CHEUNG, Howard, et al. Covariance-based uncertainty analysis of reference equations of state. Journal of Chemical & Engineering Data, 2020, 65.2: 503-522.

Covariance-based uncertainty analysis of reference equations of state

Howard Cheung,[*,†] Jérôme Frutiger,[*,‡] Ian H. Bell,[*,¶] Jens Abildskov,[‡] Gürkan Sin,[‡] and Shengwei Wang[§]

†*Carbon Exchange (Hong Kong) Ltd., Shatin, Hong Kong*
‡*Process and Systems Engineering Center (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark (DTU), 2800 Kgs. Lyngby, Denmark*
¶*Applied Chemicals and Materials Division, National Institute of Standards and Technology, Boulder, CO, 80305, USA*
§*Department of Building Services Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong*

E-mail: howard.cheung@carbonexchange.net; jfru@kt.dtu.dk; ian.bell@nist.gov

Abstract

This work presents a detailed methodology for uncertainty analysis applied to a reference equation of state (EOS) based on Helmholtz energy. With increasing interest in uncertainties of thermal process models, it is important to quantify the property uncertainties from the EOS. However, literature relating to EOS EOS development and parameter estimation either does not report uncertainties or report underestimated values. This work addresses the issue by introducing a covariance-based methodology of uncertainty analysis based on a linear approximation. The uncertainty ranges of the EOS properties (95 % confidence intervals) are calculated from the experimental values and the EOS model structure through the parameter covariance matrix and subsequent linear error propagation. In this case study, the Helmholtz-based EOS of propane is analyzed.The uncertainty methodology is general and is applicable to any novel or existing EOS because it does not re-train the EOS. The study demonstrates the insights a thorough uncertainty analysis can give for EOS users and developers. Uncertainties vary strongly as a function of the state point, and uncertainties of saturation properties are much larger than the uncertainties of the vapor region due to the use of Maxwell criteria to calculate the saturation properties.

1 Introduction

In recent years, there have been several applications of uncertainty analysis of thermal systems. This includes the selection of working fluids for Rankine cycles,[1] virtual sensors of air conditioning systems,[2,3] chiller control systems,[4] and evaluation of chiller performance models for fault detection and diagnostics (FDD) algorithms.[5] There have been systematic efforts in describing the uncertainty of both fundamental physicochemical models[6] as well as correlation-based property models for process engineering applications.[7,8] The uncertainties of equations of state (EOS) have not yet been adequately studied; most studies neglect the impact of the uncertainty of the EOS. On the other hand, the accuracy of the EOS is commonly considered in the literature, but analyses of accuracy only consider the difference between the output predicted by the model and

experimental data. The consideration of accuracy instead of uncertainty is common among studies of EOS of different substances.[9–12] To this extent, the accuracy is different from the uncertainty of the output of the EOS, which is the range of statistically possible outcomes of the model given different property observations and is usually reported by the 95 % confidence interval. Uncertainty of the EOS may also be as important as the other uncertainties. For example, the accuracy of the curve fit for the saturation pressure for the refrigerant blend R410A[13] is 0.5 % of the pressure value, and this is on par with the measurement uncertainty of refrigerant pressure.[14] It is crucial, in the scope of good modeling practices, to take the uncertainties of EOS into account in order to establish the application range and the reliability of a thermal systems model.

In this manuscript a clear distinction is made between accuracy and uncertainty. The authors are aware that some studies[13] also consider uncertainty to be the difference between modeled and measured data, and do not use the term accuracy. However, for the sake of clarity, we intend to use the term accuracy and uncertainty separately in this manuscript.

Accuracy and uncertainty are different measures and are both important to assess the performance of a model.[15,16] Analysis of model accuracy in the literature of multiparameter reference EOS has mostly so far only involved the calculation of the closeness between some model results and known measurement values.[9–12] While a comparison of "closeness" can validate if a model yields an accurate representation of measurement values, it cannot describe the overall reliability of model results, and this model accuracy analysis does not tell the whole story for property estimates that are not associated with any known measurement values. The effect of the lack of information can only be quantified by uncertainty quantification.[17] Some variables calculated by the EOS such as enthalpy and entropy can only be measured through differences relative to a reference point, and it is critical to conduct uncertainty calculations for the estimates of these variables by approaches such as covariance matrix calculation to show the variability of the estimated enthalpy and entropy differences. Furthermore, uncertainty propagation also helps a user to understand how the uncertainty of an EOS affects the result of a thermodynamic cycle model or other process models, especially, when the operation characteristics of the cycle cannot be measured and hence an analysis of model accuracy is unavailable. To do so, it is necessary to make an analysis of the uncertainty of the EOS in addition to the commonly performed analyses on their accuracy.

Some studies[14,18] consider the accuracy of EOS of refrigerants[9–12] as the uncertainties of EOS, and they ignore the effect of the development of the EOS to the uncertainties of EOS. Two studies were conducted to account for the uncertainty due to the covariance.[19,20] Both of them propagated the uncertainty of measurement data that were used to fit the EOS for the uncertainty of the EOS. One of them required complete knowledge of the uncertainty of measurement data to build the EOS which is unavailable for some EOS.[19] The other one only involved covariance between the input data and the EOS and did not involve covariance between the EOS parameters – a crucial part of EOS uncertainty.[20] In addition, the method[20] required re-fitting of the EOS which would be too difficult to be used for all existing EOS. Other studies were conducted to show how to compare the propagation of the EOS parameter uncertainties to the output of a process model. The uncertainties of two EOS (cubic-type and SAFT-type) were analyzed. Two of the methods did not use the covariance matrix, but used a bootstrap method as an alternative to estimate the uncertainties of the EOS parameters.[21,22] While the methods performed well for specific applications, it was too computationally expensive for general use. There is a need to find another method that is less computationally expensive to quantify the uncertainty of EOS of fluids. Lasala et al.[23] presented a methodology to quantify and propagate the uncertainty of cubic equations of state as well-as residual Helmholtz energy-based mixing rules for fluids used in CO2 Capture processes using the covarianced-based method. However,

this method has not been applied to Helmholtz-based equations of state for pure components with large number of parameters.

There is also an increasing industrial interest in the analysis of the uncertainty of EOS and the state-of-the-art is to conduct the uncertainty analysis using only the accuracy of the EOS in the literature[12] or the propagation of the uncertainties from the inputs[24] for computationally efficient calculation. However, these two techniques are insufficient to cover important uncertainty components in the estimation of a thermodynamic property of a pure substance as shown in Table 1.

Table 1: The steps and sources of uncertainties in the estimation of thermodynamic properties of a pure substance from an EOS

Step	Process in the task	Potential sources of uncertainties	Uncertainties propagated from previous steps	Uncertainty propagation and quantification procedures
1	Choose a model form (i.e. an EOS), a regression method and the corresponding objective function for the thermodynamic properties of a pure substance	(a) Structural uncertainty		
2	Choose the state points of the fluid at which we can measure properties and collect data to train the model	(b) Epistemic uncertainty in the choices of training data		
3	Conduct the experiments to measure the state variables such as temperature T, pressure p and density ρ for training data	(c) Stochastic and epistemic uncertainty from measurements of training data		Uncertainty analysis of the experiments for (c)
4	Perform the parameter estimation process	(d) Epistemic uncertainty of the numerical methods in the parameter estimation	(a) + (b) + (c)	Uncertainty quantification based on model covariance matrix for (a) and (b), and uncertainty propagation (e.g. through linear error propagation) for (c)
5	Obtain the input variables to EOS (e.g. temperature T and pressure p) at new conditions	(e) Subjective and stochastic uncertainties of the measurement of input variables at the new conditions	(a) + (b) + (c)	Uncertainty propagation (e.g. through linear error propagation) for (a), (b) and (c), and uncertainty analysis in the experiments for (e)
6	Evaluate the EOS to estimate the thermodynamic property of the substance at the new conditions	(f) Epistemic uncertainty of the numerical methods in the evaluation of EOS	(a) + (b) + (c) + (e)	Uncertainty propagation (e.g. through linear error propagation) for (a), (b), (c) and (e)

Table 1 shows the sources of uncertainties in the various steps to build an EOS of a pure substance and to use the EOS to estimate a thermodynamic property of the substance. It categorizes the uncertainty sources by structural uncertainties, epistemic (systematic) uncertainties, and stochastic uncertainties according to Sin *et al.*[25] In step 1, there are structural uncertainties for the mathematical form of the EOS, because other mathematical forms could be used to construct the EOS. Step 2 adds epistemic uncertainties to the process because the choices of training data of the EOS are subjective. Step 3 adds the epistemic and stochastic uncertainties of experimental setups and sensors, and different training data points have different uncertainty values.[26] Step 4 only introduces epistemic uncertainties of the numerical errors of numerical methods in the parameter estimation. The uncertainties involved are rather typically small (less than 10^{-6} % of the estimated dependent variable)[18] and are not calculated. However, the same process also enables the calculation of uncertainties of parameters which are highly related to the uncertainties in Steps 1 and 2. Step 5 is similar to Step 3: another set of data from experiments for the inputs to the EOS is considered that involves the epistemic and stochastic uncertainties of new measurement values. Step 6 is the use of the EOS and may involve numerical iterations to solve implicit equations, and hence its uncertainties are not calculated.

Table 1 not only shows the sources of uncertainties but also a critical feature of the uncertainty of a model: the uncertainty of a model output at a condition is not only determined by the uncertainty of training data at that condition. While intuition suggests that the uncertainty of a model at an accurately measured training data point should be small, the uncertainty of a model output at the data point also depends on other factors such as the structural uncertainty and the subjective uncertainties in the choices of training data. For example, if the structure of the model propagates the large uncertainty of model outputs at some conditions to model outputs at other conditions, the uncertainty of model outputs at other conditions may become larger than expected even if the corresponding training data points are observed accurately. Hence factors that affect the uncertainties of models in addition to measurement uncertainties of model inputs must be considered.

While there are many important components of uncertainties as shown in Table 1, previously reported methods only cover part of them. The "uncertainties" mentioned in the work of Lemmon *et al.*[12] only partially involve the calculation of the sum of squares of errors between estimation and measurement values of thermodynamic properties and part of the uncertainty components (a) and (b) in Table 1. Uncertainties propagated by the method of Kline and McClintock[24] only involve uncertainty component (e). Some methods also fail to address issues such as difficulty to refit the EOS and lack of knowledge of measurement uncertianties of the training data. Hence a study should be conducted to perform the uncertainty analysis of EOSs with other methods.

This paper aims to describe a covariance-based method to calculate uncertainty due to both the sum of squares of residuals and the covariance matrix of the EOS of fluids with the same applicable range as the EOS, using the example of a multiparameter reference EOS of propane (CAS registry number: 74-98-6).[12] The developed method does not require refitting of the equation of state of the fluid and can be used for an EOS of any pure substance. The method is written in a general form. If full and reliable measurement data are available, the methodology allows to take measurement uncertainty into account. However, if comprehensive information of the uncertainty of the training data is not available, the methodology shows how the EOS prediction uncertainty can be calculated without measurement uncertainty. Hence, the method invites both developers and users of EOS to consider the uncertainty through the covariance.

To validate the method, study its applicability, and illustrate its applications, the uncertainty is visualized in the saturation dome of the temperature-entropy and pressure-enthalpy diagrams. The uncertainties of sound speed,

specific heat, and density at different temperature and pressure are also visualized to study how they change under various conditions.

2 Uncertainty of parameter estimation by parameter covariance matrix

In this paper, the uncertainty calculation presented is based on the parameter covariance matrix.[27,28] The parameter covariance matrix quantifies the uncertainties of the parameters in an equation from a linear regression. It is calculated based on a linear approximation to a regression problem - the commonly performed regression for the fitting of the EOS parameters to experimental values - and is helpful for the derivation of the uncertainty of output of the regression equation - the equation of state.

As an example, we consider a property y that is described by a model $F(\mathbf{X}, \theta)$, with $\mathbf{X}$ as the matrix of input variables and θ being a vector of model parameters.[29] $\mathbf{X}$ represents the experimental measurements. The uncertainty analysis is performed after the successful identification of the best parameter estimates θ^* through non-linear regression (fitting to experimental values by minimization of an objective function) and is begun by the calculation of its parameter covariance matrix.

The calculation of the parameter covariance matrix allows linear error propagation for nonlinear regression models. Considering the above-mentioned property y described by $F(\mathbf{X}, \theta)$, the underlying assumption of this method for uncertainty analysis is that the errors ϵ (i.e. the differences between the model prediction and the experiments) are independently distributed and defined by a Gaussian distribution white noise (normal distribution with zero mean and unit standard deviation σ) as expressed by Eq. (1).

$$y = F(\mathbf{X}, \theta) + \epsilon \qquad \epsilon \sim N(0, \sigma^2) \qquad (1)$$

where y is a dependent variable of a regression equation, $\mathbf{X}$ is a matrix of independent variable of a regression equation, F is a function of a regression equation, θ is a parameter vector of a regression equation, ϵ is an error of a regression equation, σ is a standard deviation, $N(0, \sigma^2)$ is a Normal distribution with mean at zero and standard deviation at σ.

To calculate the parameter covariance matrix from the results of Eq. (1), other terms such as the sum of squared errors and the Jacobian of function $F(\mathbf{X}, \theta)$ in Eq. (1) are needed. From the parameter estimation, the weighted sum of squared errors SSE between the experimental and predicted data (y^{exp} and $F(\mathbf{X}, \theta)$) can be quantified as Eq. (2).

$$SSE = \Sigma_i w_i \cdot (y_i^{\text{exp}} - F(\theta))^2 \qquad (2)$$

where SSE is the sum of squares of errors, exp means experimental data, and w_i is the weighting factor in multi-variable non-linear regression

The Jacobian of function $F(\mathbf{X}, \theta)$ represents the local sensitivity of the property model $F(\mathbf{X}, \theta)$ with respect to the parameter values θ. It can be calculated by taking partial derivatives of $F(\mathbf{X}, \theta)$ as shown in Eq. (3).

$$J(\mathbf{X}, \theta) = \begin{bmatrix} \frac{\partial F(x_1, \theta)}{\partial \theta_1} & \cdots & \frac{\partial F(x_1, \theta)}{\partial \theta_k} & \cdots & \frac{\partial F(x_1, \theta)}{\partial \theta_m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F(x_i, \theta)}{\partial \theta_k} & \cdots & \frac{\partial F(x_i, \theta)}{\partial \theta_k} & \cdots & \frac{\partial F(x_i, \theta)}{\partial \theta_m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F(x_n, \theta)}{\partial \theta_1} & \cdots & \frac{\partial F(x_n, \theta)}{\partial \theta_k} & \cdots & \frac{\partial F(x_n, \theta)}{\partial \theta_m} \end{bmatrix} \qquad (3)$$

Its adaptation and calculation steps for EOS will be discussed in later sections in detail. The parameter w_i is the weighting factor of the residuals $(y_i^{\text{exp}} - F(\theta))$. When there is no weighting needed, w_i is assigned to be 1. In multivariate non-linear regression,[27] when fitting the experimental data of different properties and orders of magnitude (e.g., pressure and temperature), a weighting factor is needed to normalize the residuals, and w_i will be set to values such as $1/(y_i^{\text{exp}})^2$.

