Vapour-Liquid Equilibrium of Mixtures

Module 4 · 2105603 Advanced Chemical Engineering Thermodynamics

Soorathep Kheawhom

9 August 2026

A parameter set is not a result.

A result is a parameter set
plus a statement of where it may be used.

INTRODUCTION

What this module establishes

Four descriptions of the same equilibrium, then the full working cycle: data, consistency, regression, and a stated domain of validity.

The chain of reasoning

An equation of state gives P-V-T behaviour → equilibrium is governed by chemical potential → fugacity gives chemical potential a usable form → equality of fugacity predicts phase equilibrium.

Module 4 is where that chain meets measurement.

Five questions

- Which description of VLE, and what does it cost?
- What is actually measured, and how well?
- Can this dataset be believed?
- Which model, fitted against what?
- Where may the answer be used?

PART I

Four descriptions of one equilibrium

Section 4.1 Raoult, modified Raoult, gamma-phi, phi-phi

Each rung adds physics the rung below omitted, and each addition costs something.

Raoult’s law, and its two assumptions

State the two assumptions separately, because they fail separately.

$$y_i P = x_i P_i^{\mathrm{sat}}$$

Assumption 1 — ideal liquid

Every molecule sees the same environment whatever the composition. True only when the two species are chemically almost identical: benzene and toluene, hexane and heptane.

Assumption 2 — ideal vapour

$\hat{\varphi}_i = 1$. Fails at high pressure, and also at low pressure when the vapour associates. Acetic acid dimerises in the vapour at one atmosphere.

What it costs

Nothing but $P_i^{\mathrm{sat}}(T)$. No binary parameter, no data, no fitting. That is why it survives as the first estimate.

Modified Raoult's law

Liquid non-ideality admitted; the vapour still treated as an ideal gas.

$$y_i P = x_i \gamma_i P_i^{\text{sat}}$$

Physical meaning

γ_i measures how much the liquid environment differs from the pure liquid. It is a property of the *mixture*, derived from an excess Gibbs energy, and it carries the reference state with it — Lewis-Randall here, so $\gamma_i \to 1$ as $x_i \to 1$.

Engineering implication

This is the description behind almost every distillation column ever designed at moderate pressure. Two fitted parameters per binary pair, from binary VLE data. Everything in Sections 4.2 to 4.4 is about where those two numbers come from and what they are worth.

The gamma-phi formulation

Both phases non-ideal — but described by two different kinds of model.

$$y_i \hat{\varphi}_i P = x_i \gamma_i P_i^{\text{sat}} \varphi_i^{\text{sat}} \exp\left[\frac{v_i^L \left(P - P_i^{\text{sat}}\right)}{RT}\right]$$

$\hat{\varphi}_i$

Vapour fugacity coefficient, from an equation of state. The two-term virial equation is enough below a few bar.

φ_i^{sat}

The same quantity for the pure saturated vapour. It is the reference the liquid fugacity is measured against, and it does not cancel.

Poynting factor

Corrects the liquid from P_i^{sat} to P. Small below a few bar, but it is not zero, and omitting it silently is not the same as knowing it is small.

The phi-phi formulation

One equation of state for both phases. The only option above the critical temperature of a component.

$$y_i \hat{\varphi}_i^V = x_i \hat{\varphi}_i^L$$

Why it is needed

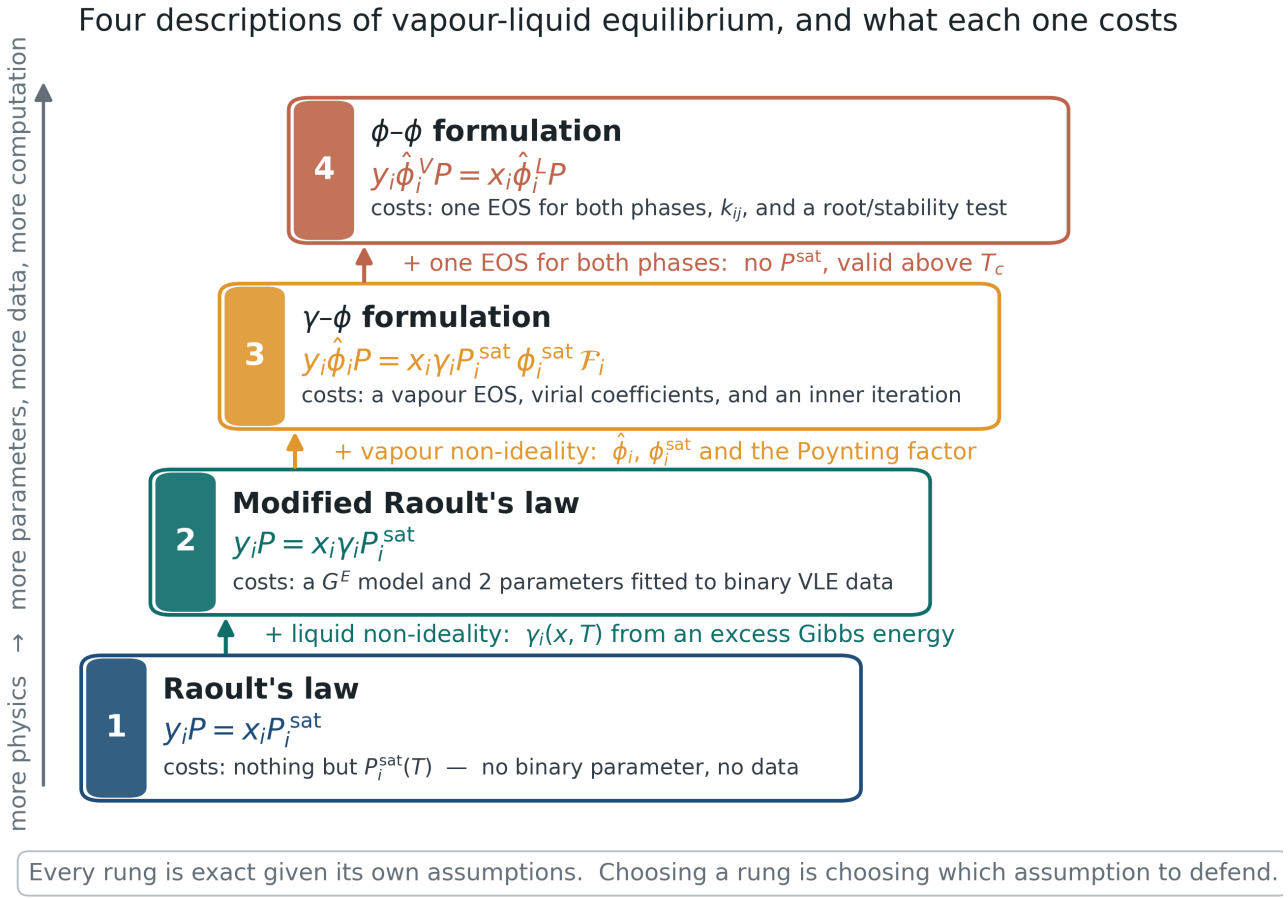
A supercritical component has no P_i^{sat}, so the whole right-hand side of the gamma-phi equation is undefined. Natural gas, hydrogen, supercritical CO_2: there is no vapour pressure to refer to.

