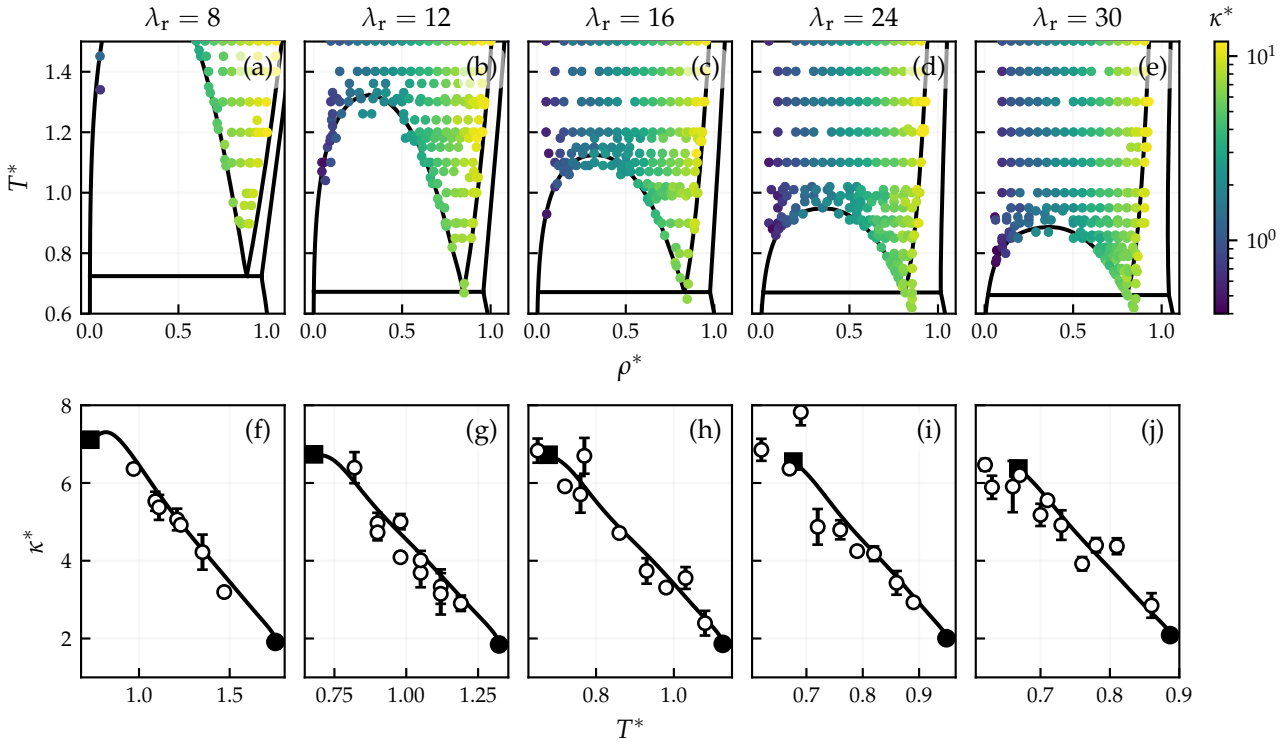

Figure S.1: **Thermal conductivity modelling across the saturation line. (a)** - **(e)** Data distribution in the density-temperature space. **(f)** - **(j)** Liquid thermal conductivity across the saturation line as predicted by Eq. (S.7).