The covariance matrix of the parameters is in its general form written as a function of the Jacobian matrix, $J(\mathbf{X}, \theta^*)$, and the measurement

uncertainty matrix, V:[27]

$$COV(\theta^*) = (J(\mathbf{X}, \theta^*)^{\mathrm{T}} \cdot V^{-1} \cdot J(\mathbf{X}, \theta^*))^{-1} \quad (4)$$

$J(\mathbf{X}, \theta^*)$ expresses the derivatives of $F(\mathbf{X}, \theta)$ with respect to all parameters and is evaluated at parameter estimates θ^*. V is a matrix containing the measurement uncertainty for each data point as diagonal elements. Furthermore, correlation information between the respective measurement values can be represented as off-diagonal elements. The elements of V are represented as variances σ^2, where $\sigma_{k,k}$ would represent the standard deviation of measurement k and n the number of measured data points:

$$V = \begin{bmatrix} \sigma_{1,1}^2 & \dots & \sigma_{1,k}^2 & \dots & \sigma_{1,n}^2 \\ \dots & \dots & \dots & \dots & \dots \\ \sigma_{1,k}^2 & \dots & \sigma_{i,k}^2 & \dots & \sigma_{k,n}^2 \\ \dots & \dots & \dots & \dots & \dots \\ \sigma_{n,1}^2 & \dots & \sigma_{n,k}^2 & \dots & \sigma_{n,n}^2 \end{bmatrix} \quad (5)$$

If accurate measurement uncertainties are available, Eq. (5) can be used directly. However, correlation information is often not available for the measurement values or it is assumed that the measurements were uncorrelated. Hence, V is approximated as a diagonal matrix with measurement uncertainties:

$$V = \begin{bmatrix} \sigma_{1,1}^2 & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \sigma_{k,k}^2 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \sigma_{n,n}^2 \end{bmatrix} \quad (6)$$

$$V = \mathrm{diag}[\sigma_{1,1}^2, \dots, \sigma_{k,k}^2, \dots, \sigma_{n,n}^2] \quad (7)$$

If the measurement uncertainties are not properly reported for all data points, assuming arbitrary uncertainties can lead to over- or underestimated uncertainties in θ and y. In order to produce a simple assumption for the value of σ^2, which is the variance of the errors ϵ, σ^2 can be estimated by the sum of squared errors as shown in Eq. (8).

$$\sigma^2 \approx \frac{SSE}{n - m} \quad (8)$$

where n is the number of data points and m is the number of parameters in a regression equation

In Eq. (8), SSE is the value of the sum of squared errors of the objective function obtained from the least-squares parameter estimation method, n is the number of data points, and m is the number of parameters.

Hence, the covariance matrix of parameters can be re-written by Eq. (9).[27]

$$COV(\theta^*) = \sigma^2 (J(\mathbf{X}, \theta^*)^{\mathrm{T}} J(\mathbf{X}, \theta^*))^{-1} \quad (9)$$

where COV is a covariance matrix, θ^* is an estimated parameter vector of a regression equation and J is a Jacobian matrix.

In Eq. (9), SSE is the value of the sum of squared errors objective function obtained from the least-squares parameter estimation method, n is the number of data points, and m is the number of parameters.

With $COV(\theta^*)$, the interdependence of the parameters in the property model $F(\mathbf{X}, \theta)$ can be quantified. This is done by calculating the corresponding elements of the parameter correlation matrix obtained by Eq. (10).

$$Corr(\theta_i^*, \theta_j^*) = \frac{COV(\theta_i^*, \theta_j^*)}{\sqrt{Var(\theta_i^*) Var(\theta_j^*)}} \quad (10)$$

where $Corr$ means correlated coefficient and Var is a Variance

In Eq. (10), $COV(\theta_i^*, \theta_j^*)$ is the respective element of the covariance matrix, and $Var(\theta_i^*)$ and $Var(\theta_j^*)$ are the variances of the respective parameters.

The covariance matrix of the property model predictions can be approximated by linear error propagation through the Jacobian of a vector of independent variables $\vec{x}$ and the covariance of the parameter estimates as shown in Eq. (11).

$$COV(y^{\mathrm{pred}}) = J(\vec{x}, \theta^*) COV(\theta^*) J(\vec{x}, \theta^*)^{\mathrm{T}} \quad (11)$$

where pred means predicted values.

With the covariance matrices of the param-

eters and the property model predictions, the uncertainty of the parameters and the property model predictions can be quantified using the confidence intervals of the parameters and property model predictions. If the assumptions behind the model are satisfied (as assumed in the previous steps), the parameter estimates will follow a Student t-distribution and the confidence interval of parameters can be expressed as Eq. (12).

$$\theta^*_{1-\gamma_t/2} = \theta^* \pm \sqrt{\mathrm{diag}(COV(\theta^*))} \cdot t(n-m, \gamma_t/2) \quad (12)$$

where γ_t is the Student t distribution percentile and diag is a vector formed by the diagonal entries of a matrix.

Similarly, the confidence intervals of the property predictions are given by Eq. (13).

$$\begin{aligned} & y^{\mathrm{pred}}_{1-\gamma_t/2} - y^{\mathrm{pred}} \\ = \; & \pm\sqrt{\mathrm{diag}(COV(y^{\mathrm{pred}}))} \cdot t(n-m, \gamma_t/2) \end{aligned} \quad (13)$$

In Eqs. (12) and (13), $t(n-m, \gamma_t/2)$ is the Student t-distribution value corresponding to the $\gamma_t/2$ percentile of the Student t-distribution, $\mathrm{diag}(COV(\theta^*))$ represents the diagonal elements of $COV(\theta^*)$, and $\mathrm{diag}(COV(y^{\mathrm{pred}}))$ are the corresponding diagonal elements of $COV(y^{\mathrm{pred}})$.

The confidence intervals from Eqs. (12) and (13) can be regarded as the uncertainty of the parameters and the property model predictions, respectively, in very simple regression models with uncorrelated parameters. They quantify the range of possible outcomes of the EOS if the regression process is repeated with data points obtained at other experimental conditions. However, since all parameters are correlated with each other to some degree, the basic theory is too simple for complex models like EOS which estimate different types of outputs based upon various types of experimental data. The forthcoming sections describe how the theory can be applied to the EOS to calculate its uncertainty.

The authors would also like to highlight the problem of systematic measurement errors. The assumption of ideally and independently distributed measurement errors (according to Eq. (1)) would correspond to completely uncorrelated measurement data (white noise). However, it is likely that data points from the same data source will be correlated and systematically higher or lower than another source. These systematic errors depend on many factors such as the experimental set-up, the location and the experimentalists themselves. In the current methodology these correlated measurement errors have not been systematically incorporated, since it is out of the scope of the current work.

3 Equation of state and its uncertainty calculation

The thermodynamic properties of pure fluids were first calculated by using the ideal gas laws and variations of ideal gas laws.[30,31] In recent years, EOS have been developed that are based on the fundamental formulation of the non-dimensionalized Helmholtz energy $\alpha = a/(RT)$ with temperature T and density ρ as independent variables. Extensive literature is available on these highly flexible empirical multiparameter equations of state.[9,12,31] The detailed formulations and derivations for the equations of state used in this paper can be found in the supporting material, where all the property equations and variables of the EOS are listed.

3.1 Uncertainty of properties with T and ρ as inputs

We use the following formulation of the non-dimensionalized Helmholtz energy α with temperature and density as independent variables:

$$\begin{aligned} \alpha(\left[\theta_{\mathrm{EOS}}, T_c, \rho_c\right], T, \rho) = & \alpha^0(\left[\theta_{\mathrm{EOS}}, T_c, \rho_c\right], T, \rho) \\ & + \alpha^r(\left[\theta_{\mathrm{EOS}}, T_c, \rho_c\right], T, \rho) \end{aligned} \quad (14)$$

where α is the dimensionless Helmholtz energy, α^0 is the ideal gas component of dimensionless Helmholtz energy, α^r is the dimensionless Helmholtz energy due to intermolecular

forces, δ the is reduced density, τ is the reciprocal of reduced temperature and EOS means equation of state.

We use the notation $\left[\theta_{\mathrm{EOS}}, T_{\mathrm{c}}, \rho_{\mathrm{c}}\right]$ to clearly indicate that θ_{EOS}, T_{c} and ρ_{c} are model parameters of the EOS. ρ_{c} is the density at the critical point and T_{c} the critical temperature. In some EOS (one example would be that of refrigerant R-134a[32]), a reference state other than the critical point is used to reduce the temperature and density, but that is not further discussed here. In this work, we consider the EOS of propane of Lemmon *et al.*[12] and the EOS parameters θ_{EOS} can be obtained from the same work. From the non-dimensionalized Helmholtz energy EOS other fluid properties can be derived and calculated given temperature T in K, density ρ in mol/m^3 and parameters $\left[\theta_{\mathrm{EOS}}, T_{\mathrm{c}}, \rho_{\mathrm{c}}\right]$: the pressure p in Pa, the molar internal energy $u$in J/mol, the molar enthalpy $h$in J/mol, the molar entropy s in J/mol-K, the molar Gibbs energy g in J/mol-K, the molar Helmholtz energy a in J/mol, the molar isochoric heat capacity c_v in J/mol-K, the molar isobaric heat capacity c_p in J/mol-K, the speed of sound w in m/s, the fugacity coefficient ϕ, the second virial coefficient B in m^3/mol, the third virial coefficient C in m^6/mol, and the ideal gas isobaric heat capacity c_{p0} in J/mol-K.[12,31]

In the literature, Eq. (14) is conventionally written in terms of the reduced density $\delta = \rho/\rho_{\mathrm{c}}$ and reciprocal reduced temperature $\tau = T_{\mathrm{c}}/T$, where δ and τ are the independent variables of the EOS and θ_{EOS} are the parameters. However, there are usually considerable differences among investigators in the measurement of the critical temperature T_{c} and critical density ρ_{c} as shown by Lemmon *et al.*[12] This means that it is necessary to take into account the measurement uncertainties of the critical point when calculating the uncertainty of the EOS outputs.

The covariance-based uncertainty analysis method, outlined in Section 2, is now described for the Helmholtz-based EOS. Equation (14) can be written analogously to Eq. (1) assuming ideally and independently distributed errors defined by a Gaussian distribution white noise.

$$y = F(\mathbf{X}, \vartheta) + \epsilon \qquad \epsilon \sim N(0, \sigma^2) \tag{15}$$

$$\mathbf{y} := [p, u, h, s, g, a, c_v, c_p, \omega, \phi, B, C, c_{p0}] \tag{16}$$

$$\mathbf{X} := (T, \rho) \tag{17}$$

$$\vartheta := [\theta_{\mathrm{EOS}}, T_{\mathrm{c}}, \rho_{\mathrm{c}}] \tag{18}$$

T_c and ρ_c are defined according to the experimental results of thermodynamic properties around the critical point, and we assume that the parameter estimation for θ_{EOS} (parameter optimization and fitting to a large set of experimental data) has been completed by the developers of the EOS . The set of parameters called θ^*_{EOS} is known from the literature.[11] Hence, the parameter estimates θ^*_{EOS} are used in this study for the uncertainty analysis and re-training of the data is not necessary.

The parameters of the EOS are given by $\theta_{\mathrm{EOS}} = [\theta_1, ..., \theta_j, ..., \theta_m]$, with m being the number of parameters.

However, experimental data need to be collected or taken from the work of the developers of the EOS because it is needed to calculate the errors (differences between experimental data and predicted values by the EOS). For example, in the case of propane, the EOS has been fitted to experimental data of p, p_{sat}, c_v, c_p, w, B, c_{p0} as functions of T and ρ. These variables are called fitted (or observed) variables because some of their values are measured and included in the fitting process. Other variables (e.g., u, h, s, g, a, C and ϕ) are called (purely) predicted variables because none of their experimental measurements are included in the fitting process of the EOS. Some EOSs may have different sets of fitted variables and predicted variables depending on experimental data availability, but the proposed method focuses on using the aforementioned set of fitted and predictive variables for simplicity. The parameter covariance matrix (i.e. the uncertainty information of the parameter estimates θ^*_{EOS}) is obtained

through the use of data. The uncertainties of T, ρ, p, c_v, c_p, w, B and c_{p0} as well as u, h, s, C, ϕ, a and g estimated by the EOS are then obtained by linear error propagation.

For the fitted properties, there are a number of n experimental data points respectively, e.g. $p^{\mathrm{exp}} = [p_1^{\mathrm{exp}}, ..., p_{n_p}^{\mathrm{exp}}]$, $c_v^{\mathrm{exp}} = [c_{v\,1}^{\mathrm{exp}}, ..., c_{v\,n_{c_v}}^{\mathrm{exp}}]$, etc.

We describe specifically how the uncertainty analysis described in Section 2 can be applied to the estimation of the uncertainty of the predictions of EOS outputs. We will use the notation introduced above for $\alpha(\left[\theta_{\mathrm{EOS}}, T_{\mathrm{c}}, \rho_{\mathrm{c}}\right], T, \rho)$, where the parameters for the EOS are θ_{EOS}, T_{c} and ρ_{c}, and the independent variables are T and ρ.

Equation (9) described the covariance matrix of the parameters determined from the variances of errors σ^2 and the Jacobian $J(\theta^*)$, where σ^2 is obtained from the sum of squared errors SSE of the objective function from the least-square parameter estimation. Hence, in the case of the fundamental EOS for propane, as described by Lemmon *et al.*,[12] SSE is described as the sum of squared relative errors – the difference between the predicted and experimental property value divided by the experimental value (least square regression). Hence, for the example of the fitted property c_p, its variance σ^2 should be written as

$$\sigma_{c_p}^2 \approx \frac{SSE_{c_p}}{n_{c_p} - m_{c_p}} \tag{19}$$

where n_{c_p} is the number of data points for c_p and m_{c_p} is the number of EOS parameters that are needed to calculate c_p using the EOS (note $n_{c_p} - m_{c_p} > 0$). SSE_{c_p} is the value of the sum of squared relative errors objective function from the parameter estimation.[12] In order to obtain SSE_{c_p} the parameter estimation does not need to be retrained:

$$SSE_{c_p} = \sum_{i=1}^{n_{c_p}} \left[\frac{{c_{p}}_i^{\mathrm{exp}} - c_p\left(\theta_{\mathrm{EOS}}^*, \tau_i, \delta_i\right)}{{c_{p}}_i^{\mathrm{exp}}} \right]^2 \tag{20}$$

where i is the index for the experimental condition (T_i, ρ_i), ${c_{p}}_i^{\mathrm{exp}}$ is the experimental value of specific heat and $c_p(\theta_{\mathrm{EOS}}^*, \tau_i, \delta_i)$ is the predicted specific heat value. The estimates for variances of the other fitted properties are obtained for $\sigma_{c_v}^2$, σ_w^2, σ_B^2, $\sigma_{c_{p0}}^2$ and $\sigma_{p_{\mathrm{c}}}^2$. The estimate for the error of T_{c} and ρ_{c} ($\sigma_{T_{\mathrm{c}}}^2$ and $\sigma_{\rho_{\mathrm{c}}}^2$) can be calculated from the standard deviation of the experimental measurements. Lemmon *et al.*[12] used a different objective function for the residual errors in the pressure as shown in Eq. (21), in order to obtain similar magnitudes of the liquid and vapor phase.

$$SSE_p = \sum_i \left[\frac{p_i^{\mathrm{exp}} - p(\theta_{\mathrm{EOS}}^*, \tau_i, \delta_i)}{\rho_i^{\mathrm{exp}} \left(\dfrac{\partial p_i}{\partial \rho_i} \right)_T} \right]^2 \tag{21}$$

where p_i^{exp} is the experimental pressure value, $p(\theta_{\mathrm{EOS}}^*, \tau_i, \delta_i)$ is the predicted pressure value, ρ_i^{exp} is the experimental density value and $\left.\dfrac{\partial p_i}{\partial \rho_i}\right|_T$ is the partial derivative of the pressure with respect to the density calculated at θ_{EOS}^*.

In order for the uncertainty assessment of the EOS to be consistent with the EOS itself, the exact objective function used in the training of the EOS is needed, but this is very difficult to achieve in practice. The state-of-the-art fitting process includes addition and removal of data points and constraints in an iterative fashion, and it is not possible to obtain the weights that were ultimately used in the regression process. Therefore, the weights that were used in the fitting process of Lemmon *et al.*[12] are unknowable, and an estimation of the objective function (i.e. the corresponding SSE) is required.

While Eq. (21) is appropriate for pressure data points with density and temperature as independent variables, it is not suitable for vapor pressure data. For vapor pressure data, the desired residue would be

$$SSE_{p_{\mathrm{sat}}} = \sum_i \left[\frac{p_i^{\mathrm{exp}} - p_{\mathrm{sat}}(\theta_{\mathrm{EOS}}^*, \tau_i)}{p_i^{\mathrm{exp}}} \right]^2, \tag{22}$$

though this form is not suitable because the evaluation of p^{sat} requires an iterative solution for the vapor pressure, which is a calculation fraught with potential perils. For that reason, it is preferrable, and common practice, to mini-

mize the difference in Gibbs energy between the saturated liquid and vapor phases for the given temperature and pressure (see for instance Bell et al.[33]). In this case, the densities must be solved for in each phase, but the full resolution of the Maxwell criteria is not required.