What it costs

One equation of state that must describe a dense liquid and a dilute vapour with the same parameters, a binary interaction parameter k_{ij} fitted to data, and a root selection problem — the cubic has three roots and choosing wrongly gives a confidently wrong answer.

The ladder

Each rung is exact given its own assumptions. Choosing a rung is choosing which assumption to defend.



F4.1 · The four descriptions, with the physics added and the cost incurred at each step.

Choosing a description

A decision table, not a ranking.

SYSTEM	DESCRIPTION	BECAUSE
Benzene / toluene, 1 atm	Raoult	Chemically similar; no binary data needed
Ethanol / water, 1 atm	Modified Raoult	Strongly non-ideal liquid, near-ideal vapour
Acetic acid / water, 1 atm	Gamma-phi	Vapour associates; $\hat{\varphi} \neq 1$ at 1 atm
Methane / propane, 80 bar	Phi-phi	One component is supercritical
CO_2 / ethanol, 100 bar	Phi-phi with a G^E mixing rule	Supercritical *and* polar — see the appendix

The last row is the case neither classical description handles, and it is why the appendix to this deck exists.

PART II

From measurement to activity coefficient

Section 4.2 A reported γ is not a measurement.

It is four measured numbers and two assumptions.

What a VLE apparatus measures

Four numbers per point, each with its own uncertainty, and none of them is γ.

Temperature

Calibrated thermometer in the boiling liquid. Typically ± 0.05 K. Its error enters through $P_i^{\mathrm{sat}}(T)$, which is exponential in T.

Pressure

Manometer or transducer. Typically ± 0.05 kPa. Usually the *best* measured variable, which is why Barker's method uses it.

Liquid composition

Sample and analyse — refractometry, gas chromatography, density. Typically ± 0.001 mole fraction, and only if the calibration is honest.

Vapour composition

Condense and analyse. The hardest measurement: partial condensation, holdup, and entrainment all bias it. Typically ± 0.002, often worse.

Computing gamma

Two assumptions convert four measurements into an activity coefficient.

$$\gamma_i = \frac{y_i \Phi_i P}{x_i P_i^{\mathrm{sat}}(T)}, \qquad \Phi_i = \frac{\hat{\varphi}_i}{\varphi_i^{\mathrm{sat}}} \exp\left[-\frac{v_i^L (P - P_i^{\mathrm{sat}})}{RT}\right]$$

Assumption A — the vapour

Setting $\Phi_i = 1$ is modified Raoult. For acetone/chloroform at 50 °C the virial correction shifts $\ln \gamma_1$ by up to 0.008 — small, but **systematic**, and a consistency test reacts to systematic errors, not to scatter.

Assumption B — the vapour pressure

P_i^{sat} comes from a correlation, not from this experiment. Antoine coefficients are copied between references with unit errors more often than any other quantity in this subject. Check the correlation against the normal boiling point before you trust one activity coefficient.

Where the error actually comes from

Propagate, and the ranking is not the one students expect.

$$\sigma^2(\ln \gamma_1) = \left(\frac{\sigma_y}{y_1}\right)^2 + \left(\frac{\sigma_P}{P}\right)^2 + \left(\frac{\sigma_x}{x_1}\right)^2 + \left(\frac{\mathrm{d}\ln P_1^{\mathrm{sat}}}{\mathrm{d}T}\,\sigma_T\right)^2$$

The temperature term

P^{sat} is exponential in T, so $\mathrm{d}\ln P^{\mathrm{sat}}/\mathrm{d}T$ is of order $0.03\ \mathrm{K}^{-1}$ near the normal boiling point. A 0.05 K error is then worth 0.0015 in $\ln\gamma$ — comparable to the composition term.

The dilute ends

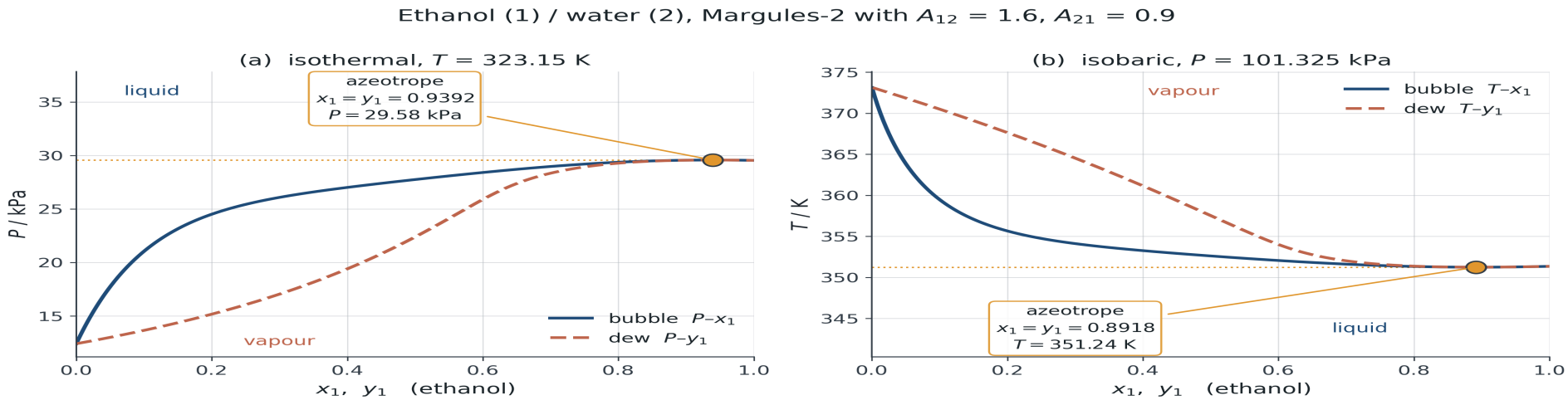
σ_x/x_1 blows up as $x_1 \to 0$. At $x_1 = 0.02$ a 0.001 error in composition is 5 % in γ_1. That is exactly where the model is least constrained and where γ^{∞} is read off.