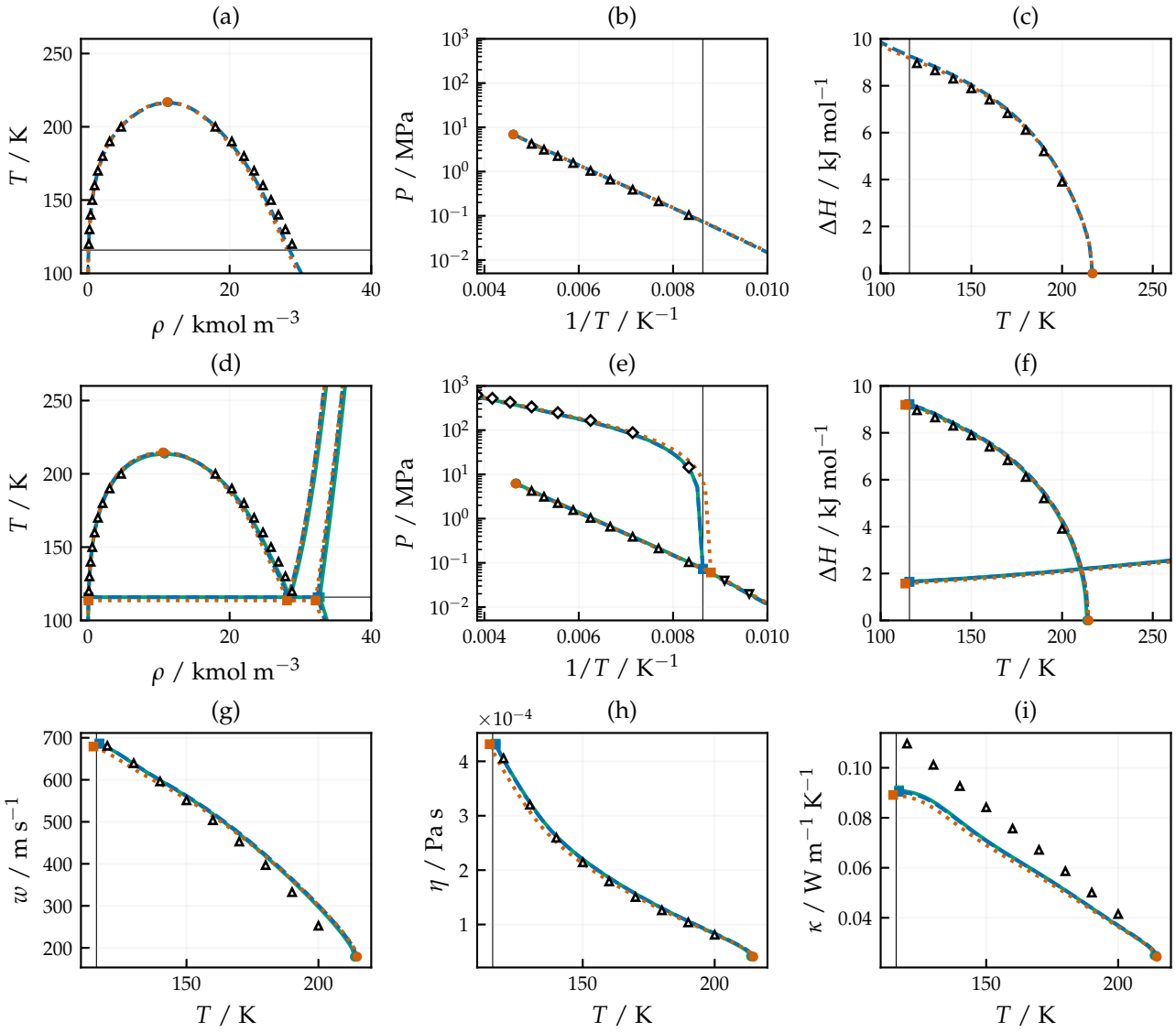

Figure S.2: **Selected thermophysical properties of Krypton.** **(a)** and **(d)** Phase envelope. **(b)** and **(e)** Clapeyron plot. **(c)** and **(f)** Enthalpy of phase change. **(g)** Speed of sound. **(h)** Shear viscosity. **(i)** Thermal conductivity. **(a)** - **(c)** Phase equilibria modelled using the FE-ANN EoS developed Ref. [1] **(d)** - **(g)** Phase equilibria and speed of sound modelled using the FE-ANN(s) EoS developed in Ref. [3]. FE-ANN/FE-ANN(s) EoSs are based on Eq. (S.3). **(h)** and **(i)** The shear viscosity and thermal conductivity are modelled using the ANN-based models developed Ref. [2], based on Eqs. (S.6) and (S.8). Dashed blue lines and symbols refer to results using parameters optimised with OF_1 (VLE only). Dotted orange lines and symbols refer to results using parameters optimised with OF_2 (VLE + shear viscosity). Solid green lines and symbols refer to results using parameters optimised with OF_3 (VLE + SLE + SVE). Refer to Eq. (S.2) and Table S.2 for further details about the objective functions. Filled circle is the critical point. Filled square is the triple point. Pseudo-experimental data obtained from NIST TRC [8]. Upright triangles are VLE data, diamonds are SLE data, and downward triangles are SLE data. For reference, the solid black line refers to the triple temperature.