To describe the minimization process in terms of SSE, Eq. (23) is used with vapor pressure data.

$$SSE_g = \sum_i \left[\frac{g(\theta^*_{\mathrm{EOS}}, \tau_i, \frac{\rho_{\mathrm{min},i}}{\rho_c}) - g(\theta^*_{\mathrm{EOS}}, \tau_i, \frac{\rho_{\mathrm{max},i}}{\rho_c})}{RT_i^{\mathrm{exp}}} \right]^2 \tag{23}$$

$$\rho_{\mathrm{min},i} = \min(\vec{\rho}(T_i^{\mathrm{exp}}, p_i^{\mathrm{exp}})) \tag{24}$$

$$\rho_{\mathrm{max},i} = \max(\vec{\rho}(T_i^{\mathrm{exp}}, p_i^{\mathrm{exp}})) \tag{25}$$

where ρ_{min} is the minimum density calculated by iteratively solving for density from the specified temperature and pressure, ρ_{max} is the maximum density calculated from the same iterative calculation of density, R is the universal gas constant, T_i^{exp} is the experimental temperature reading in the vapor pressure data.

One may wonder if the SSE in Eq. (23) is suitable to describe the deviation in pressure differences between estimated and measured vapor pressure. To verify that Eq. (23) is approximately equivalent to Eq. (22), the relative deviation between the estimated and pressure pressure is plotted with the difference of the Gibbs energy of saturated liquid and vapor divided by the product of the gas constant and T, as well as more accurate approximations to Eq. (22). The derivations in the supplemental material explain how this non-dimensionalization can be obtained, following the assumptions that a) a first-order series expansion of Gibbs energy with respect to pressure difference can be used, b) the vapor phase derivative of Gibbs energy with respect to pressure at constant temperature is much greater in magnitude than that of the liquid phase, and c) the vapor phase can be treated as an ideal gas.

Figure 1 presents numerical values for each of the approximations to the vapor pressure residue of propane. This figure demonstrates that $\Delta g/(RT)$ provides a fair representation of the saturation pressure residue although the second-order expansion yields a superior evaluation of the vapor pressure residue.

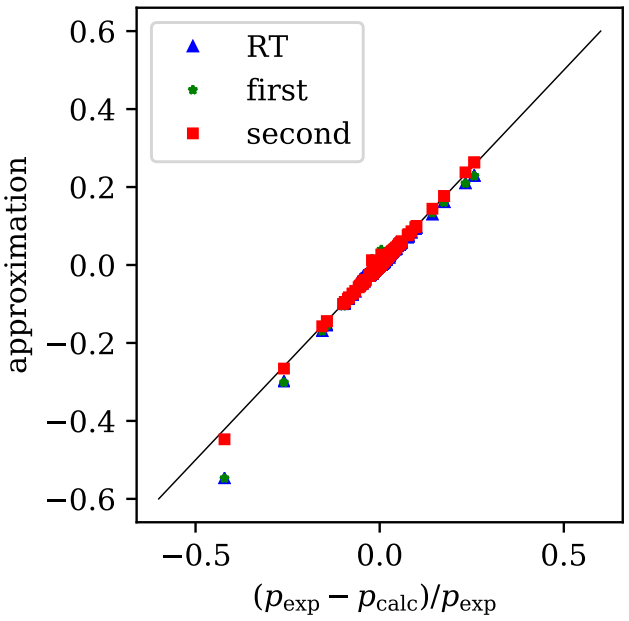


Figure 1: Comparing three different approximations to the deviation in pressure with relative error in pressure calculation. For more information, see the derivations in the supplemental material (RT: $res = \Delta g/(RT)$, first: a first-order expansion in pressure difference, dropping the liquid derivative without any further simplifications, second: a second order expansion in pressure difference without any further simplifications)

The results show that the difference of Gibbs energy can effectively describe the difference between estimated and measured vapor pressure. However, since the differences of pressure are not really used in the estimation of the coefficients in the Helmholtz-energy-based EOS, Eq. (23) is used in this study.

A diagonal matrix containing all the error variances can then be built as follows in Eq. (26).

The V matrix in Eq. (26) contain $2n_{p_v}$ number of σ_g^2 because each σ_g^2 is calculated by two predicted values: the predicted values of Gibbs energy of saturated vapor and the predicted values of Gibbs energy of saturated liquid.

$$V = \mathrm{diag}[\underbrace{\sigma_p^2, ..., \sigma_p^2}_{n_p}, \underbrace{\sigma_g^2, ..., \sigma_g^2}_{2n_{p_v}}, \underbrace{\sigma_{c_v}^2, ..., \sigma_{c_v}^2}_{n_{\mathrm{cv}}}, \underbrace{\sigma_{c_p}^2, ..., \sigma_{c_p}^2}_{n_{\mathrm{cp}}}, \underbrace{\sigma_w^2, ..., \sigma_w^2}_{n_w}, \underbrace{\sigma_B^2, ..., \sigma_B^2}_{n_B}, \underbrace{\sigma_{n_{c_{p0}}}^2, ..., \sigma_{n_{c_{p0}}}^2}_{n_{c_{p0}}}, \underbrace{\sigma_{T_\mathrm{c}}^2, ..., \sigma_{T_\mathrm{c}}^2}_{n_{T_\mathrm{c}}}, \underbrace{\sigma_{\rho_\mathrm{c}}^2, ..., \sigma_{\rho_\mathrm{c}}^2}_{n_{\rho_\mathrm{c}}}] \tag{26}$$

The Jacobian for all of the fitted properties needs to be calculated at the specific experimental measurement point of the variables T and ρ (or τ and δ respectively for their dimensionless form), and at the specific fitted parameter values θ^*_{EOS}. However, since the variances σ_p^2, σ_g^2, $\sigma_{c_p}^2$, $\sigma_{c_v}^2$, σ_w^2, σ_B^2, $\sigma_{c_{p0}}^2$, $\sigma_{T_\mathrm{c}}^2$ and $\sigma_{p_\mathrm{c}}^2$ used relative errors, the elements of the Jacobians need to be normalized by their corresponding experimental value. As an example, the general analytical expression of the Jacobian of p can be written as Eq. (27).

While the Jacobians of an experimental data point of most fitted variables can be calculated in a similar way as Eq. (27), the Jacobian of each experimental observation of vapor pressure are calculated differently because they have two predicted values of Gibbs energy in each difference term in Eq. (23). To involve the partial derivatives of both predicted values in the SSE_g, the Jacobian of vapor pressure is a 2 × (m+2) matrix as Eq. (28)

The first row of the Jacobian in Eq. (28) corresponds to the predicted values of Gibbs energy of the maximum density value at the vapor pressure, and the second row of the Jacobian corresponds to the predicted values of Gibbs energy of the minimum density values at the corresponding temperature.

The Jacobian of the fitted data of p at multiple experimental data points of τ and δ, evaluated at the parameter estimates θ^*_{EOS} is then given by Eq. (29), where $J_p(\theta^*_{\mathrm{EOS}}, \tau, \delta)$ is a n_p x (m+2) matrix with $\theta^*_{\mathrm{EOS}} = [\theta_1^*, ..., \theta_m^*]$, $\tau = [\tau_1, ..., \tau_{n_p}]$ and $\delta = [\delta_1, ..., \delta_{n_p}]$. p_i^{exp} are the corresponding experimental values used to normalize the Jacobian. The method to calculate the partial derivatives is discussed in the Supplementary Materials for reference.

Assuming that the uncertainties of critical properties have no effect on the estimation of the coefficients θ_{EOS} and do not propagate through their estimation process, the Jacobian of p evaluated at the experimental values can be simplified by setting the last two columns of values in Eq. (29) as zeros.

It is important to notice that J_p is evaluated at the parameter estimates θ^*_{EOS}, but is a function of T and ρ. In analogy, the Jacobian for the other fitted data can be obtained, giving J_{c_v}, J_{c_p}, J_w, J_B and $J_{c_{p0}}$. Since we are accounting for the effect of the uncertainties of critical densities and critical temperature, their Jacobians are given as

$$\begin{aligned} J_{T_\mathrm{c}} &= \begin{bmatrix} 0_{1\mathrm{x}m} & \frac{\partial T_\mathrm{c}}{\partial T_{\mathrm{c},1}}\frac{1}{T_{\mathrm{c},1}} & 0 \\ \cdots & \cdots & \cdots \\ 0_{1\mathrm{x}m} & \frac{\partial T_\mathrm{c}}{\partial T_{\mathrm{c},i}}\frac{1}{T_{\mathrm{c},i}} & 0 \\ \cdots & \cdots & \cdots \\ 0_{1\mathrm{x}m} & \frac{\partial T_c}{\partial T_{\mathrm{c},n_{T_\mathrm{c}}}}\frac{1}{T_{\mathrm{c},n_{T_\mathrm{c}}}} & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0_{1\mathrm{x}m} & \frac{1}{n_{T_\mathrm{c}}T_{\mathrm{c},1}} & 0 \\ \cdots & \cdots & \cdots \\ 0_{1\mathrm{x}m} & \frac{1}{n_{T_\mathrm{c}}T_{\mathrm{c},i}} & 0 \\ \cdots & \cdots & \cdots \\ 0_{1\mathrm{x}m} & \frac{1}{n_{T_\mathrm{c}}T_{\mathrm{c},n_{T_\mathrm{c}}}} & 0 \end{bmatrix} \end{aligned} \tag{30}$$

$$\begin{aligned} J_{\rho_\mathrm{c}} &= \begin{bmatrix} 0_{1\mathrm{x}m} & 0 & \frac{\partial \rho_\mathrm{c}}{\partial \rho_{\mathrm{c},1}}\frac{1}{\rho_{\mathrm{c},1}} \\ \cdots & \cdots & \cdots \\ 0_{1\mathrm{x}m} & 0 & \frac{\partial \rho_\mathrm{c}}{\partial \rho_{\mathrm{c},i}}\frac{1}{\rho_{\mathrm{c},i}} \\ \cdots & \cdots & \cdots \\ 0_{1\mathrm{x}m} & 0 & \frac{\partial \rho_\mathrm{c}}{\partial \rho_{\mathrm{c},n_{\rho_\mathrm{c}}}}\frac{1}{\rho_{\mathrm{c},n_{\rho_\mathrm{c}}}} \end{bmatrix} \\ &= \begin{bmatrix} 0_{1\mathrm{x}m} & 0 & \frac{1}{n_{\rho_\mathrm{c}}\rho_{\mathrm{c},1}} \\ \cdots & \cdots & \cdots \\ 0_{1\mathrm{x}m} & 0 & \frac{1}{n_{\rho_\mathrm{c}}\rho_{\mathrm{c},i}} \\ \cdots & \cdots & \cdots \\ 0_{1\mathrm{x}m} & 0 & \frac{1}{n_{\rho_\mathrm{c}}\rho_{\mathrm{c},n_{\rho_\mathrm{c}}}} \end{bmatrix}. \end{aligned} \tag{31}$$

The combined Jacobian of the fitted data p, p_v, c_v, c_p, w, B and c_{p0} can constructed as in Eq. (32).

The Jacobian $J_{\mathrm{tot}}(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho, p)$ is a $(n_p+2n_{p_v}+n_{c_v}+n_{c_p}+n_w+n_B+n_{c_{p0}}+n_{T_\mathrm{c}}+n_{\rho_\mathrm{c}})$

$$J_p(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T_i, \rho_i) = \frac{1}{p_i^{\exp}} \left[\frac{\partial p}{\partial \theta_1}(\theta^*_{\mathrm{EOS}}, T_i, \rho_i) \quad \dots \quad \frac{\partial p}{\partial \theta_M}(\theta^*_{\mathrm{EOS}}, T_i, \rho_i) \quad \frac{\partial p}{\partial T_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T_i, \rho_i) \quad \frac{\partial p}{\partial \rho_c}(\theta^*_{\mathrm{EOS}}, T_i, \rho_i)\right] \tag{27}$$

$$J_{p_v}(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T_i, p_i) = \frac{1}{RT_i} \times \begin{bmatrix} \frac{\partial g}{\partial \theta_1}(\theta^*_{\mathrm{EOS}}, T_i, \rho_{\min,i}) & \dots & \frac{\partial g}{\partial \theta_M}(\theta^*_{\mathrm{EOS}}, T_i, \rho_{\min,i}) & \frac{\partial g}{\partial T_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T_i, \rho_{\min,i}) & \frac{\partial g}{\partial \rho_c}(\theta^*_{\mathrm{EOS}}, T_i, \rho_{\min,i}) \\ \frac{\partial g}{\partial \theta_1}(\theta^*_{\mathrm{EOS}}, T_i, \rho_{\max,i}) & \dots & \frac{\partial g}{\partial \theta_M}(\theta^*_{\mathrm{EOS}}, T_i, \rho_{\max,i}) & \frac{\partial g}{\partial T_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T_i, \rho_{\max,i}) & \frac{\partial g}{\partial \rho_c}(\theta^*_{\mathrm{EOS}}, T_i, \rho_{\max,i}) \end{bmatrix} \tag{28}$$

$$J_p = \begin{bmatrix} \frac{\frac{\partial p}{\partial \theta_1}\big|_{\theta^*_{\mathrm{EOS}}, T_1, \rho_1}}{p_1^{\exp}} & \dots & \frac{\frac{\partial p}{\partial \theta_m}\big|_{\theta^*_{\mathrm{EOS}}, T_1, \rho_1}}{p_1^{\exp}} & \frac{\frac{\partial p}{\partial \tau_1}\big|_{\theta^*_{\mathrm{EOS}}, T_1, \rho_1} \frac{\partial \tau_1}{\partial T_\mathrm{c}}}{p_1^{\exp}} & \frac{\frac{\partial p}{\partial \delta_1}\big|_{\theta^*_{\mathrm{EOS}}, T_1, \rho_1} \frac{\partial \delta_1}{\partial \rho_\mathrm{c}}}{p_1^{\exp}} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\frac{\partial p}{\partial \theta_1}\big|_{\theta^*_{\mathrm{EOS}}, T_i, \rho_i}}{p_i^{\exp}} & \dots & \frac{\frac{\partial p}{\partial \theta_m}\big|_{\theta^*_{\mathrm{EOS}}, T_i, \rho_i}}{p_i^{\exp}} & \frac{\frac{\partial p}{\partial \tau_i}\big|_{\theta^*_{\mathrm{EOS}}, T_i, \rho_i} \frac{\partial \tau_i}{\partial T_\mathrm{c}}}{p_i^{\exp}} & \frac{\frac{\partial p}{\partial \delta_i}\big|_{\theta^*_{\mathrm{EOS}}, T_i, \rho_i} \frac{\partial \delta_i}{\partial \rho_\mathrm{c}}}{p_i^{\exp}} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\frac{\partial p}{\partial \theta_1}\big|_{\theta^*_{\mathrm{EOS}}, T_{n_p}, \rho_{n_p}}}{p_{n_p}^{\exp}} & \dots & \frac{\frac{\partial p}{\partial \theta_m}\big|_{\theta^*_{\mathrm{EOS}}, T_{n_p}, \rho_{n_p}}}{p_{n_p}^{\exp}} & \frac{\frac{\partial p}{\partial \tau_{n_p}}\big|_{\theta^*_{\mathrm{EOS}}, T_{n_p}, \rho_{n_p}} \frac{\partial \tau_{n_p}}{\partial T_\mathrm{c}}}{p_{n_p}^{\exp}} & \frac{\frac{\partial p}{\partial \delta_{n_p}}\big|_{\theta^*_{\mathrm{EOS}}, T_{n_p}, \rho_{n_p}} \frac{\partial \delta_{n_p}}{\partial \rho_\mathrm{c}}}{p_{n_p}^{\exp}} \end{bmatrix} \tag{29}$$

$$J_{\mathrm{tot}}(\left[\theta^*_{\mathrm{EOS}}, T_c, \rho_\mathrm{c}\right], T, \rho, p) = \begin{bmatrix} J_p(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho) \\ J_{p_\mathrm{v}}(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, p) \\ J_{c_v}(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho) \\ J_{c_p}(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho) \\ J_w(\left[\theta^*_{\mathrm{EOS}}, T_c, \rho_\mathrm{c}\right], T, \rho) \\ J_B(\left[\theta^*_{\mathrm{EOS}}, T_c, \rho_\mathrm{c}\right], T) \\ J_{c_{p0}}(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T) \\ J_{T_\mathrm{c}} \\ J_{\rho_\mathrm{c}} \end{bmatrix} \tag{32}$$

x $(m+2)$ matrix. J_{p_v} are functions of T and p because their SSE in Eq. (23) have no density values as inputs. J_B and $J_{c_{p0}}$ are not functions of δ because B and c_{p0} depend on temperature only.