The consequence

Quoting γ to four figures from a thermometer read to 0.1 K is arithmetic, not measurement. Report the uncertainty or do not report the figure.

Isothermal, isobaric, and the azeotrope

The two kinds of dataset, and the one composition where γ comes for free.



Illustrative Margules parameters from the course blueprint, not a regression of measured ethanol/water data. Azeotrope located with vlekit.flash.azeotrope.

Isothermal is easier to use

At constant T the Gibbs-Duhem equation has no enthalpy term. Isobaric data are more common — that is how a column runs — and harder, because T varies across the composition range.

The azeotrope is a free measurement

At $x_i = y_i$ the equilibrium condition gives $\gamma_1/\gamma_2 = P_2^{\text{sat}}/P_1^{\text{sat}}$ directly, with no vapour sample at all. In an old dataset it is often the only point that can be trusted without qualification.

F4.2 · The same system as P-x-y and as T-x-y, with the azeotrope marked on both.

PART III

Thermodynamic consistency

Section 4.3 Can this dataset be believed?

Gibbs-Duhem makes a binary VLE dataset over-determined, so the data can be tested against thermodynamics with no model assumed correct.

The premise

One of the four measured variables is redundant. The size of the prediction error is the test.

$$x_1 \frac{\mathrm{d} \ln \gamma_1}{\mathrm{d}x_1} + x_2 \frac{\mathrm{d} \ln \gamma_2}{\mathrm{d}x_1} = 0 \qquad (\text{constant } T,\ P)$$

Satisfied by construction

Any model derived from a single G^E satisfies this **identically** — the residual is zero to machine precision. Finding otherwise means the code is wrong, not the fluid.

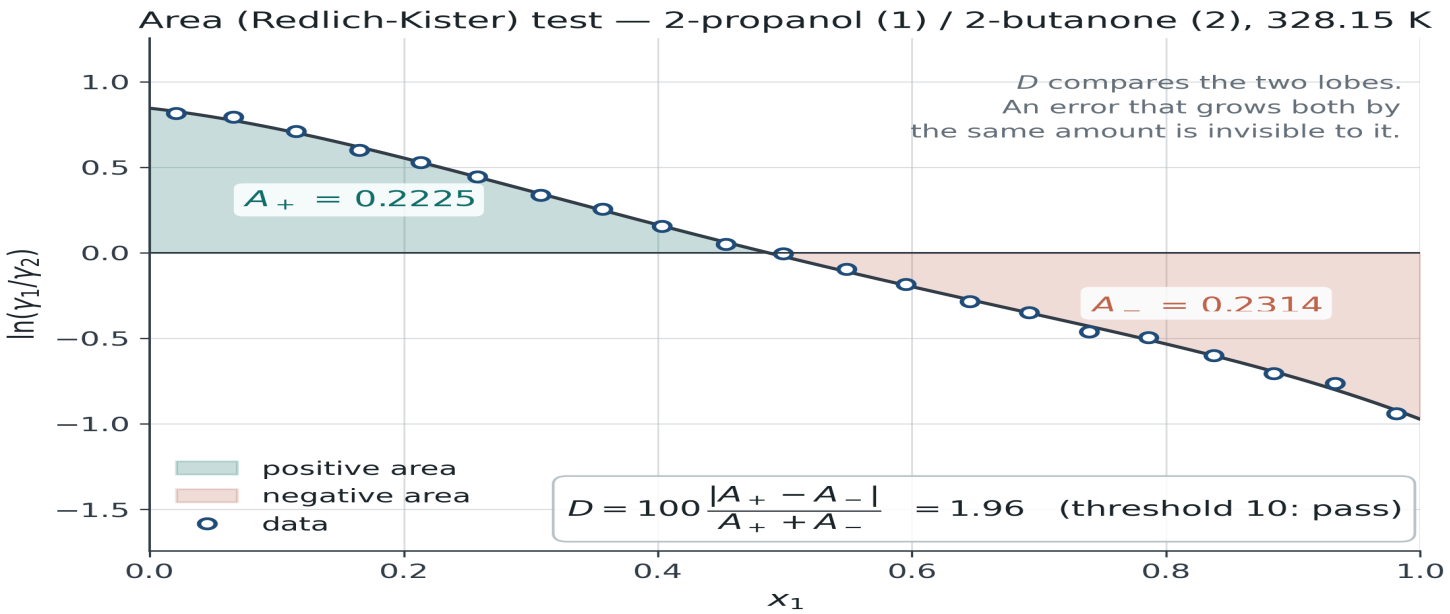
Satisfied by measurement

Data satisfy it only to within their errors. The two statements look alike on a slide and are completely different in practice. Everything that follows is a way of deciding whether the gap is noise or a real error.

The area test

The oldest test, the most quoted, and the weakest.

$$\int_0^1 \ln\frac{\gamma_1}{\gamma_2}\,\mathrm{d}x_1 = 0, \qquad D = 100\,\frac{|A^+ - A^-|}{A^+ + A^-}$$



F4.3 · The two areas whose difference must vanish, with D reported.

Why the area test is weak

An integral cannot see an error that changes sign. This is a structural limitation, not a matter of threshold.