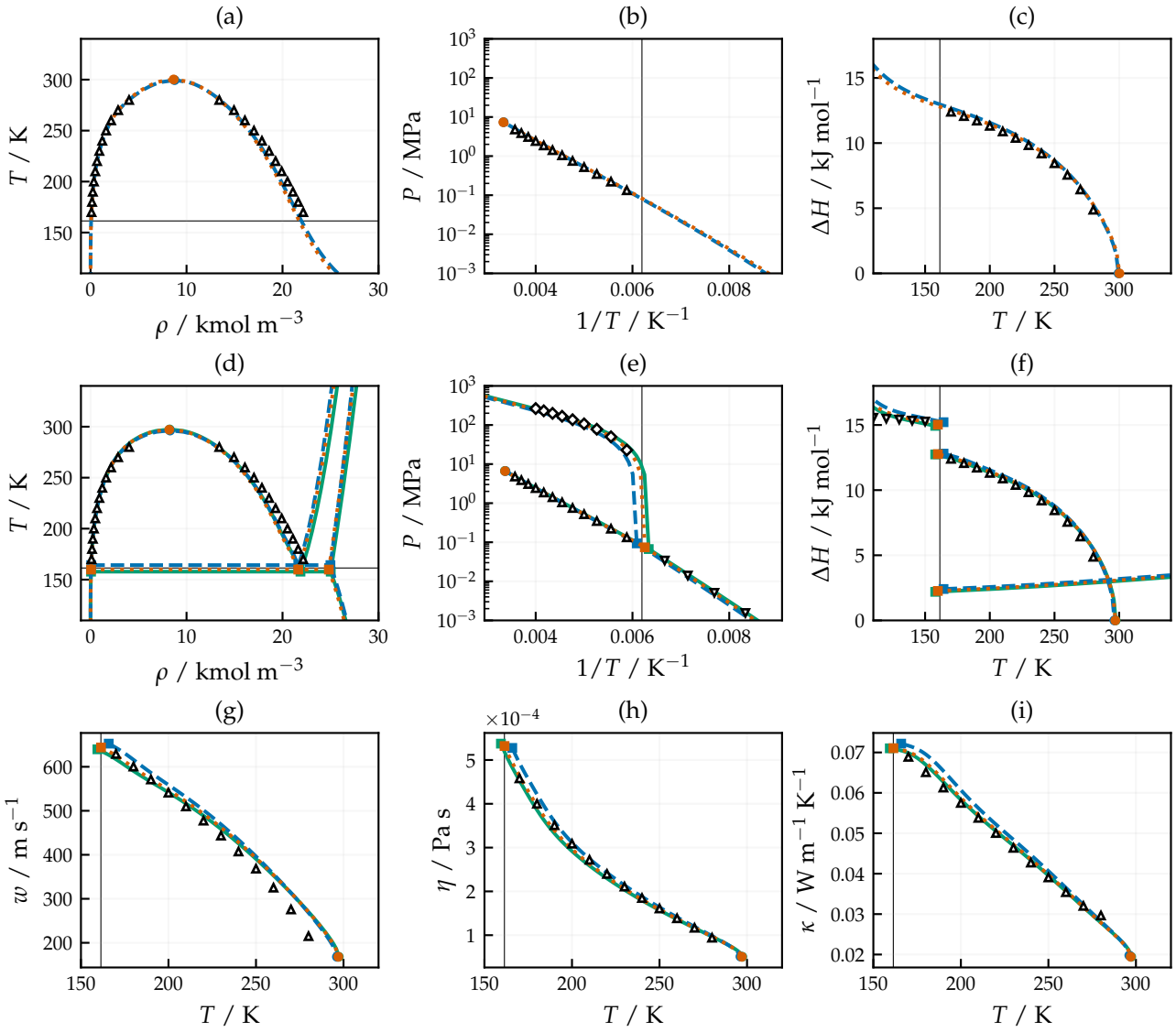

Figure S.3: **Selected thermophysical properties of Xenon.** **(a)** and **(d)** Phase envelope. **(b)** and **(e)** Clapeyron plot. **(c)** and **(f)** Enthalpy of phase change. **(g)** Speed of sound. **(h)** Shear viscosity. **(i)** Thermal conductivity. **(a)** - **(c)** Phase equilibria modelled using the FE-ANN EoS developed Ref. [1] **(d)** - **(g)** Phase equilibria and speed of sound modelled using the FE-ANN(s) EoS developed in Ref. [3]. FE-ANN/FE-ANN(s) EoSs are based on Eq. (S.3). **(h)** and **(i)** The shear viscosity and thermal conductivity are modelled using the ANN-based models developed Ref. [2], based on Eqs. (S.6) and (S.8). Dashed blue lines and symbols refer to results using parameters optimised with OF_1 (VLE only). Dotted orange lines and symbols refer to results using parameters optimised with OF_2 (VLE + shear viscosity). Solid green lines and symbols refer to results using parameters optimised with OF_3 (VLE + SLE + SVE). Refer to Eq. (S.2) and Table S.2 for further details about the objective functions. Filled circle is the critical point. Filled square is the triple point. Pseudo-experimental data obtained from NIST TRC [8]. Upright triangles are VLE data, diamonds are SLE data, and downward triangles are SLE data. For reference, the solid black line refers to the triple temperature.