In Eq. (32), with the exception of J_{p_v}, all output variables are assumed to be associated with measurements of temperature and density. However, in actual experiments, other sets of variables like temperature and pressure are more natural independent variables. Hence, conversion for the measurement data point is needed before it can be used to calculate the Jacobian. For data with temperature and pressure measurement but unknown density values, the density values are calculated from the temperature and pressure measurement via the EOS. The conversion method for density of different variables is outlined in Table 2.

According to Eq. (9), the covariance matrix can be given as in Eq. (33).

The size of $COV(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right])$ depends on the number of parameters in the EOS. For example, the EOS of propane has 97 parameters that are obtained from measured data in the literature. Taking into account T_c and ρ_c, its

$$COV(\left[\theta^*_{EOS}, T_\mathrm{c}, \rho_\mathrm{c}\right]) = \left[J_\mathrm{tot}(\left[\theta^*_\mathrm{EOS}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho)^\mathrm{T} \cdot V^{-1} \cdot J_\mathrm{tot}(\left[\theta^*_\mathrm{EOS}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho)\right]^{-1} \tag{33}$$

Table 2: Methods of data conversion for density values in Jacobian calculation

Type of measurement	Density values
c_{p0}	Not needed
p, ρ, T	From ρ measurement only
B	Not needed
w, c_p and c_v	Calculated by T and p from EOS

covariance matrix is a 99×99 square matrix.[12] The EOS of refrigerant R-22 has 55 parameters and its covariance matrix is a 57×57 square matrix.[9]

Following Eq. (9), the variance σ_p^2 is only multiplied with the elements of $J_p(\theta^*_\mathrm{EOS}, \tau, \delta)$; the same holds for the other properties. In order to illustrate this, it is possible to re-write Eq. (32) in the following way

$$V^{-1} \cdot J_\mathrm{tot}(\left[\theta^*_\mathrm{EOS}, T_\mathrm{c}, \rho_c\right], T, \rho, p) = \begin{bmatrix} 1/\sigma_p^2 \cdot J_p \\ 1/\sigma_g^2 \cdot J_{p_v} \\ 1/\sigma_{c_v} \cdot J_{c_v} \\ 1/\sigma_{c_p}^2 \cdot J_{c_p} \\ 1/\sigma_w^2 \cdot J_w \\ 1/\sigma_B^2 \cdot J_B \\ 1/\sigma_{c_{p0}}^2 \cdot J_{c_{p0}} \\ 1/\sigma_{T_\mathrm{c}}^2 \cdot J_{T_\mathrm{c}} \\ 1/\sigma_{\rho_\mathrm{c}}^2 \cdot J_{\rho_\mathrm{c}} \end{bmatrix} \tag{34}$$

The variances σ are essentially weighting factors for the elements of the Jacobian.

To avoid refitting the EOS for the covariance between the coefficients and the critical values, it is assumed that the uncertainties of critical properties have a negligible effect on the estimation of the coefficients θ_EOS. The covariance can be written as

$$\begin{aligned} &COV(\left[\theta^*_\mathrm{EOS}, T_\mathrm{c}, \rho_\mathrm{c}\right]) \\ &= \begin{bmatrix} COV(\theta^*_\mathrm{EOS}) & 0 & 0 \\ 0 & \frac{\sigma^2_{T_\mathrm{c}}}{T_\mathrm{c}^2} & 0 \\ 0 & 0 & \frac{\sigma^2_{\rho_\mathrm{c}}}{\rho_\mathrm{c}^2} \end{bmatrix} \end{aligned} \tag{35}$$

where $COV(\theta^*_\mathrm{EOS})$ is the submatrix of $COV(\left[\theta^*_\mathrm{EOS}, T_\mathrm{c}, \rho_\mathrm{c}\right])$ describing the covariances of the EOS parameters but not T_c and ρ_c. $COV(\theta^*_\mathrm{EOS})$ can also be obtained directly by using Eqs. (32) and (33) when T_c and ρ_c are assumed to be known perfectly and have therefore no uncertainty.

Calculating $COV(\theta^*_\mathrm{EOS})$ based on Eq. (35), it is possible to calculate the respective 95% confidence interval of the EOS parameter values θ_EOS

$$\begin{aligned} &\theta_{\mathrm{EOS},1-\gamma_\mathrm{t}/2} \\ &= \theta^*_\mathrm{EOS} \pm \sqrt{\mathrm{diag}(COV(\theta^*_\mathrm{EOS}))} \cdot t(n-m, \gamma_\mathrm{t}/2) \end{aligned} \tag{36}$$

In Eq. (36) $t(n-m, \gamma_\mathrm{t}/2)$ is the Student t-distribution value corresponding to the $\gamma_\mathrm{t}/2$ percentile of the Student t-distribution.

The covariance of the prediction of a certain property is obtained from the respective Jacobian and the parameter covariance matrix and is dimensionalized so that its unit should be the square of that of the property. For example, the covariance $COV(p)$ for a pressure p is given by Eq. (37) and the jacobian vector in Eq. (38).

The $COV(\left[\theta^*_\mathrm{EOS}, T_\mathrm{c}, \rho_\mathrm{c}\right])$ is a matrix with elements independent of the input variables to EOS. However, the adjusted Jacobian $j_p(\left[\theta^*_\mathrm{EOS}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho)$ depends on the value of T or ρ for the prediction. In complete analogy the covariance matrix is obtained for purely predicted properties such as entropy s by Eqs. (39)

$$\begin{aligned} COV(p) &= J_p(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho) \cdot COV(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right]) \cdot J_p(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho)^\mathrm{T} \cdot p^2 \\ &= j_p(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho) \cdot COV(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right]) \cdot j_p(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho)^\mathrm{T} \end{aligned} \tag{37}$$

$$j_p(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho) = \left[\frac{\partial p}{\partial \theta_1}(\theta^*_{\mathrm{EOS}}, T, \rho) \quad \dots \quad \frac{\partial p}{\partial \theta_M}(\theta^*_{\mathrm{EOS}}, T, \rho) \quad \frac{\partial p}{\partial T_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T, \rho) \quad \frac{\partial p}{\partial \rho_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T, \rho)\right] \tag{38}$$

and (40). The covariance matrix of other predicted properties (e.g. u, h, a, g, ϕ and C) can be calculated in a similar manner.

The calculation of the two covariance matrices is numerically not trivial. The parameter covariance matrix $COV(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right])$ is theoretically obtained through an inversion (see Eq. (33)). However, both the Jacobians and the parameter covariance matrices are very large sparse matrices and the parameter covariance matrix can also be positive semi-definite, which means it has columns that are numerically close to being linearly dependent. We recommend the following procedure to overcome these issues:

1. In order to store more significant digits of matrix elements and to allow more precise matrix operations compared to ordinary algebraic calculation of matrices, we recommend the usage of arbitrary precision methods, such as *mpmath* in the Python programming language.

2. For the calculation of the covariance matrices of the properties, e.g. $COV(p)$ or $COV(s)$, we have used LU decomposition. The calculation is shown in the supporting information.

Having obtained the respective covariance matrices of the properties, it is now possible to calculate the 95% confidence interval of the respective properties. Eqs. (41) and (42) allow for a calculation of the 95% confidence interval for pressure p and entropy s. The same holds for the other properties.

$$p^\mathrm{pred}_{1-\gamma_t/2} = p^\mathrm{pred} \pm \sqrt{\mathrm{diag}(COV(p))} \cdot t(n-m, \gamma_\mathrm{t}/2) \tag{41}$$

$$s^\mathrm{pred}_{1-\gamma_t/2} = s^\mathrm{pred} \pm \sqrt{\mathrm{diag}(COV(s))} \cdot t(n-m, \gamma_\mathrm{t}/2) \tag{42}$$

It is important to notice that in this study the sum of squared errors and the Jacobian of the fitted properties used relative values, i.e. the residuals and the derivatives have been divided by the corresponding experimental value. This corresponded to the objective function for the fitting of the parameters used by Lemmon *et al.*[12] However, this means that the covariance matrices provide the relative uncertainties, which subsequently need to be multiplied with the corresponding predicted thermodynamic property from the EOS, in order to obtain the absolute uncertainty range as shown in Eqs. (41) and (42).

Similar methods can be used to calculate the uncertainties of the differences of properties because covariance of differences or sums of properties can be calculated in a similar manner. This is especially important for properties like u, s, h, a and g, in which only the difference of values matters in EOS applications. For example, to calculate the uncertainty of the difference of two entropy values $s_1(T_1, \rho_1)$ and $s_2(T_2, \rho_2)$, one can first calculate the Jacobian of the difference between the two values by Eq. (43) and calculate the covariance using

$$COV(s) = j_s(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho) \cdot COV(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right]) \cdot j_s(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho)^\mathrm{T} \tag{39}$$

$$j_s(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho) = \left[\frac{\partial s}{\partial \theta_1}(\theta^*_{\mathrm{EOS}}, T, \rho) \quad \dots \quad \frac{\partial s}{\partial \theta_M}(\theta^*_{\mathrm{EOS}}, T, \rho) \quad \frac{\partial s}{\partial T_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T, \rho) \quad \frac{\partial s}{\partial \rho_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T, \rho)\right] \tag{40}$$

Eq. (44).

$$\begin{aligned} & j_{s_1-s_2}(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho)^\mathrm{T} \\ = & \begin{bmatrix} \frac{\partial s_1}{\partial \theta_1}(\theta^*_{\mathrm{EOS}}, T_1, \rho_1) - \frac{\partial s_2}{\partial \theta_1}(\theta^*_{\mathrm{EOS}}, T_2, \rho_2) \\ \vdots \\ \frac{\partial s_1}{\partial \theta_m}(\theta^*_{\mathrm{EOS}}, T_1, \rho_1) - \frac{\partial s_2}{\partial \theta_m}(\theta^*_{\mathrm{EOS}}, T_2, \rho_2) \\ \frac{\partial s_1}{\partial T_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T_1, \rho_1) - \frac{\partial s_2}{\partial T_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T_2, \rho_2) \\ \frac{\partial s_1}{\partial \rho_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T_1, \rho_1) - \frac{\partial s_2}{\partial \rho_\mathrm{c}}(\theta^*_{\mathrm{EOS}}, T_2, \rho_2) \end{bmatrix} \end{aligned} \tag{43}$$

The uncertainty of the difference between s_1 and s_2 can be calculated by Eq. (45).

3.2 Properties with inputs other than T and ρ

In practical applications of EOS, there are many scenarios where the inputs to the EOS are not temperature and density. For example, a user may be given the measured temperature and pressure of a fluid to find the speed of sound, and the user must use an iterative solver to find the density of the fluid first before calculating the speed of sound. For substances that are either superheated vapor or subcooled liquid, if the output variable required is neither temperature nor density, the calculation of the uncertainty of the output variable will follow that in Section 3.1 after calculating the missing temperature and density values by using an iterative solver on the EOS. However, if the output variable is either temperature or density, the uncertainty of the output variable should be calculated by propagating the uncertainty of the input variable due to the EOS to the output using the linearization of the EOS.[24,34] This is done by Eqs. (46) and (47).

$$\begin{aligned} & \Delta T^{\mathrm{pred}} \\ = & \left| \left(\frac{\partial T^{\mathrm{pred}}}{\partial x_1}\Big|_{x_2} \Delta x_1 \right) + \left(\frac{\partial T^{\mathrm{pred}}}{\partial x_2}\Big|_{x_1} \Delta x_2 \right) \right| \end{aligned} \tag{46}$$

$$\begin{aligned} & \Delta \rho^{\mathrm{pred}} \\ = & \left| \left(\frac{\partial \rho^{\mathrm{pred}}}{\partial x_1}\Big|_{x_2} \Delta x_1 \right) + \left(\frac{\partial \rho^{\mathrm{pred}}}{\partial x_2}\Big|_{x_1} \Delta x_2 \right) \right| \end{aligned} \tag{47}$$

where Δx is the standard uncertainty of variable x, T^{pred} and ρ^{pred} are the temperature and density values predicted by using an iterative solver with the EOS, x_1 and x_2 are the independent variables used to predict T^{pred} or ρ^{pred}, Δx_1 and Δx_2 are the EOS uncertainties of the variables x_1 and x_2 calculated by using the temperature and density values in Eq. (42), and the partial derivatives are obtained according to Thorade and Saadat.[35]

If any of x_1 and x_2 in Eqs. (46) and (47) are T or ρ, their corresponding uncertainties in these equations will be zero.

Since the use of Eqs. (46) and (47) is a result of the difficulty to define covariance of temperature and density, calculating the uncertainties of the sums and differences of temperature and density with equations similar to Eq. (45) is impossible. Hence the calculation of the uncertainties of the sums and differences of temperature and density is the same as that of ordinary variables and follows the method in Kline and McClintock.[24]

3.3 Saturated liquid and vapor properties

Calculation of thermodynamic properties of a fluid that is based on temperature and density

$$COV(s_1 - s_2) = j_{s_1-s_2}(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho) \cdot COV(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right]) \cdot j_{s_1-s_2}(\left[\theta^*_{\mathrm{EOS}}, T_\mathrm{c}, \rho_\mathrm{c}\right], T, \rho)^\mathrm{T} \quad (44)$$

$$s^{\mathrm{pred}}_{1,1-\gamma_\mathrm{t}/2} - s^{\mathrm{pred}}_{2,1-\gamma_t/2} = s_1 - s_2 \pm \sqrt{\mathrm{diag}(COV(s_1 - s_2))} \cdot t(n-m, \gamma_\mathrm{t}/2) \quad (45)$$

values for a homogeneous phase can be achieved by the procedure in Section 3.1. However, users of EOS are often asked for fluid properties of saturated liquid and vapor of a fluid given a temperature or pressure value only. According to Lemmon *et al.*,[12] this is achieved by finding two density values for which Gibbs energy and pressure values are equivalent at the given temperature value – this is the Maxwell's criteria. The smaller density value corresponds to the density of saturated vapor whereas the larger density value corresponds to the density of saturated liquid, and the other properties of saturated liquid and vapor can be calculated from the density values. The uncertainty of the thermodynamic properties can be calculated in the same manner as that in Section 3.1 once the density values are found. The calculation of the residuals in the Maxwell criteria is given by Eqs. (48) and (49).

$$\Delta res_1 = g(\rho_\mathrm{v}, T) - g(\rho_\mathrm{l}, T) \quad (48)$$

$$\Delta res_2 = p(\rho_\mathrm{v}, T) - p(\rho_\mathrm{l}, T), \quad (49)$$

where Δres_1 and Δres_2 are residual functions of the Maxwell criteria. By finding the density values that lead to zero values for both residual functions, the saturation densities of the fluid at a certain temperature T can be found.

The uncertainty of the Gibbs energy difference in Eq. (48) given by Δres_1 can be calculated by calculating the covariance of the Gibbs energy values from Eq. (52).

The uncertainties of the density values can then be calculated by propagating the uncertainties of the residual functions as shown in Eqs. (50) and (51).

$$\Delta\rho_\mathrm{l} = \left| \left(\frac{\partial res_1}{\partial \rho_\mathrm{l}}\bigg|_T \right)^{-1} \Delta res_1 + \left(\frac{\partial res_2}{\partial \rho_\mathrm{l}}\bigg|_T \right)^{-1} \Delta res_2 \right| \quad (50)$$

$$\Delta\rho_\mathrm{v} = \left| \left(\frac{\partial res_1}{\partial \rho_\mathrm{v}}\bigg|_T \right)^{-1} \Delta res_1 + \left(\frac{\partial res_2}{\partial \rho_\mathrm{v}}\bigg|_T \right)^{-1} \Delta res_2 \right| \quad (51)$$

However, Δres_2 in Eqs. (50) and (51) involves uncertainty of pressure of saturated liquid that changes very nonlinearly with respect to density. The rapid changes of the sensitivity of saturated liquid pressure with density renders linear error propagation to be invalid to calculate the uncertainties in Eqs. (50) and (51). To calculate Δres_2 for the linear error propagation in Eqs. (50) and (51), the uncertainty of the Gibbs energy difference in Eq. (52) is used to approximate the uncertainty of difference of pressure values as shown in Eq. (53).

$$\Delta res_2 = \frac{p \Delta res_1}{RT} \quad (53)$$

Equation (53) is used because the relative difference of Gibbs energy between saturated liquid and vapor changes proportionally with the difference of estimated and measured saturated pressure as shown in Fig. 1.