The construction

Perturbing y_1 shifts the measured ratio by $\delta \ln(\gamma_1/\gamma_2) = \delta y_1 / [\, y_1(1-y_1)\,]$.

Choose $\delta y_1 = \varepsilon\, y_1(1-y_1)(2x_1-1)$ and the ratio shifts by $\varepsilon(2x_1-1)$ — odd about $x_1 = \frac{1}{2}$, so it contributes **exactly nothing** to the integral.

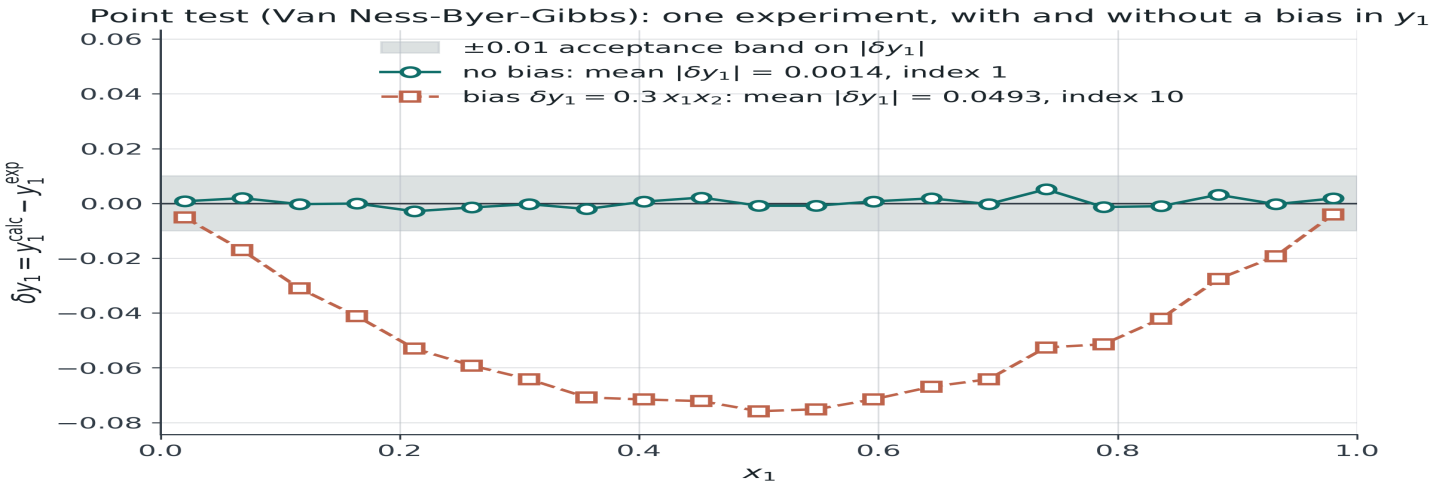
The result

At $\varepsilon = 0.4$ the dataset is grossly wrong. The area test reports $D = 1.3$ against a threshold of 10 — healthier than the merely noisy control — while the point and direct tests both return the worst possible index.

An area test is a necessary condition. It is never a sufficient one.

The point test

Van Ness, Byer and Gibbs: fit P-x only, then predict y. Errors cannot cancel, because nothing is integrated.



2-propanol (1) / 2-butanone (2) at 328.15 K, generated from NRTL(0.35, 0.62); Barker fit to *P-x* only, y_1 predicted.

The logic

Barker's method uses only P and x. If the data obey Gibbs-Duhem, y is already implied by the P-x curve, so the fit must reproduce it. Mean $|\delta y_1| < 0.01$ is the conventional pass.

The trap

A large residual in P means the **model** failed, not the data. Check $\mathrm{rms}\,\delta P$ before blaming the measurement, and use a model flexible enough to follow the P-x curve.

F4.4 · Point-test residuals for a consistent and an inconsistent dataset on common axes.

The direct test, and the 1-10 index

The same redundancy, expressed in the variable Gibbs-Duhem is actually written in. This is what modern journals expect.

$$\delta = \ln\left(\frac{\gamma_1}{\gamma_2}\right)_{\text{exp}} - \ln\left(\frac{\gamma_1}{\gamma_2}\right)_{\text{fitted}}$$

INDEX	RMS Δ	VERDICT
1 – 3	< 0.075	Good. Publishable without comment
4 – 5	0.075 – 0.125	Marginal. Requires an explanation
6 – 10	> 0.125	Reject, or diagnose the fault

The fitted values come from a Barker fit to P-x alone, so δ measures the data, not the fit — provided the model follows P-x.

Two more tests, and what they add

Different tests are blind to different things. That is the reason to run more than one.

Infinite dilution

Kojima’s test compares $(G^E/RT)/x_1x_2$ with $\ln(\gamma_1/\gamma_2)$ at each pure end point, where the two must coincide. Catches the dilute region, which the area test averages away and where γ^∞ is read off.

Herington $D - J$

The area test with an empirical allowance $J = 150\,\Delta T/T_{\min}$ for the temperature variation of an isobaric set. Widely required by referees; Wisniak has shown it both passes bad data and fails good data.