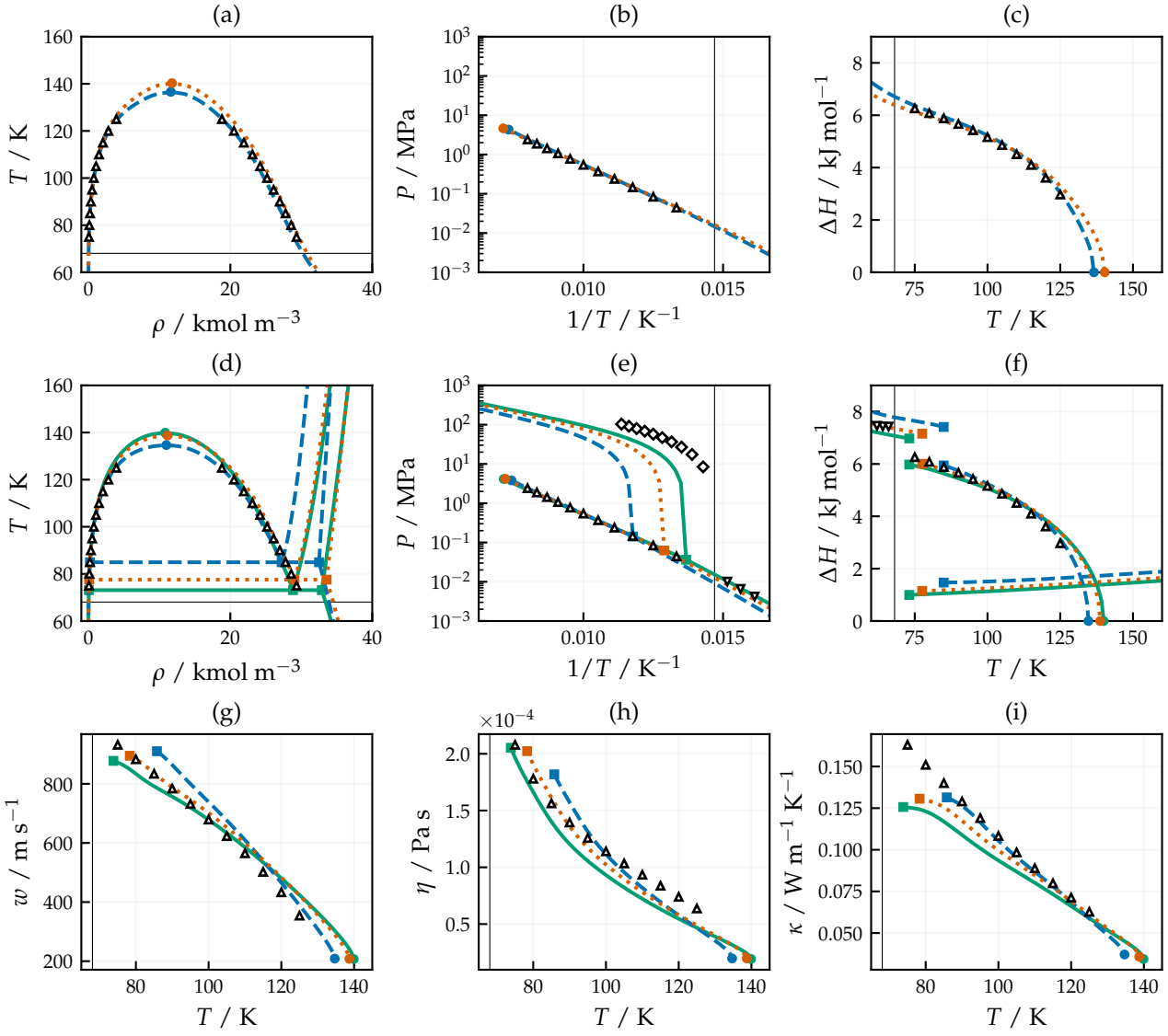

Figure S.4: **Selected thermophysical properties of Carbon monoxide.** **(a)** and **(d)** Phase envelope. **(b)** and **(e)** Clapeyron plot. **(c)** and **(f)** Enthalpy of phase change. **(g)** Speed of sound. **(h)** Shear viscosity. **(i)** Thermal conductivity. **(a)** - **(c)** Phase equilibria modelled using the FE-ANN EoS developed Ref. [1] **(d)** - **(g)** Phase equilibria and speed of sound modelled using the FE-ANN(s) EoS developed in Ref. [3]. FE-ANN/FE-ANN(s) EoSs are based on Eq. (S.3). **(h)** and **(i)** The shear viscosity and thermal conductivity are modelled using the ANN-based models developed Ref. [2], based on Eqs. (S.6) and (S.8). Dashed blue lines and symbols refer to results using parameters optimised with OF_1 (VLE only). Dotted orange lines and symbols refer to results using parameters optimised with OF_2 (VLE + shear viscosity). Solid green lines and symbols refer to results using parameters optimised with OF_3 (VLE + SLE + SVE). Refer to Eq. (S.2) and Table S.2 for further details about the objective functions. Filled circle is the critical point. Filled square is the triple point. Pseudo-experimental data obtained from NIST TRC [8]. Upright triangles are VLE data, diamonds are SLE data, and downward triangles are SLE data. For reference, the solid black line refers to the triple temperature.