Since residual functions calculate the differences of Gibbs energy and pressure respectively, the uncertainties of the corresponding Gibbs energy and pressure at saturation are given by the uncertainties of the residual functions as shown in Eqs. (54) and (55).

$$\Delta res_1 = \sqrt{\mathrm{diag}(COV(g(\rho_\mathrm{v}, T) - g(\rho_\mathrm{l}, T)))} \cdot t(n-m, \gamma_\mathrm{t}/2) \tag{52}$$

$$\Delta g_\mathrm{sat} = \Delta res_1 \tag{54}$$

$$\Delta p_\mathrm{sat} = \Delta res_2 \tag{55}$$

where sat means saturation.

With the approximation in Eq. (53), the uncertainties of p and g at saturation in Eqs. (54) and (55) and the definition of fugacity, the uncertainty of fugacity coefficient ϕ at saturation can be derived to be zero and the relative uncertainty of fugacity equals to the relative uncertainty of pressure at saturation in Eq. (55).

However, other thermodynamic properties at the saturated states are calculated by solving the original EOS with the density values from the Maxwell criteria and the original temperature input to the Maxwell criteria. The calculation would propagate the uncertainties calculated from Eqs. (50) and (51), and the uncertainties from Eqs. (50) and (51) will be added to the uncertainties from Eq. (42) to calculate the uncertainties of these thermodynamic properties. For example, the uncertainty of entropy of saturated liquid and saturated vapor are calculated by adding the uncertainties as Eqs. (56) and (57).

$$\Delta s_\mathrm{l} = \sqrt{\begin{array}{l}\mathrm{diag}(COV(s(T, \rho_\mathrm{l}))) \cdot t(n-m, \gamma_\mathrm{t}/2)^2 \\ + \left(\frac{\frac{\partial s_l}{\partial T}\big|_\mathrm{sat}}{\frac{\partial \rho_l}{\partial T}\big|_\mathrm{sat}} \Delta\rho_\mathrm{l} \right)^2 \end{array}} \tag{56}$$

$$\Delta s_\mathrm{v} = \sqrt{\begin{array}{l}\mathrm{diag}(COV(s(T, \rho_\mathrm{v}))) \cdot t(n-m, \gamma_\mathrm{t}/2)^2 \\ + \left(\frac{\frac{\partial s_\mathrm{v}}{\partial T}\big|_\mathrm{sat}}{\frac{\partial \rho_\mathrm{v}}{\partial T}\big|_\mathrm{sat}} \Delta\rho_\mathrm{v} \right)^2 \end{array}} \tag{57}$$

where l means saturated liquid and v is saturated vapor.

When the uncertainty of the differences between the entropy of saturated liquid or vapor and an entropy value in a homogeneous phase is needed, the calculation can be carried out by Eqs. (58) and (59).

The uncertainty of the entropy of vaporization can be calculated by Eq. (60).

Uncertainties of other properties (e.g., u, h, a, c_p, c_v and ϕ) of saturated liquid and vapor can be calculated in a similar manner as Eqs. (56) and (57). Numerically, the uncertainties can also be calculated from Eq. (42) with density and temperature of the saturated liquid and vapor, but Eq. (42) does not include the uncertainty in the EOS caused by the Maxwell criteria and underestimates the uncertainties. Hence when the uncertainties of properties at saturation are needed, Eqs. (56) and (57) should be used to calculate the uncertainties of entropy of the saturated liquid and vapor instead.

Properties other than temperature can be used as inputs to the EOS to calculate the thermodynamic properties of saturated states; for instance pressure can be used as an input to find the saturation temperature of a fluid at that pressure. To calculate the uncertainty of saturation temperature, the uncertainty of pressure in Eq. (55) can be used to propagate the uncertainty of the EOS to the saturation temperature value, and the uncertainty of saturation temperature can be calculated by Eq. (61).

$$\Delta T_\mathrm{sat} = \left| \left(\frac{\mathrm{d}T}{\mathrm{d}p}\bigg|_\mathrm{sat} \right) \Delta p_\mathrm{sat} \right| \tag{61}$$

where $\frac{\mathrm{d}T}{\mathrm{d}p}|_\mathrm{sat}$ is calculated by the Clapeyron relation as shown in Eq. (62).[30]

$$\frac{\mathrm{d}T}{\mathrm{d}p}\bigg|_\mathrm{sat} = \frac{1/\rho_\mathrm{l} - 1/\rho_\mathrm{v}}{s_\mathrm{v} - s_\mathrm{l}} \tag{62}$$

and the uncertainty in vapor pressure Δp_sat is obtained from Eq. (53).

$$\Delta(s_{\mathrm{l}}-s_{1\mathrm{p}})=\sqrt{\mathrm{diag}(COV(s(T,\rho_{\mathrm{l}})-s_{1\mathrm{p}}(T_{1\mathrm{p}},\rho_{1\mathrm{p}})))\cdot t(n-m,\gamma_{\mathrm{t}}/2)^2+\left(\frac{\frac{\partial s_{\mathrm{l}}}{\partial T}\Big|_{\mathrm{sat}}}{\frac{\partial \rho_{\mathrm{l}}}{\partial T}\Big|_{\mathrm{sat}}}\Delta\rho_{\mathrm{l}}\right)^2} \tag{58}$$

$$\Delta(s_{\mathrm{v}}-s_{1\mathrm{p}})=\sqrt{\mathrm{diag}(COV(s(T,\rho_{\mathrm{v}})-s_{1\mathrm{p}}(T_{1\mathrm{p}},\rho_{1\mathrm{p}})))\cdot t(n-m,\gamma_{\mathrm{t}}/2)^2+\left(\frac{\frac{\partial s_{\mathrm{v}}}{\partial T}\Big|_{\mathrm{sat}}}{\frac{\partial \rho_{\mathrm{v}}}{\partial T}\Big|_{\mathrm{sat}}}\Delta\rho_{\mathrm{v}}\right)^2} \tag{59}$$

$$\Delta(s_{\mathrm{v}}-s_{\mathrm{l}})=\sqrt{\begin{array}{c}\mathrm{diag}(COV(s(T,\rho_{\mathrm{v}})-s(T,\rho_{\mathrm{l}})))\cdot t(n-m,\gamma_{\mathrm{t}}/2)^2\\+\left(\frac{\frac{\partial s_{\mathrm{l}}}{\partial T}\Big|_{\mathrm{sat}}}{\frac{\partial \rho_{\mathrm{l}}}{\partial T}\Big|_{\mathrm{sat}}}\Delta\rho_{\mathrm{l}}-\frac{\frac{\partial s_{\mathrm{v}}}{\partial T}\Big|_{\mathrm{sat}}}{\frac{\partial \rho_{\mathrm{v}}}{\partial T}\Big|_{\mathrm{sat}}}\Delta\rho_{\mathrm{v}}\right)^2\end{array}} \tag{60}$$

3.4 Thermodynamic quality

When defining the thermodynamic state of liquid and vapor phases in equilibrium of a pure substance, an additional property called thermodynamic quality as defined by Eq. (63).

$$q=\frac{1/\rho-1/\rho_{\mathrm{l}}(T)}{1/\rho_{\mathrm{v}}(T)-1/\rho_{\mathrm{l}}(T)} \tag{63}$$

where q is thermodynamic quality, ρ_{l} is the density of saturated liquid, ρ_{v} is the density of saturated vapor, q is the thermodynamic quality, and $\rho_{\mathrm{v}}\leq\rho\leq\rho_{\mathrm{l}}$.

Although q is not given as part of an output of a part of Helmholtz-based EOS, it is commonly computed as an internal function in software packages calculating thermodynamic properties, and hence it is necessary to define its uncertainty calculation. From Eq. (63), it can be seen that the uncertainty of thermodynamic quality due to EOS mainly comes from the calculation of the density of the saturated liquid and vapor. By propagating the uncertainty of the density of saturated liquid and vapor, the uncertainty of thermodynamic quality due to the EOS can be calculated by Eq. (64).

$$\Delta q=\left|\frac{\partial q}{\partial\rho_{\mathrm{v}}}\Delta\rho_{\mathrm{v}}+\frac{\partial q}{\partial\rho_{\mathrm{l}}}\Delta\rho_{\mathrm{l}}\right| \tag{64}$$

$$\frac{\partial q}{\partial\rho_{\mathrm{l}}}=\frac{1-q}{\rho_{\mathrm{l}}^2\left(\frac{1}{\rho_{\mathrm{v}}}-\frac{1}{\rho_{\mathrm{l}}}\right)} \tag{65}$$

$$\frac{\partial q}{\partial\rho_{\mathrm{v}}}=\frac{q}{\rho_{\mathrm{v}}^2\left(\frac{1}{\rho_{\mathrm{v}}}-\frac{1}{\rho_{\mathrm{l}}}\right)} \tag{66}$$

where Δq is the uncertainty of the thermodynamic quality, and $\Delta\rho_{\mathrm{v}}$ and $\Delta\rho_{\mathrm{l}}$ are uncertainties calculated in Section 3.3.

If either temperature or density are not given as inputs to the EOS, they will be first calculated by solving the EOS iteratively. The temperature and density values will be used to calculate uncertainty of the thermodynamic quality according to Eq. (64).

3.5 Other thermodynamic properties of two-phase mixtures

The calculation method of the uncertainty of the property from the EOS for two-phase systems with input variables other than temperature and density is different from that in Sections 3.1 and 3.2, because it involves the calculation of the thermodynamic quality. In this case, the uncertainties calculated in Section 3.3 related to the Maxwell criteria are also involved in the calculation of the uncertainties of proper-

ties using thermodynamic quality as one of the inputs.

When the thermodynamic quality is 0 (saturated liquid) or 1 (saturated vapor), the uncertainty of the properties can be calculated based on the equations in Section 3.3, and their uncertainties are given by Eqs. (54), (55) and (61). However, for all other properties, their uncertainties are calculated with a different method. For example, if the uncertainty of an entropy value is calculated with an intermediate thermodynamic property as one of its inputs, to include the uncertainty propagated from the use of the Maxwell criteria, its calculation will be carried out with Eq. (67).

$$\begin{aligned}\Delta s_{2\mathrm{p}} &= \sqrt{\begin{array}{l}\left(\frac{\partial s_{2\mathrm{p}}}{\partial q}|_T \Delta q\right)^2 \\ +\mathrm{diag}(COV(s_{2\mathrm{p}})) \cdot t(n-m,\gamma_\mathrm{t}/2)^2\end{array}} \\ &= \sqrt{\begin{array}{l}[(s_\mathrm{v}-s_\mathrm{l})\Delta q]^2 \\ +\mathrm{diag}(COV(s_{2\mathrm{p}})) \cdot t(n-m,\gamma_\mathrm{t}/2)^2\end{array}}\end{aligned} \tag{67}$$

where the subscript 2p means two-phase fluid.

For the uncertainty of its difference with entropy values in single phase, the uncertainty can be calculated by Eq. (68).

$$\begin{aligned}&\Delta(s_{2\mathrm{p}}(T_{2\mathrm{p}},q_{2\mathrm{p}})-s_{1\mathrm{p}}) \\ &= \sqrt{\begin{array}{l}[(s_\mathrm{v}(T_{2\mathrm{p}})-s_\mathrm{l}(T_{2\mathrm{p}}))\Delta q_{2\mathrm{p}}]^2 \\ +\mathrm{diag}(COV(s_{2\mathrm{p}}-s_{1\mathrm{p}})) \\ \cdot t(n-m,\gamma_\mathrm{t}/2)^2\end{array}}\end{aligned} \tag{68}$$

where 1p is single phase.

When the uncertainty of its difference with entropy values of the saturated liquid or vapor is needed, the uncertainty can be calculated by Eqs. (69) and (70).

The uncertainty of the difference between a pair of two-phase entropy states can be calculated by Eq. (71). The uncertainty of properties u, h and a can also be calculated in a manner similar to Eqs. (67) to (70).

The uncertainty of density can be calculated from

$$\Delta\rho_{2\mathrm{p}} = \sqrt{\left(\frac{-1/\rho^2}{1/\rho_\mathrm{v}-1/\rho_\mathrm{l}}\Delta q\right)^2 + (\Delta\rho)^2}, \tag{72}$$

where $\Delta\rho$ comes from Eq. (47).

The properties c_p and w are undefined for two-phase states and therefore do not have any uncertainty values associated with them.

3.6 Summary

Figure 2 summarizes the choice of uncertainty calculation method of the EOS based on the input variables and the phase of the fluid where the uncertainty of a property value is needed.

When the input variables are temperature and density, the calculation method follows the basic uncertainty calculation method derived in Section 3.1 because the calculation does not involve the determination of the phase of the fluid nor any iterative calculation. If the input variables are not temperature and density, the EOS would involve iterative calculation. For superheated vapor and subcooled liquid, the method should follow Section 3.2. For saturated vapor and liquid, the calculation involves the Maxwell criteria and Section 3.3 should be followed. For two-phase states, if the output is Gibbs energy, pressure, fugacity coefficient or temperature, the calculation method would be the same as that of saturated liquid and vapor, and the method in Section 3.3 should be followed. If the output is thermodynamic quality, the uncertainty calculation method in Section 3.4 should be used. Otherwise, the uncertainty calculation method in Section 3.5 should be followed.

When the differences of properties are needed, the calculation steps of the uncertainties follow the flowchart in Fig. 3

If the difference involves T, ρ, or q, whose covariance cannot be calculated directly, the uncertainty of the differences will be calculated by propagating the uncertainties of individual property values with methods in Kline and McClintock.[24] If the inputs are all T and ρ, no Maxwell criteria will be involved even for two-

$$\Delta(s_{2\mathrm{p}}(T_{2\mathrm{p}},q_{2\mathrm{p}})-s_{\mathrm{l}}(T_{\mathrm{l}}))=\sqrt{\begin{array}{l}\mathrm{diag}(COV(s_{2\mathrm{p}}(T_{2\mathrm{p}},q_{2\mathrm{p}})-s_{\mathrm{l}}(T_{\mathrm{l}})))\cdot t(n-m,\gamma_{\mathrm{t}}/2)^2\\ +\left(\dfrac{\left.\dfrac{\partial s_{\mathrm{l}}(T_{\mathrm{l}})}{\partial T}\right|_{\mathrm{sat}}}{\left.\dfrac{\partial \rho_{\mathrm{l}}(T_{\mathrm{l}})}{\partial T}\right|_{\mathrm{sat}}}\Delta\rho_{\mathrm{l}}(T_{\mathrm{l}})-[(s_{\mathrm{v}}(T_{2\mathrm{p}})-s_{\mathrm{l}}(T_{2\mathrm{p}}))\Delta q_{2\mathrm{p}}]\right)^2\end{array}} \tag{69}$$

$$\Delta(s_{2\mathrm{p}}(T_{2\mathrm{p}},q_{2\mathrm{p}})-s_{\mathrm{v}}(T_{\mathrm{v}}))=\sqrt{\begin{array}{l}\mathrm{diag}(COV(s_{2\mathrm{p}}(T_{2\mathrm{p}},q_{2\mathrm{p}})-s_{\mathrm{v}}(T_{\mathrm{v}})))\cdot t(n-m,\gamma_{\mathrm{t}}/2)^2\\ +\left(\dfrac{\left.\dfrac{\partial s_{\mathrm{v}}(T_{\mathrm{v}})}{\partial T}\right|_{\mathrm{sat}}}{\left.\dfrac{\partial \rho_{\mathrm{v}}(T_{\mathrm{v}})}{\partial T}\right|_{\mathrm{sat}}}\Delta\rho_{\mathrm{v}}(T_{\mathrm{v}})-[(s_{\mathrm{v}}(T_{2\mathrm{p}})-s_{\mathrm{l}}(T_{2\mathrm{p}}))\Delta q_{2\mathrm{p}}]\right)^2\end{array}} \tag{70}$$

phase substances, and the uncertainty of property differences can be calculated directly following Section 3.1. If the difference involves two-phase states, the property calculation follows that listed in Section 3.5. If the difference involves saturated liquid or vapor, the method to calculate the uncertainty of property differences in Section 3.3. Otherwise, the calculation method of uncertainty methods listed in Section 3.1 will be used.