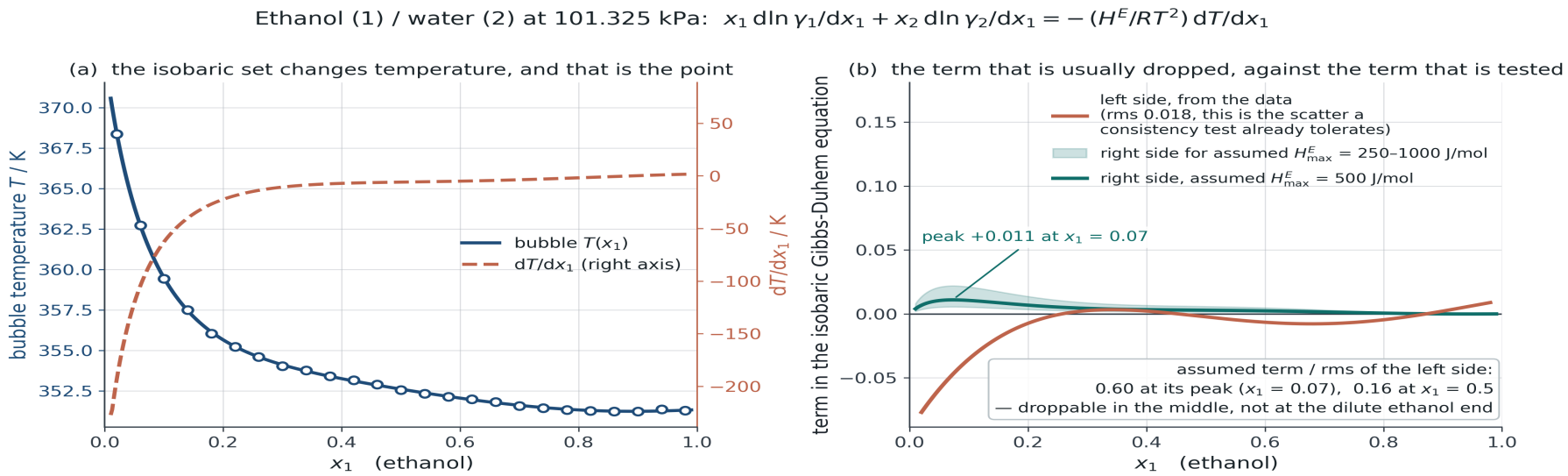
Wisniak L/W

Built on an energy balance rather than on G^E, so it responds to a different combination of the measurements. Worth quoting when a dataset is contested.

Isobaric data: the term that does not vanish

At constant P the temperature varies across the composition range, and the excess enthalpy term returns.

$$x_1 \frac{\mathrm{d}\ln\gamma_1}{\mathrm{d}x_1} + x_2 \frac{\mathrm{d}\ln\gamma_2}{\mathrm{d}x_1} = -\frac{H^E}{RT^2}\frac{\mathrm{d}T}{\mathrm{d}x_1}$$

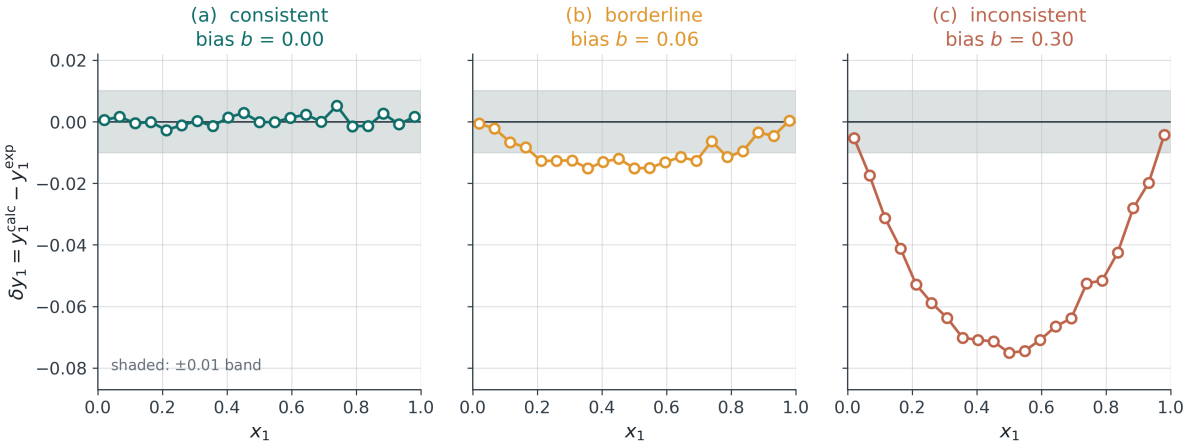


ASSUMED VALUE: $H^E = 4H^E_{max}x_1x_2$ with H^E_{max} = 500 J/mol (band: 250–1000 J/mol). This is an assumed magnitude for an alcohol/water mixture near its boiling range, not a measurement; no calorimetric data are used anywhere in this package. $\mathrm{d}T/\mathrm{d}x_1$ and the left-hand side are computed, from vlekit.regress.bubble_T and vlekit.consistency.gibbs_duhem_residual respectively.

F4.9 · Size of the enthalpy term against the left-hand side, ethanol/water at 101.325 kPa. The H^E magnitude is assumed and labelled as such on the figure.

Three datasets, five tests

2-propanol (1) / 2-butanone (2), isobaric at 101.325 kPa — the same experiment at three bias levels $\delta y_1 = b\,x_1 x_2$



consistency test	threshold	consistent	borderline	inconsistent
area (Redlich-Kister), D	10	0.86	9.66	**47.22**
Herington, $\lvert D - J \rvert$	10	2.05	6.75	**44.31**
point (VNBG), mean $\lvert \delta y_1 \rvert$	0.01	0.0014 (index 1)	0.0095 (index 4)	**0.0492 (index 10)**
direct (Van Ness), rms δ	0.05	0.0117 (index 1)	0.0471 (index 2)	**0.2398 (index 10)**
infinite dilution (Kojima)	30	24.0	16.1	**56.0**

Green = pass, red = fail. Every value is what vlekit.consistency.run_all returned for the set plotted above it.
At b = 0.06 every test passes — the area test by 0.3 of its threshold, the point test by 0.0005 — yet the point index is already 4, which is marginal.
Between b = 0.06 and b = 0.08 four of the five verdicts flip together: these are not independent detectors, but one residual read at different thresholds.
Report a dataset with all five, never with the one that happens to pass.

F4.5 · The same experiment at three bias levels, with every statistic tabulated.

Diagnosis, not verdict

A failure is the beginning of an investigation. Each fault has a signature.