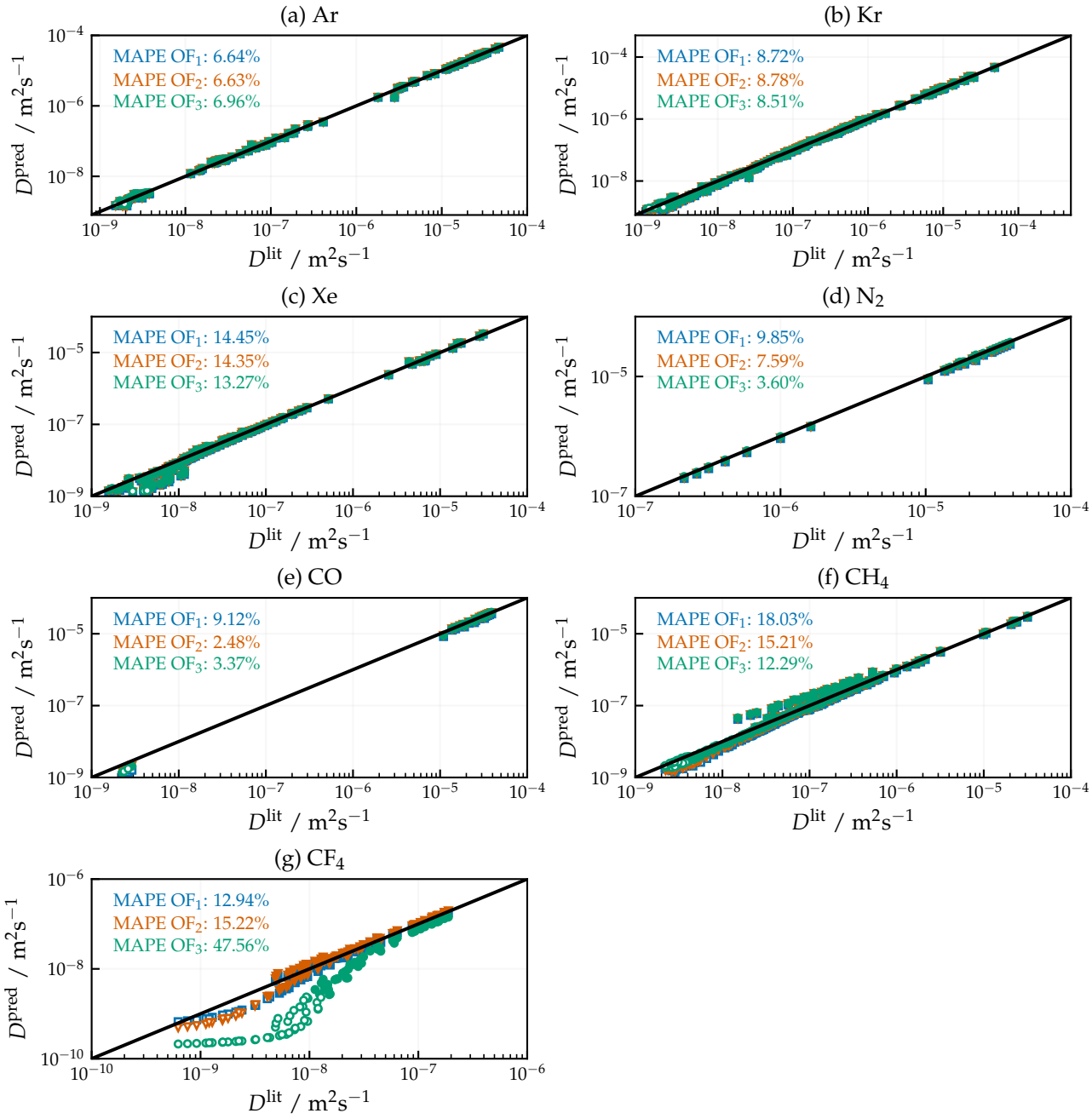

Figure S.5: **Self-diffusivity modelling of quasi-spherical molecules (a)** Argon. **(b)** Krypton. **(c)** Xenon. **(d)** Nitrogen. **(e)** Carbon Monoxide. **(f)** Methane. **(g)** Tetrafluoromethane. Experimental data collected by Allers *et al.* [5]. Filled symbols refer to fluid data points, and empty symbols refer to solid data points as predicted by the FE-ANN(s) EoS [3]. Blue squares refer to results using parameters optimised with OF_1 (VLE only). Orange downward triangles refer to results using parameters optimised with OF_2 (VLE + shear viscosity). Green circles refer to results using parameters optimised with OF_3 (VLE + SLE + SVE). Refer to Eq. (S.2) and Table S.2 for further details about the objective functions. Self-diffusivity computed using Eq. (S.5)

References

1. Chaparro, G. & Müller, E. A. Development of thermodynamically consistent machine-learning equations of state: Application to the Mie fluid. en. *The Journal of Chemical Physics* **158,** 184505. doi:10.1063/5.0146634 (2023).
2. Chaparro, G. & Müller, E. A. Simulation and Data-Driven Modeling of the Transport Properties of the Mie Fluid. en. *The Journal of Physical Chemistry B* **128,** 551–566. doi:10.1021/acs.jpcb.3c06813 (2024).