The Python code used to calculate the uncertainties is listed in the Supplementary Materials for reference.

4 Results and Discussion

To illustrate the application of the proposed uncertainty calculation method, it is applied to the EOS of propane.[12] Its EOS is Helmholtz-energy-explicit, and the proposed uncertainty calculation method can be applied to it to demonstrate how uncertainties of the EOS affect the properties estimated by the EOS.

4.1 Experimental data

The experimental data used in this work were obtained via NIST ThermoDataEngine (TDE) #103b.[36–39] The data sources are summarized in Table 3, and a more detailed list of the experimental data considered is provided in the Supporting Data. These data span several different types, including densities, saturation properties (vapor pressure, latent heat of vaporization, etc.) and properties in homogeneous phases (speed of sound, heat capacities, etc.).

The resultant normalized σ values in Eq. (26) that are calculated based on the deviations between the data in Table 3 and the estimation from the EOS are listed in Table 4. These σ values come from the relative deviations between the EOS and the measurement results. The relative deviation between measurement data and the EOS estimation of pressure from the data sets of liquid and vapor density yield the largest values in Table 4. It can be seen that the major source of uncertainties come from the deviations of the measurement data of density of liquid and vapor with the EOS estimation. The relative deviations between measurement and

$$\Delta(s_{2\mathrm{p},1}(T_{2\mathrm{p},1},q_{2\mathrm{p},1})-s_{2\mathrm{p},2}(T_{2\mathrm{p},2},q_{2\mathrm{p},2}))=\sqrt{\begin{array}{l}\mathrm{diag}(COV(s_{2\mathrm{p},1}-s_{2\mathrm{p},2}))\cdot t(n-m,\gamma_{\mathrm{t}}/2)^2\\ +([(s_{\mathrm{v}}(T_{2\mathrm{p},1})-s_{\mathrm{l}}(T_{2\mathrm{p},1}))\Delta q_{2\mathrm{p},1}]-[(s_{\mathrm{v}}(T_{2\mathrm{p},2})-s_{\mathrm{l}}(T_{2\mathrm{p},2}))\Delta q_{2\mathrm{p},2}])^2\end{array}} \tag{71}$$

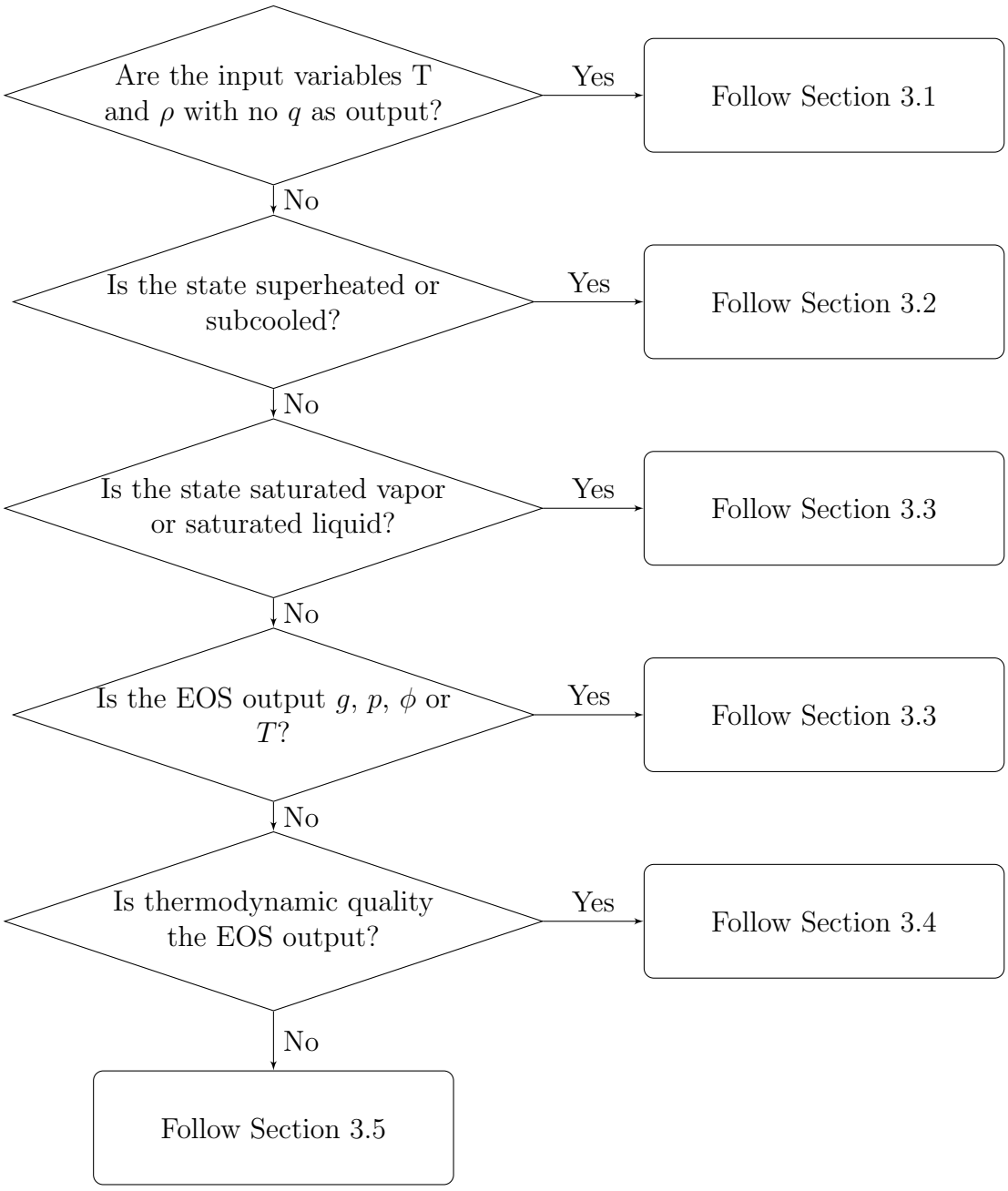


Figure 2: Flowchart showing the choice of equations to calculate the uncertainties of EOS based on the input variables and state of the fluid

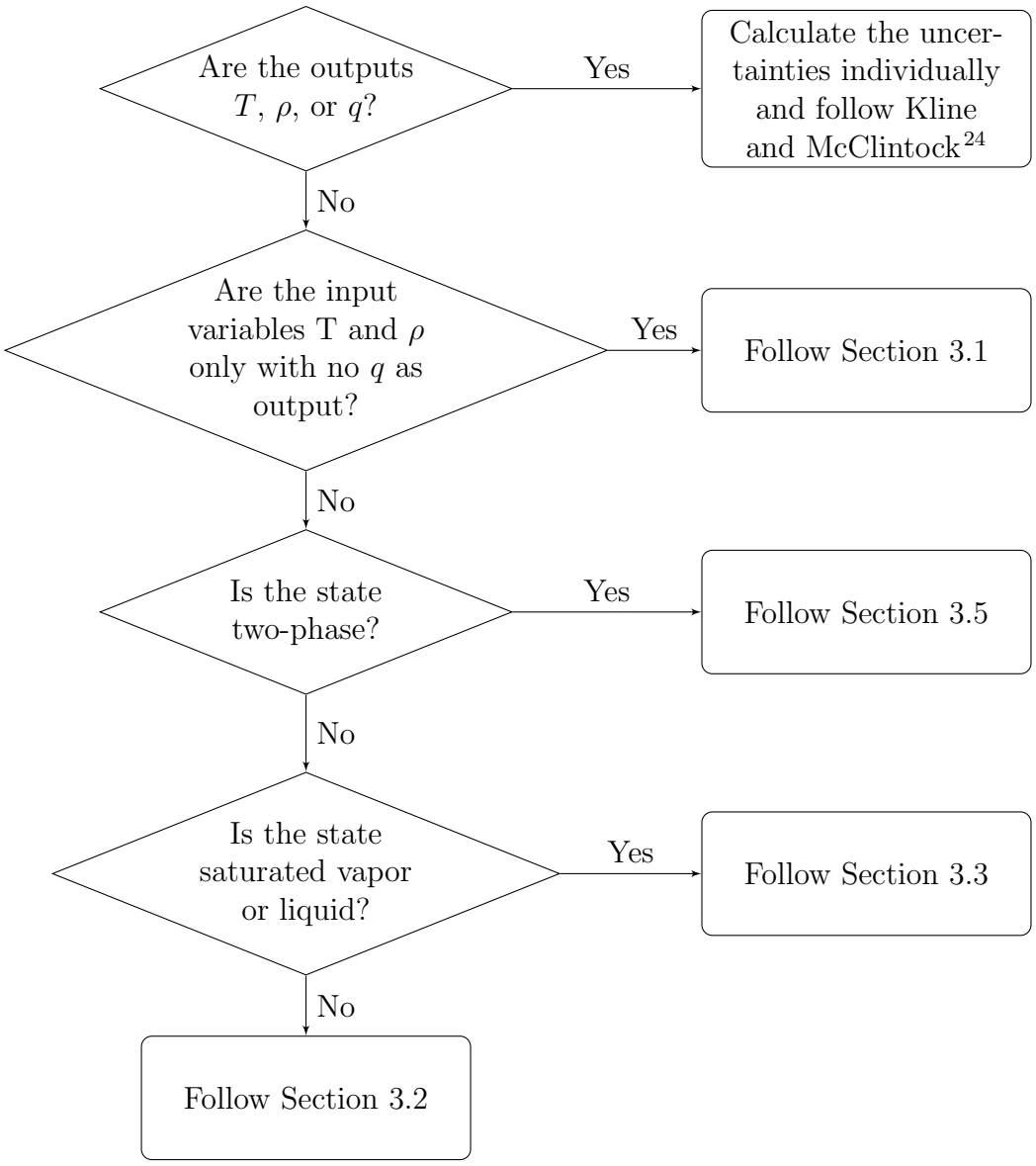


Figure 3: Flowchart showing the choice of equations to calculate the uncertainties of property differences

the EOS estimation of specific heat capacity of liquid and vapor, saturation pressure, second virial coefficient and speed of sound also contribute to the uncertainties of the EOS significantly. However, the relative deviations between measurement and the EOS estimation of ideal gas specific heat capacity, critical temperature and critical density are much smaller than that of the other properties, and they do not contribute to the uncertainty of the EOS as significantly as the measurement of the other properties.

Table 4 also shows the corresponding standard measurement uncertainty of different property values that are converted. These uncertainty values were taken from ThermoDataEngine and represent a combination of experimental uncertainties ascribed by the individual researcher as well as, when appropriate, expert evaluation to increase the uncertainty to a more reasonable value if the claimed uncertainty is not reasonable. These standard uncertainty values are calculated by averaging the standard measurement uncertainties of the measurement data, and they can be compared directly with the normalized σ values. The results show that the noramlized σ are not significantly larger than their corresponding

Table 3: Summary of experimental data from TDE considered in this study. Any experimental data points flagged by TDE as being unreliable were not included.

Data type	N	T range K	p range MPa
T_c	45		
p_c	38		
ρ_c	18		
p_{sat}	1392	85 to 369	1.6332×10^{-10} to 4.91426
Liquid ρ	3746	88 to 369	0.3546 to 1072.7
Saturated Liquid ρ	528	86 to 369	
Vapor ρ	3494	243 to 609	0.040934 to 1047
Saturated Vapor ρ	100	230 to 369	
Second Virial Coefficient B	167	183 to 559	
Enthalpy of Vaporization	50	186 to 362	
Liquid c_p	108	89 to 366	0.101325 to 5.43
Vapor c_p	138	273 to 573	0.049033 to 10.3421
c_{p0}	50	148 to 353	
Homogeneous Phase c_p	110	100 to 374	
Liquid w	655	90 to 339	1.92 to 60.58
Saturated Vapor w	59	90 to 325	
Vapor w	423	225 to 375	0.01008 to 0.8513
Homogenous Phase w	593	239 to 498	9.80665×10^{-6} to 101.337

measurement uncertainties with the exception of the second virial coefficient and the speed of sound. For these two variables, the normalized σ are at least 100% larger than their measurement counterparts. The normalized σ for p from data sets of liquid and vapor ρ is also much larger than the relative standard measurement uncertainty of p values at saturation. The results show that the inaccuracy of EOS estimation may also play a critical role in the uncertainty of EOS besides the measurement uncertainty of the experimental data.

We present the results in three parts. First the parameter identifiability is analyzed. Then the uncertainties of saturation properties are studied. Finally the variation of uncertainties with the thermodynamic state points are presented and discussed.

4.2 Covariance, correlation matrix and identifiability

Considering the parameter covariance matrix (as calculated by Eq. 9) it is also possible to assess the parameter identifiability. The square-root of the diagonal elements of the covariance matrix (i.e. $\sqrt{\text{diag}(COV(\theta^*_{\text{EOS}}))}$) correspond to the parameter standard deviation. If the standard deviations with respect to the parameter values are large, parameters are not practically identifiable. This means that there are not enough data to estimate the parameters with the current model and objective function. We chose as a metric for practical identifiability: $\left|\sqrt{\text{diag}(COV(\theta^*_{\text{EOS}}))}/\theta^*_{\text{EOS}}\right|$. In the given study nearly all the parameters (besides n_{16} and n_{17}) have a small standard deviation compared to the parameter value, i.e. $\left|\sqrt{\text{diag}(COV(\theta^*_{\text{EOS}}))}/\theta^*_{\text{EOS}}\right| < 0.1$ (see the Supporting Information). This result implies that the amount of experimental data used for the parameter fitting (see Lemmon *et al.*[12]) and the uncertainty analysis is sufficiently high to guarantee the identification of the parameters from the data. A small amount of experimental data would lead to poor identifiability and therefore parameter values with large standard

Table 4: Normalized σ values from deviations between measurement and EOS estimation in Eq. (26) for propane, and their corresponding average relative standard measurement uncertainty

Property	Normalized σ	Corresponding average relative standard uncertainty
c_{p0}	5.104×10^{-3}	1.448 %
c_p	8.113×10^{-2}	4.391 %
c_v	7.021×10^{-2}	1.081 %
Saturated liquid and vapor p	2.661×10^{-2}	5.228 %
p from data sets of liquid and vapor ρ	1.465×10^{-1}	Not available
ρ_c	4.447×10^{-6}	1.306 %
T_c	7.767×10^{-6}	0.100 %
B	3.880×10^{-2}	1.623 %
w	6.388×10^{-2}	0.196 %

deviations. It is recommended that developers of Helmholtz-type EoS analyze the parameter covariance matrix with respect to parameter identifiability in order to ensure that sufficient experimental data has been used for the fitting of the respective parameters.

The parameter correlation matrix (obtained by Eq. 10), contains information of the correlation coefficients between the parameters; the parameter correlation matrix is attached in the supplementary material. Several of the parameters are highly correlated, corresponding to high correlation coefficients $> \pm 0.7$ (a correlation coefficient equal to 1 would correspond to a perfect correlation). This means that many parameters are not independent from each other, due to the structure of the equations: Parameters increase or decrease, when other ones increase or decrease. However, due to the fact that a sufficiently large amount of experimental data has been used for the calculation, the high correlation coefficients did not lead to high parameter uncertainties.

4.3 Computational cost to calculate an uncertainty

To examine if the computational cost of the proposed method is lower than the alternative methods,[1,22] the computational cost of the methods are analyzed by comparing the most computational expensive portion of the methods. In the alternative methods,[1,22] the most computationally expensive step lies in the Monte Carlo method which requires at least 1000 computations of the EOS before the convergence for each uncertainty value. In contrast, the most computational expensive portion of the method is the calculation of the Jacobian vector in Eq. (11). While the Jacobian matrix in Eq. (32) and the Covariance matrix in Eq. (35) are more computationally expensive than the calculation of a Jacobian vector, their values, upon calculation from training data, are stored in memory for all other uncertainty calculation of properties of the same substance. Their calculation does not repeat, and the computational cost is negligible compared to the computational cost of the uncertainty method in the long run. In this study, each entry of the Jacobian vector is calculated by performing differentiation on the EOS, and the computational cost of each entry of the vector is at most three times of the computational cost of a computation of the EOS. The number of entries in each Jacobian vector equals to the number of parameters in an EOS plus two as shown in Eq. (38) (i.e. 99 for propane in this study), and the computational cost of a Jacobian vector is at most 300 computations of the EOS. Thus the computational cost of the proposed method is around 30% of the computational cost of the alternative

methods, and the computation of the proposed method is faster than the alternatives.