Constant offset in δy	Vapour sampling bias — partial condensation, or a GC calibration offset
Residual bowed, one sign throughout	Wrong Antoine constants, or the vapour-phase correction omitted when it mattered
Sign change near the azeotrope	Composition analysis non-linear across the range
Fails at one end only	Impurity concentrating in the volatile or the heavy component
Large $\mathrm{rms}\,\delta P$ as well	The model, not the data. Refit with a more flexible G^E before concluding anything

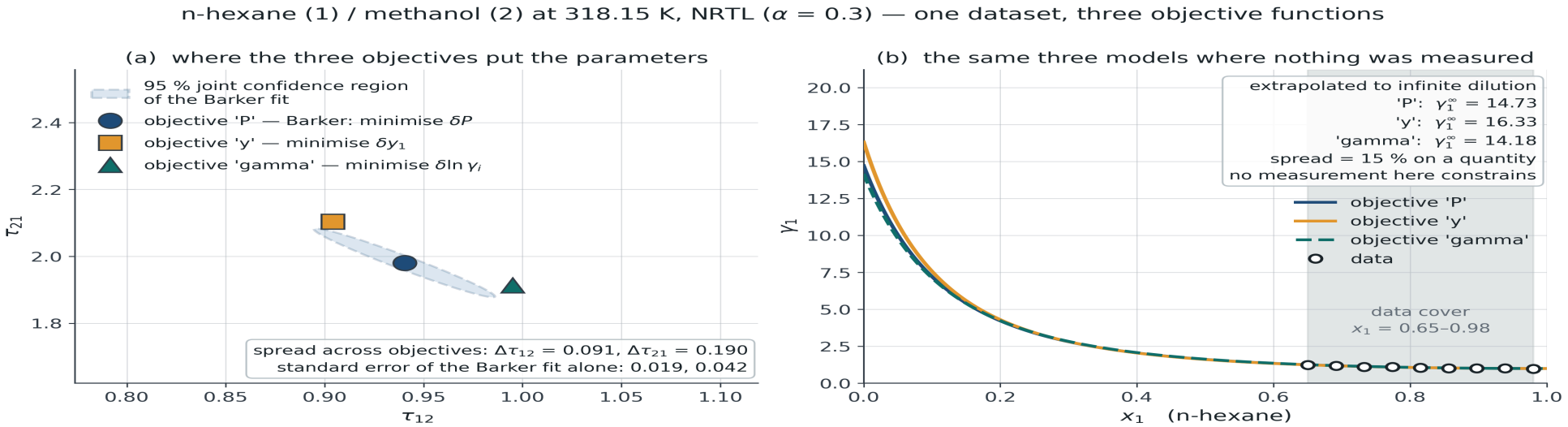
PART IV

Building a model from data

Section 4.4 Regression converges to something for almost any input. The question is never whether a fit exists. It is what the fit is worth.

The objective function is a modelling decision

Same data, same model, four objectives, four answers. This is not numerical noise; it is a choice about which measurement you trust.



The spread is not small

Across objectives the parameters move by several times the standard error that any one of them reports. An uncertainty that ignores the choice of objective is understating the real one.

Minimise what you measured well

Fitting $\ln \gamma$ weights the dilute ends heavily — precisely where the measurement is worst. Fitting y fits the least accurate variable. Fitting P uses the best one.

F4.6 · Parameters and predictions from three objective functions on one dataset.

Barker’s method

Fit to P-x alone. Do not fit y at all.

$$P^{\text{calc}}(x_1) = x_1\gamma_1 P_1^{\text{sat}} + x_2\gamma_2 P_2^{\text{sat}}, \qquad \min_{\boldsymbol{\theta}} \sum_k \left(P_k^{\text{calc}} - P_k^{\text{exp}}\right)^2$$

It uses the best data

P and x are the two accurately measured variables. Nothing in the objective depends on the vapour sample.

It makes the tests possible

A y that was never fitted can be compared with a y that was measured. That comparison is the point test and the direct test. Fit y and you destroy your own diagnostic.

Many datasets have no y

Static-cell measurements report P-x only, by design. Barker is not a compromise for those data — it is the intended method.

Does the parameter earn its place?

A model with more parameters cannot fit worse. The residual alone will always choose the larger model.

$$\mathrm{AIC} = n \ln \frac{\mathrm{SSE}}{n} + 2k$$

What AIC answers

Whether the improvement in fit is worth the parameter. A difference below about 2 is conventionally no evidence at all — the data cannot separate those models.

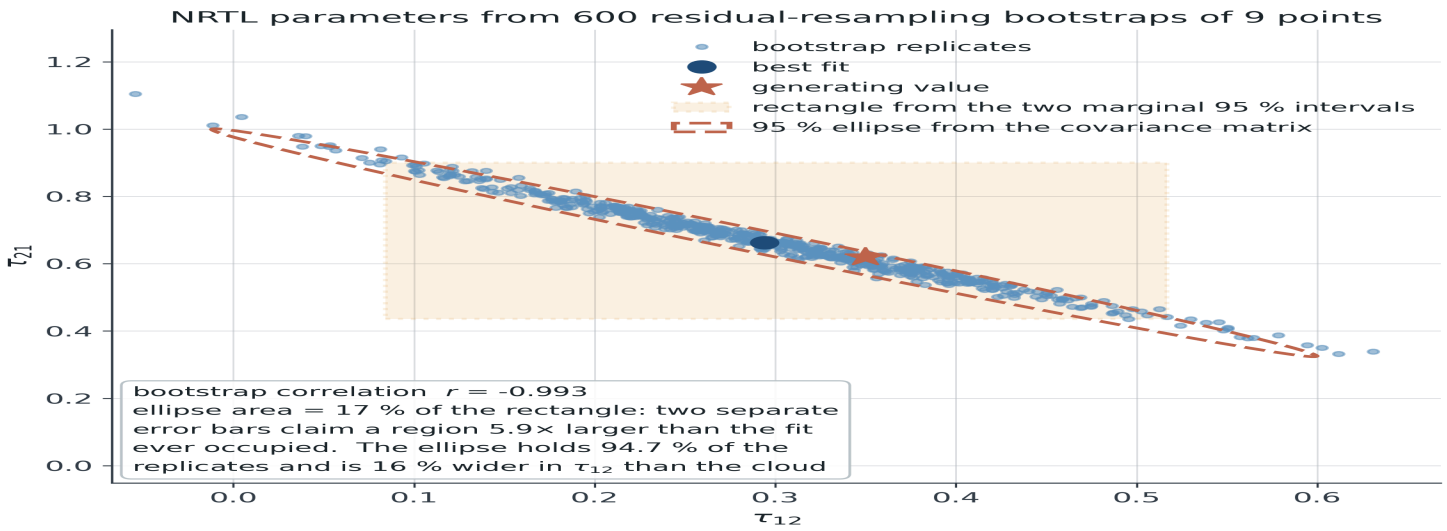
Typical outcome

On twenty-one points at realistic noise, Margules, van Laar, Wilson, NRTL and UNIQUAC land within 2 AIC units of one another. Any claim that one "describes the system better" needs more than that dataset.