3. Chaparro, G. & Müller, E. A. Development of a Helmholtz free energy equation of state for fluid and solid phases via artificial neural networks. en. *Communications Physics* **7,** 406. doi:10.1038/s42005-024-01892-3 (2024).
4. Newell, D. B. *et al.* The CODATA 2017 values of h, e, k, and N_A for the revision of the SI. en. *Metrologia* **55,** L13–L16. doi:10.1088/1681-7575/aa950a (2018).
5. Allers, J. P., Garzon, F. H. & Alam, T. M. Artificial neural network prediction of self-diffusion in pure compounds over multiple phase regimes. en. *Physical Chemistry Chemical Physics* **23,** 4615–4623. doi:10.1039/D0CP06693A (2021).
6. Michelsen, M. L. & Mollerup, J. M. *Thermodynamic models: fundamentals & computational aspects* 2nd. eng (Tie-Line Publications, Holte, 2007).
7. Liang, Z. & Tsai, H.-L. The vibrational contribution to the thermal conductivity of a polyatomic fluid. en. *Molecular Physics* **108,** 1707–1714. doi:10.1080/00268976.2010.489520 (2010).
8. Linstrom, P. *NIST Chemistry WebBook, NIST Standard Reference Database 69* en. 1997. doi:10.18434/T4D303.