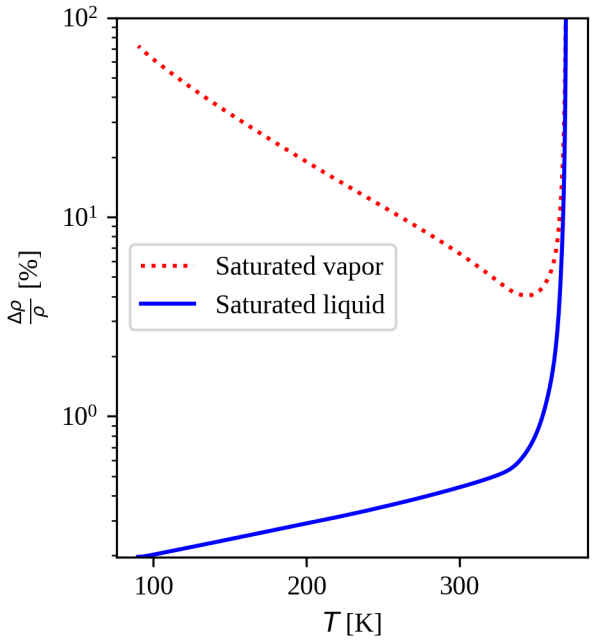
4.4 Saturated property uncertainty

To understand the magnitude of uncertainties of various properties within the applicable range of the EOS, the uncertainties of density and enthalpy along the saturation curves are plotted from the triple point temperature of 85.525 K to the critical temperature[12] with uncertainty bounds overlaid as shown in Fig. 4.

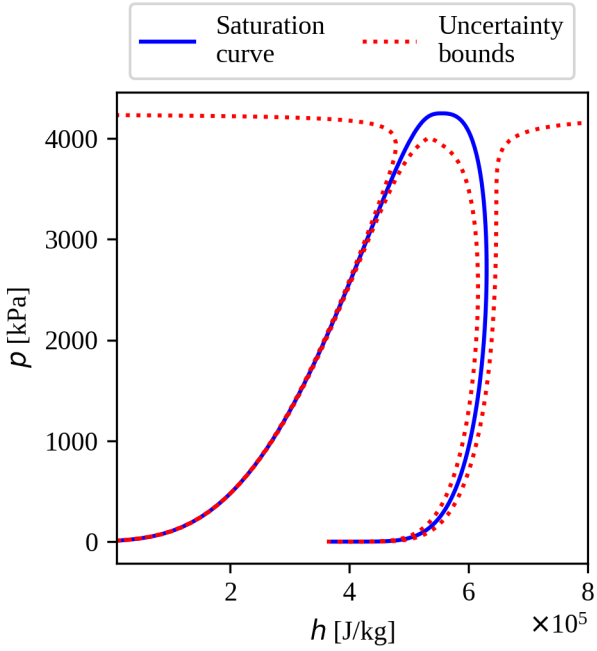
The uncertainties in Figure 4 are reasonable at most conditions except near the critical point in both diagrams and at low temperature for saturated vapor in Fig. 4a. The uncertainties of properties near the critical point are much larger than ones further from the critical point, which is also reported in Lemmon *et al.*[12] without quantification. The reason for the large uncertainties is due to the mathematical structure of the Maxwell criteria as shown in Section 3.3. According to the Maxwell criteria and Eqs. (50) and (51), the uncertainties of the density of saturated liquid and vapor and hence the uncertainties of other saturated properties depend on the derivative of densities with respect to pressure and Gibbs energy. As the temperature of a saturated substance approaches the critical point, the saturation densities change rapidly along the saturation line, and the derivatives of densities with respect to pressure become very large in magnitude, as shown in Fig. 5.

The trend of the derivative increasing to infinity at the critical temperature as shown in Fig. 5 are unavoidable because critical points are classically defined as the point where $\left.\frac{\partial \rho}{\partial p}\right|_T$ and $\left.\frac{\partial^2 \rho}{\partial p^2}\right|_T$ are infinite.[30] The uncertainties of saturated liquid and vapor properties depend on the uncertainties of their densities; hence their large derivatives with respect to pressure near the critical point results in large uncertainties as shown in Fig. 4.

Fig. 5 also explains why the uncertainty of saturated vapor is large at low temperature in Fig. 4a. The relative derivatives of density of saturated vapor with respect to pressure is increasing with decreasing temperature. These increasing derivatives propagate into the uncertainty of saturated density by Maxwell criteria, and the uncertainty of density of saturated



(a) Change of relative uncertainty of ρ of saturated liquid and vapor with temperature along the saturation line



(b) p-h diagram with uncertainties of h relative to its reference state

Figure 4: Uncertainties along saturation lines in different property diagrams from 85.525 K to the critical temperature

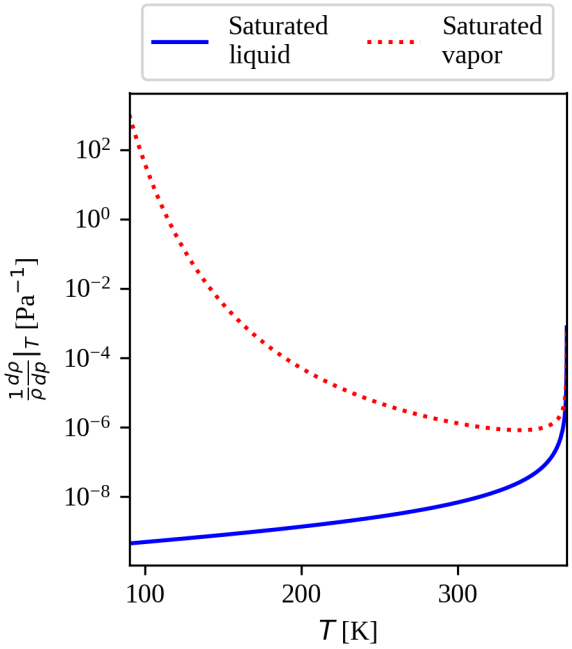
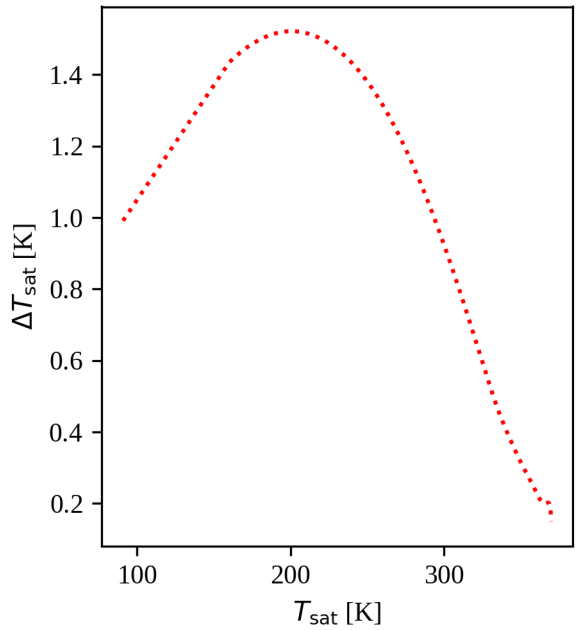
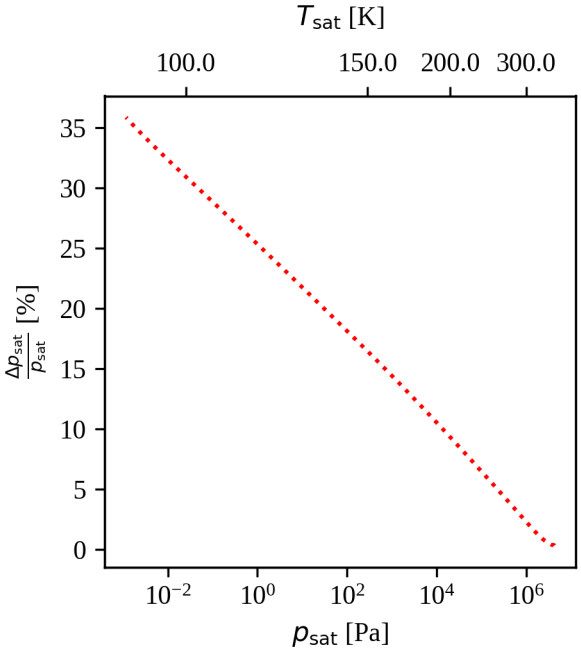


Figure 5: Changes of relative derivatives of density with respect to pressure with temperature from 85.525 K to the critical temperature of propane



(a) Changes of uncertainty of saturation temperature with temperature



(b) Changes of relative uncertainty of saturation pressure with pressure

Figure 6: Changes of saturation property uncertainties with their corresponding properties

vapor becomes large. In contrast, the relative derivative of density of saturated liquid is decreasing with decreasing temperature, and by the same uncertainty propagation mechanism, the uncertainty of density of saturated liquid becomes small at low temperature in Fig. 4a.

To examine the uncertainty of saturation pressure along the saturation line, the uncertainty of saturation temperature and the relative uncertainty of pressure calculated by Eq. (73) are plotted with the saturation properties in Fig. 6.

$$\text{Relative uncertainty of variable } x = \frac{\Delta x}{x} \quad (73)$$

.

Figure 6a shows that the uncertainty of saturation temperature becomes more significant at lower temperature. At temperature around 200 K, the uncertainty peaks at around 1.5 K, and it approaches 1 K as the temperature is lowered to the triple point of propane. However, when the uncertainty is studied from the perspective of pressure in Figure 6b, the relative uncertainty of saturation pressure is found to be increasing as the pressure drops. The cause of the large uncertainty is due to the scattering of the deviations between the estimated and measured pressure at low pressure levels as shown

in Fig. 7.

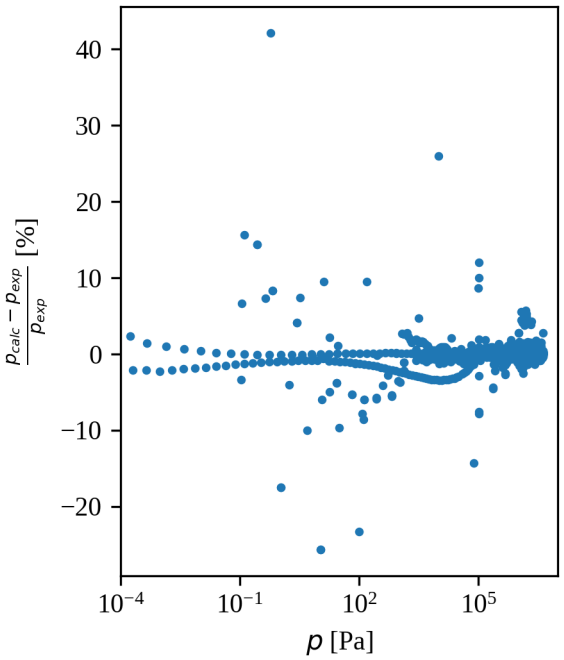


Figure 7: Relative deviations of pressure estimates

As the pressure drops, the relative deviations between the estimated and measured pressure scatter. At high pressure, the relative deviations concentrate between +10% and -10%. However, as lower pressure, the distribution of the relative deviations widens to between +40 % and -30 %. The large scattering of the relative deviation of saturation pressure at low levels results in the increase of relative uncertainties of saturation pressure with a drop of pressure.

To have a more comprehensive understanding, the relative uncertainty of pressure in the single phase region is also studied by the contour plot of its relative uncertainty over a temperature-specific-volume (T-1/ρ) plot. The contour plot is shown in Fig. 8.

Figure 8 shows that the relative uncertainties of pressure of vapor in the single-phase region are much less than the relative uncertainties of saturation pressure. There are some relative uncertainties of pressure of subcooled liquid at high densities that are high, but it fits the general understanding that the derivative of pressure with density is high in that regime and hence the relative uncertainty of pressure in the regime is high.

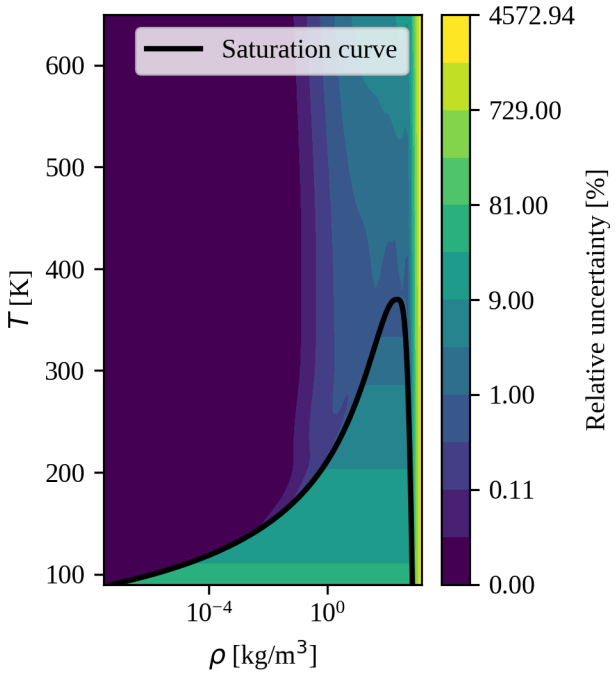


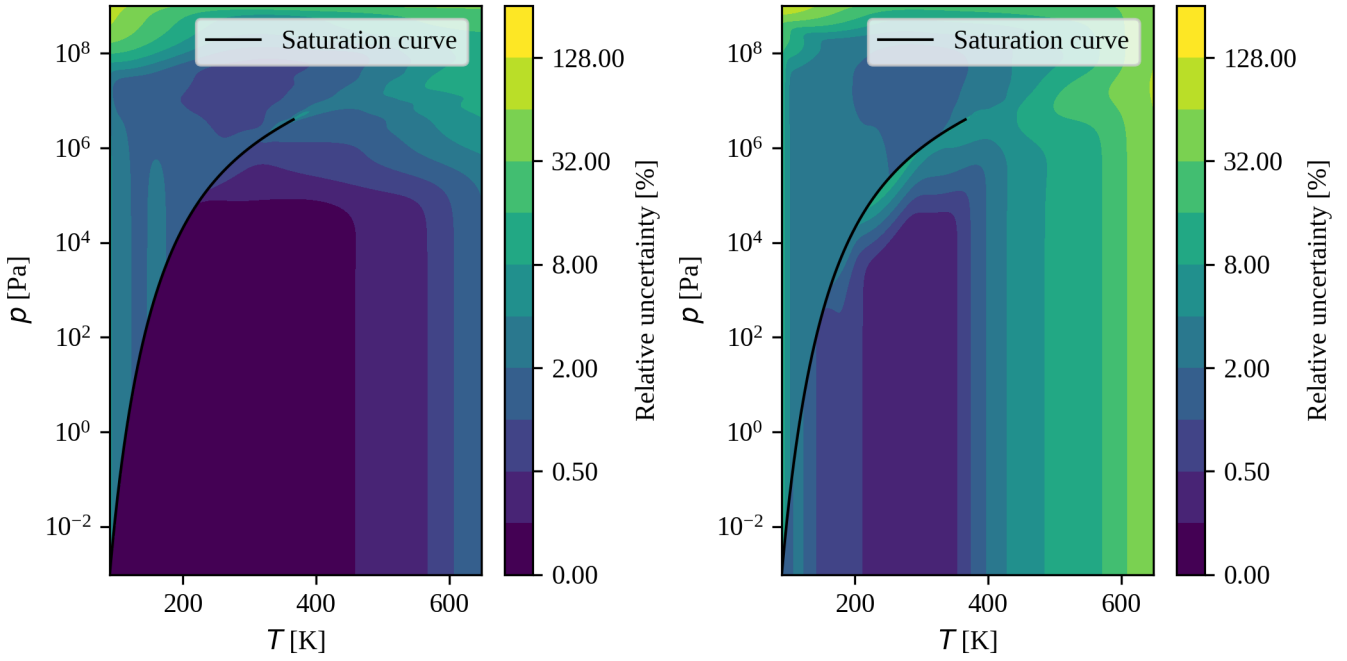
Figure 8: Relative uncertainty of pressure in a T-1/ρ diagram

4.5 Uncertainty of properties in the single-phase region

To further understand how the uncertainties of EOS change in the single-phase region, the relative uncertainties of speed of sound, isochoric specific heat capacity, isobaric specific heat capacity and density are calculated at different temperature and pressure state points, and the results are shown in Figs. 9a to 9d.