*AIC compares models fitted the same way. It does **not** compare objective functions: each minimises a different residual vector, so the likelihood being approximated is not the same likelihood.*

Parameter uncertainty is a region, not two numbers

Two standard errors quoted separately describe a rectangle the fit never occupied.



What the cloud shows

A correlation near -1: a long narrow ridge of parameter pairs that fit equally well. The two parameters are not separately identifiable from one isotherm.

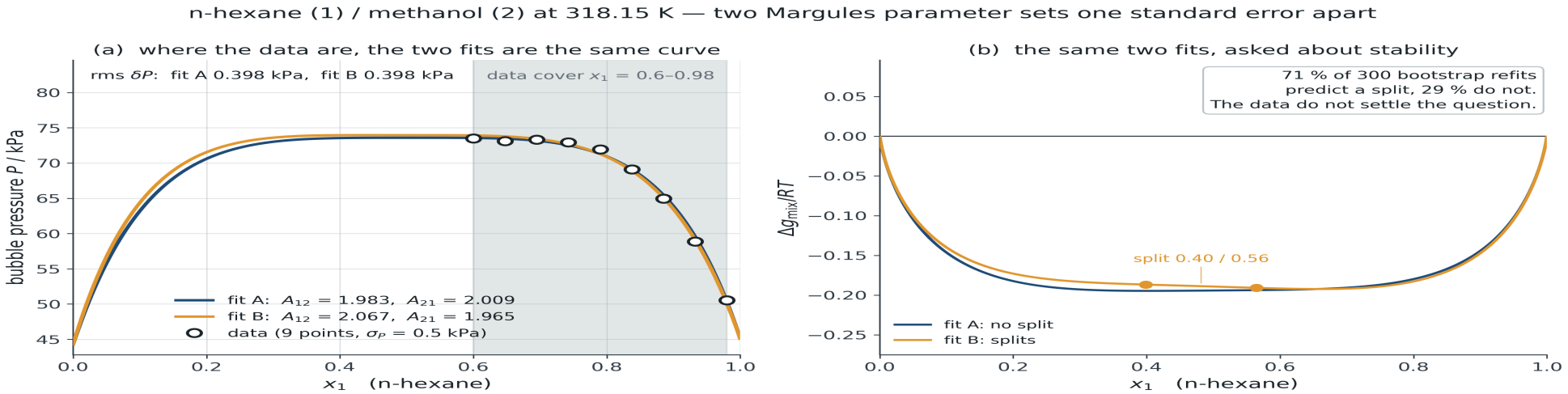
Why bootstrap

The linearised errors assume a quadratic optimum and normal residuals. For two parameters and twenty points, neither holds well. Resampling makes no such assumption and is transparent enough to explain in a report.

F4.7 · Bootstrap replicates with the linearised ellipse over them.

The consequence that actually matters

Two fits the data cannot separate can disagree completely about the region you care about.



Both sets lie one joint standard error from the same Barker fit along its least-determined direction; their residuals in P differ by 0.0004 kPa, against a scatter of 0.5 kPa.

Same evidence, different claims

Fitted only where data exist, several models agree to within the scatter. They then report γ^{∞} differing by tens of per cent, and they do not agree on whether the liquid splits into two phases.

Wilson is the clearest case

Wilson is structurally incapable of predicting a miscibility gap for **any** parameters. Fitted to a system that has one, it buys the same fit with an absurd γ^{∞} instead. That number is a symptom of the functional form, not a property of the fluid.

F4.8 · Statistically indistinguishable fits extrapolated beyond the data.

Two ways to find out whether you have learned the noise

Reporting the residual on the points you fitted is not evidence.

Cross-validation by composition

Train on the middle of the range and predict the dilute ends, then the reverse. Report both. A model that fits its training points three times better than it predicts the held-out ones has learned the scatter.

Splitting on alternate points tests almost nothing — neighbouring points carry almost the same information.

Against a prediction

UNIFAC predicts γ from group contributions with nothing fitted to your data. Your model must beat it — it saw the answer. The question is by how much, and whether the margin survives outside the fitted range.

State the margin in a quantity someone would act on: γ^{∞}, or the azeotrope composition. Not in rms.

Temperature, and the limits of two numbers

Parameters fitted at one temperature, used at another.

$$\tau_{ij} = \exp\left(-\frac{\Delta u_{ij}}{RT}\right) \quad \Longrightarrow \quad \left.\frac{\partial (G^E/RT)}{\partial T}\right|_x = -\frac{H^E}{RT^2}$$

What the extrapolation assumes

Fitting τ directly at one temperature and reusing it assumes G^E/RT is independent of T — that is, $H^E = 0$. Fitting Δu instead assumes Δu is constant, which is weaker but still an assumption.

In practice

Twenty or thirty kelvin is usually tolerable for a non-associating mixture. Across an alcohol/water column from reboiler to condenser it is not, and the error shows up exactly where the column is hardest to control.

The deliverable

A regression report that stops at the parameters is not finished.

1. **Model, parameters and objective.**
   Which G^E form, which two numbers, minimised against what, with which vapour-phase assumption.

2. **Range of validity.**
   The composition, temperature and pressure range of the data, stated as a range and not as a midpoint.

3. **Uncertainty and correlation.**
   A confidence region or a covariance, not two independent error bars.

4. **What was extrapolated.**
   Every quantity you report that lies outside the data — γ^{∞}, an azeotrope, a predicted phase split — named as an extrapolation.

PART V

Using the model

Section 4.5 Bubble, dew, flash — and the question none of them asks.

Four algorithms, not one

The same equilibrium condition with different variables specified. They do not converge equally well.