Figure 9a shows that the relative uncertainty of speed of sound is lower than 4 % in most cases except conditions near the critical point and at pressure higher than 10^8 Pa. At pressure higher than 10^8 Pa, only speed of sound data were only collected near 240 K and density data were collected between 370 K and 610 K. Other property data were not collected in the pressure regime. Thus the figure only shows an uncertainty value lower than 8% between 240 K and 610 K at pressure higher than 10^8 Pa, and the uncertainty of the speed of sound remains high at all other temperature values in the pressure regime.

Figures 9b and 9c show a similar pattern in



(a) Relative uncertainty of speed of sound

(b) Relative uncertainty of c_v

(c) Relative uncertainty of c_p

(d) Relative uncertainty of density

Figure 9: Relative uncertainties in p-T coordinates

which the relative uncertainty remains lower than 4 % in the subcooled liquid region and highly superheated vapor region. However, the relative uncertainties of the specific heat are much more significant near the the saturated vapor line at 250 K. While the c_p and c_v of superheated vapor near the saturated curve at low temperatures should behave as an ideal gas, the relative uncertainties of c_p and c_v near the saturated curve at around 250 K are much higher than their ideal gas counterparts at lower pressure. While this may be caused by the approximation to use model deviation instead of measurement uncertainty to conduct the calcula-

tion, it may also caused by potential overffiting issues in the part of the EOS related to the residual Helmholtz energy.

Overfitting arises when a regression equation contains too many coefficients and is fit for the random variation in the experimental data rather than the systematic relationship between variables. Equations with overfitting issues usually result in very accurate prediction at the experimental data points but large model uncertainties.[40] Since c_p and c_v in the vapor phase with the exception of ones near the critical point should not differ too much from their ideal gas counterparts and the EOS part of the ideal gas contribution (α^0 in Eq. (14)) is not complex enough to result in significant overfitting, the large uncertainty is likely caused by the overfitting of the residual part of the Helmholtz energy arising from the intermolecular forces (α^r in Eq. (14)). To mitigate these issues and the uncertainties, simplification of EOS or constraints to the coefficients such as penalization should be made to reduce these unexpectedly high uncertainties.[41]

Figure 9d shows a different pattern. Relative uncertainty of density is lower than 0.25 % for all subcooled liquid states and is lower than 1 % for most highly superheated vapor states. The relative uncertainty of density of the superheated vapor at higher pressures also ranges between 1 % and 4 %. The relative uncertainty only exceeds 8 % in the critical region.

In summary, the relative uncertainties of EOS are in general large in the critical region. For pressure and speed of sound, their relative uncertainties in the subcooled liquid region are higher than that of the superheated vapor region. For specific heat capacities and densities, the relative uncertainties of superheated vapor close to saturation are higher than those of the subcooled liquid and more highly superheated vapor.

5 Conclusion

In this paper, a method to calculate the uncertainty of thermodynamic properties of Helmholtz-energy-explicit EOS is developed based on the parameter covariance matrix of nonlinear regression models. The covariance matrix can be calculated from the experimental data in the literature and the Jacobian matrix of the EOS with respect to the parameters of the EOS. To ensure that the effects of the Maxwell criteria and uncertainties of the differences of properties such as enthalpy values are properly calculated, the uncertainty calculation method also involves the linearization of the EOS and the covariance of the differences of properties.

To test the applicability and validity of the calculation method, it is demonstrated by a calculation of the uncertainties of the EOS of propane. The method also enables an analysis of the parameter correlation matrix which shows how the parameters are correlated with each other. It also demonstrates how the use of the Maxwell criteria and the rapid change of properties with respect to the state of the substances around their critical point result in larger uncertainties of properties along the superheated vapor line than the uncertainties of properties in the single-phase region. The results of this study allow users to take into account the uncertainties of the EOS in process model simulations. However, the results also show some limitations of the method:

- The uncertainty analysis does not present the effect of experimental uncertainty of the training data to the EOS because of the lack of information in some data sets of propane. The authors strongly encourage future research to take into account measurement uncertainties, if full measurement uncertainties have been obtained for each data point.

- The uncertainty analysis follows a linear propagation of error approach. This method is the only computationally tractable method that could be used for a nonlinear model like the equation of state studied here. Other more advanced sampling techniques (Markov Chain Monte Carlo (MCMC) sampling methods, etc.) are too computationally expensive to be

practically applied in technical applications.

- The method assumes a Gaussian distribution for the uncertainties of the EOS, and may fail to calculate the appropriate value at conditions which uncertainties distribute differently from the Gaussian distribution.

- The method assumes that all experimental data points are not correlated with each other. However, some training data points are not associated with comprehensive information on their measurement uncertainties and hence information on correlation between data, and more research is needed before the effect of correlation can be comprehensively accounted for by an uncertainty calculation method.

- The method does not involve the structural uncertainty in Table 1.

- The state-of-the-art fitting process of the EOS includes addition and removal of data points and constraints in an iterative fashion, and it is computationally expensive to involve the constantly changing and exact objective function for the fit of the EOS into the uncertainty calculation method. Thus the method has to use an approximated objective function of the EOS for the uncertainty calculation, and it is unknown how much error is introduced by the difference as a result.

Last but not least, it is recommended that users and developers of EOS perform this type of analysis in order to obtain insights about the uncertainties of the properties calculated by the EOS. Since the method does not incorporate other sources of uncertainties such as experimental uncertainties shown in Table 1, other methods to involve the experimental uncertainties should also be used to have a comprehensive understanding of uncertainties of EOS.

Supporting Information. Supporting Information is provided in three files. One file contains the mathematical derivation related to the differences of the expressions in this paper from that in some statistical textbooks, derivation for the Gibbs energy approximation in Fig. 1, the parameter identifiability analysis results, the Python code used to calculate the uncertainties of the thermodynamic properties of propane and the list of sources of experimental data being used in the study. The covariance matrix is provided in `pcov.csv`, and the correlation matrix is provided in `pcor.csv`. A Jupyter Notebook and the associated Python code and data files are also provided in a separate zip file `jupyter-notebook-demo.zip` to illustrate the use of the code.

Acknowledgement. The authors would like to acknowledge Allan H. Harvey at National Institute of Standards and Technology, Boulder, CO, USA for his comments on the manuscript draft. Commercial equipment, instruments, or materials are identified only in order to adequately specify certain procedures. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products identified are necessarily the best available for the purpose.

References

(1) Frutiger, J.; Andreasen, J.; Liu, W.; Spliethoff, H.; Haglind, F.; Abildskov, J.; Sin, G. Working fluid selection for organic Rankine cycles Impact of uncertainty of fluid properties. *Energy* **2016**, *109*, 987–997.

(2) Cheung, H.; Braun, J. E. A general method for calculating the uncertainty of virtual sensors for packaged air conditioners. *Int. J. Refrig* **2016**, *63*, 225–236.

(3) Huang, G.; Wang, S.; Xu, X. A robust model predictive control strategy for improving the control performance of air-conditioning systems. *Energy Convers. Manage.* **2009**, *50*, 2650–2658.

(4) Shan, K.; Wang, S.; Xiao, F.; Sun, Y. Sensitivity and uncertainty analysis of cool-

ing water control strategies. *HVAC&R Research* **2013**, *19*, 435–443.

(5) Reddy, T. A.; Niebur, D.; Andersen, K. K.; Pericolo, P. P.; Cabrera, G. Evaluation of the Suitability of Different Chiller Performance Models for On-Line Training Applied to Automated Fault Detection and Diagnosis (RP-1139). *HVAC&R Research* **2003**, *9*, 385–414.

(6) Proppe, J.; Reiher, M. Reliable estimation of prediction uncertainty for physicochemical property models. *J. Chem. Theory Comput.* **2017**, *13*, 3297–3317.

(7) Hukkerikar, A.; Sarup, B.; Ten Kate, A.; Abildskov, J.; Sin, G.; Gani, R. Group-contribution+ (GC+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilib.* **2012**, *321*, 25–43.

(8) Morgan, J.; Bhattacharyya, D.; Tong, C.; Miller, D. C. Uncertainty quantification of property models: Methodology and its application to CO2-loaded aqueous MEA solutions. *AIChE J.* **2017**, *61*, 1822–1839.

(9) Kamei, A.; Beyerlein, S. W.; Jacobsen, R. T. Application of nonlinear regression in the development of a wide range formulation for HCFC-22. *Int. J. Thermophys.* **1995**, *16*, 1155–1164.

(10) Wagner, W.; Pruß, A. The IAPWS Formulation 1995 for the Thermodynamic Properties of Ordinary Water Substance for General and Scientific Use. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387–535.

(11) Lemmon, E. W.; Jacobsen, R. T. Equations of State for Mixtures of R-32, R-125, R-134a, R-143a, and R-152a. *J. Phys. Chem. Ref. Data* **2004**, *33*, 593.

(12) Lemmon, E. W.; McLinden, M. O.; Wagner, W. Thermodynamic Properties of Propane. III. A Reference Equation of State for Temperatures from the Melting Line to 650 K and Pressures up to 1000 MPa. *J. Chem. Eng. Data* **2009**, *54*, 3141–3180.

(13) Lemmon, E. W. Pseudo-Pure Fluid Equations of State for the Refrigerant Blends R-410A, R-404A, R-507A, and R-407C. *Int. J. Thermophys.* **2003**, *24*, 991–1006.

(14) Cheung, H.; Bach, C. K. Prediction of uncertainty of 10-coefficient compressor maps for extreme operating conditions. IOP Conference Series: Materials Science and Engineering. 2015; p 012078.

(15) ASME, *ASME PTC 19.1-2013 Test Uncertainty Performance Test Codes*; The American Society of Mechanical Engineers: New York, 2013.

(16) JCGM, *JCGM 100:2008 Evaluation of measurement data Guide to the expression of uncertainty in measurement*; Joint Committee for Guides in Metrology, 2008.

(17) ASME, *ASME V&V 10-2006 Guide for Verification and Validation in Computational Solid Mechanics*; American Society of Mechanical Engineers: New York, 2006.

(18) Cheung, H.; Wang, S. A comparison of the effect of empirical and physical modeling approaches to extrapolation capability of compressor models by uncertainty analysis: A case study with common semi-empirical compressor mass flow rate models. *Int. J. Refrig* **2018**, *86*, 331–343.

(19) Feistel, R.; Lovell-Smith, J. W.; Saunders, P.; Seitz, S. Uncertainty of empirical correlation equations. *Metrologia* **2016**, *53*, 1079.

(20) Lovell-Smith, J. W.; Saunders, P.; Feistel, R. Unleashing Empirical Equations with Nonlinear Fitting and GUM Tree Calculator. *Int. J. Thermophys.* **2017**, *38*, 148.

(21) Frutiger, J.; Bell, I.; O'Connell, J. P.; Kroenlein, K.; Abildskov, J.; Sin, G. Uncertainty assessment of equations of state with application to an organic Rankine cycle. *Mol. Phys.* **2017**, *115*, 1225–1244.

(22) Lötgering-Lin, O.; Schöniger, A.; Nowak, W.; Gross, J. Bayesian Model Selection Helps To Choose Objectively between Thermodynamic Models: A Demonstration of Selecting a Viscosity Model Based on Entropy Scaling. *Ind. Eng. Chem. Res.* **2016**, *55*, 10191–10207.

(23) Lasala, S.; Chiesa, P.; Privat, R.; Jaubert, J.-N. Modeling the Thermodynamics of Fluids Treated by CO2 Capture Processes with PengRobinson + Residual Helmholtz Energy-Based Mixing Rules. *Industrial & Engineering Chemistry Research* **2017**, *56*, 2259–2276.

(24) Kline, S. J.; McClintock, F. A. Describing uncertainties in single-sample experiments. *Mech. Eng.* **1953**, *75*, 3–8.

(25) Sin, G.; Gernaey, K. V.; Lantz, A. E. Good modeling practice for PAT applications: Propagation of input uncertainty and sensitivity analysis. *Biotechnol. Progr.* **2009**, *25*, 1043–1053.

(26) Dong, Q.; Chirico, R. D.; Yan, X.; Hong, X.; Frenkel, M. Uncertainty Reporting for Experimental Thermodynamic Properties. *J. Chem. Eng. Data* **2005**, *50*, 546–550.

(27) Seber, G.; Wild, C. J. *Nonlinear regression*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1989.

(28) Sin, G.; Gernaey, K. V.; Neumann, M. B.; van Loosdrecht, M. C. M.; Gujer, W. Global sensitivity analysis in wastewater treatment plant model applications: Prioritizing sources of uncertainty. *Water Res.* **2011**, *45*, 639–651.

(29) Frutiger, J.; Marcarie, C.; Abildskov, J.; Sin, G. A Comprehensive Methodology for Development, Parameter Estimation, and Uncertainty Analysis of Group Contribution Based Property ModelsAn Application to the Heat of Combustion. *J. Chem. Eng. Data* **2016**, *61*, 602–613.

(30) Çengel, Y. A.; Boles, M. A. *Thermodynamics: An Engineering Approach*, 5th ed.; McGraw-Hill: Boston, 2005.

(31) Span, R. *Multiparameter Equations of State - An Accurate Source of Thermodynamic Property Data*; Springer, 2000.

(32) Tillner-Roth, R.; Baehr, H. D. A International Standard Formulation for the Thermodynamic Properties of 1,1,1,2-Tetrafluoroethane (HFC-134a) for Temperatures from 170 K to 455 K and Pressures up to 70 MPa. *J. Phys. Chem. Ref. Data* **1994**, *23*, 657–729.

(33) Bell, I. H.; Satyro, M.; Lemmon, E. W. Consistent Twu Parameters for More than 2500 Pure Fluids from Critically Evaluated Experimental Data. *J. Chem. Eng. Data* **2018**, *63*, 2402–2409.

(34) Coleman, H. W.; Steele, W. G. *Experimentation, Validation, and Uncertainty Analysis for Engineers*, 3rd ed.; Wiley: Hoboken, N.J, 2009.

(35) Thorade, M.; Saadat, A. Partial derivatives of thermodynamic state properties for dynamic simulation. *Env. Earth Sci.* **2013**, *70*, 3497–3503.

(36) Diky, V.; Chirico, R. D.; Muzny, C. D.; Kazakov, A. F.; Kroenlein, K.; Magee, J. W.; Abdulagatov, I.; Kang, J. W.; Frenkel, M. ThermoData Engine (TDE) software implementation of the dynamic data evaluation concept. 7. Ternary mixtures. *J. Chem. Inf. Model.* **2011**, *52*, 260–276.

(37) Frenkel, M.; Chirico, R. D.; Diky, V.; Yan, X.; Dong, Q.; Muzny, C. ThermoData Engine (TDE): Software Implementation of the Dynamic Data Evaluation Concept. *J. Chem. Inf. Model.* **2005**, *45*, 816–838.

(38) Diky, V.; Chirico, R. D.; Muzny, C. D.; Kazakov, A. F.; Kroenlein, K.; Magee, J. W.; Abdulagatov, I.; Kang, J. W.; Gani, R.; Frenkel, M.

ThermoData Engine (TDE): Software implementation of the dynamic data evaluation concept. 8. Properties of material streams and solvent design. *J. Chem. Inf. Model.* **2012**, *53*, 249–266.

(39) Diky, V.; Chirico, R. D.; Muzny, C. D.; Kazakov, A. F.; Kroenlein, K.; Magee, J. W.; Abdulagatov, I.; Frenkel, M. ThermoData Engine (TDE): Software implementation of the dynamic data evaluation concept. 9. Extensible thermodynamic constraints for pure compounds and new model developments. *J. Chem. Inf. Model.* **2013**, *53*, 3418–3430.

(40) Hawkins, D. M. The Problem of Overfitting. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1–12.

(41) Babyak, M. A. What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosom. Med.* **2004**, *66*, 411–421.