Bubble P

Given x, T. Direct: $\gamma_i(x, T)$ is known immediately, so P and y follow with no iteration for an ideal vapour.

Bubble T

Given x, P. One outer iteration on T, because P_i^{sat} and γ_i both move with it.

Dew P

Given y, T. The liquid composition is unknown, so γ_i must be updated every iteration. Slower and less robust than bubble.

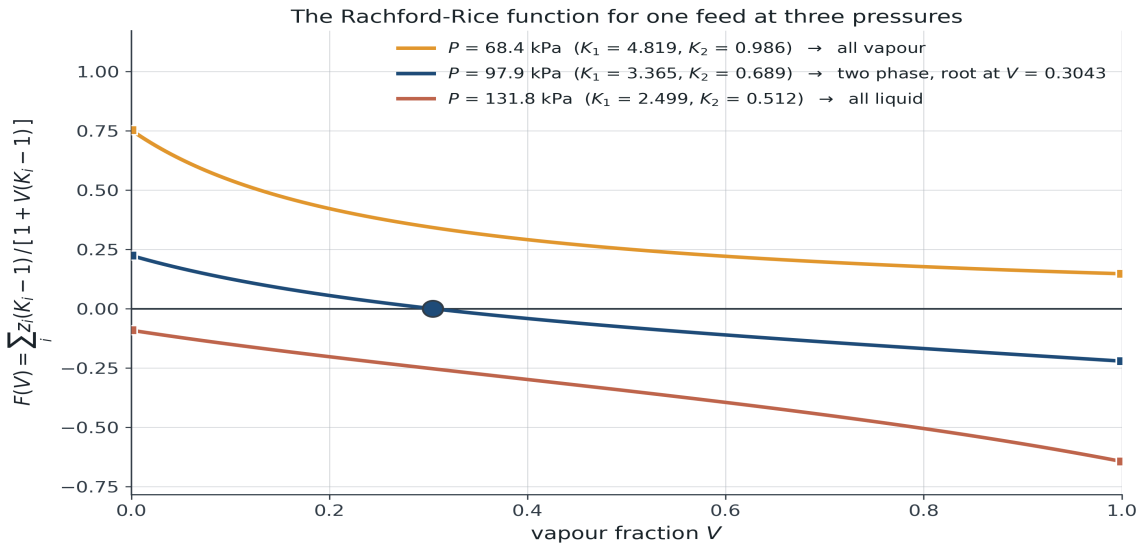
Dew T

Given y, P. Both loops at once. This is the one that fails in a flowsheet at three in the morning.

Rachford-Rice

Monotonic in the vapour fraction, and therefore safe. Show the plot before the algebra.

$$F(V) = \sum_i \frac{z_i(K_i - 1)}{1 + V(K_i - 1)} = 0, \qquad K_i = \frac{y_i}{x_i} = \frac{\gamma_i P_i^{\mathrm{sat}}}{P}$$



Ethanol (1) / water (2), Margules-2 with $A_{12} = 1.6$, $A_{21} = 0.9$ (illustrative, as in F4.2). Feed $z_1 = 0.2$ at $T = 360.00$ K; bubble 119.83 kPa, dew 75.96 kPa.
Largest dF/dV over all three curves = -0.126: F is strictly decreasing, so there is at most one root and the signs of $F(0)$ and $F(1)$ settle the phase state before any iteration begins.

F4.10 · One feed at three pressures. The signs of $F(0)$ and $F(1)$ settle the phase state before any iteration begins.

The question the flash never asks

Rachford-Rice was told there are two phases, and it will find two phases.

What is missing

Nothing in the equilibrium equations checks **how many** phases there should be. A fitted model that predicts a liquid-liquid split will still return a single perfectly reasonable-looking liquid composition inside the gap.

What answers it

The Gibbs energy of mixing and a tangent construction — not the equilibrium equations. A liquid splits when Δg_{mix} dips below its own chord. That is Module 5, and it is why the capstone spans both modules.

Mixing rules for cubic equations of state

The phi-phi route needs mixture parameters. The classical answer is quadratic in composition.

$$a_{\text{mix}} = \sum_i \sum_j x_i x_j \sqrt{a_i a_j}\,(1 - k_{ij}), \qquad b_{\text{mix}} = \sum_i x_i b_i$$

k_{ij} is fitted, not predicted

It comes from binary data, exactly like the activity model parameters, and it carries the same obligations: a range of validity and an uncertainty.

Why it fails for polar mixtures

The quadratic form assumes random mixing. Substituting it into the cubic gives an excess Gibbs energy that is thermodynamically equivalent to a one-constant Margules model: symmetric, one parameter. Hydrogen-bonding mixtures are neither symmetric nor one-parameter, and k_{ij} can scale that curve but cannot change its shape.

One idea, stated and not derived

A mixing rule can be constructed so that a cubic equation of state reproduces a chosen excess Gibbs energy model.

What that buys

The pressure range and the supercritical components of an equation of state, together with the composition behaviour of NRTL or UNIQUAC. This is how a simulator handles CO_2 with ethanol at 100 bar.

Where to find it

Huron-Vidal and Wong-Sandler. Four appendix slides follow this one, and the algebra is in `vle-mixing-rules-note.pdf` on the course page. Not lectured, not examined — but the property-method dropdown in a simulator says PRWS and PSRK, and four pages is what it takes to know what that means.

CLOSING

What you must be able to do

The module in six statements.

1. **Choose a description and defend the assumption.**
   Raoult, modified Raoult, gamma-phi or phi-phi — and say which assumption you are relying on.

2. **Turn T, P, x, y into γ with its uncertainty.**
   Including the vapour-phase assumption, stated, and the error in P^{sat}.

3. **Test a dataset and diagnose the failure.**
   All five tests, and a signature-based diagnosis rather than a verdict.

4. **Fit, and say what the fit is worth.**
   Objective chosen deliberately, confidence region reported, extrapolations named.

5. **Compare against a prediction.**
   Know what the two fitted parameters bought over UNIFAC.

6. **Know what the calculation does not check.**
   A flash assumes the number of phases. Module 5 supplies the missing test.

Appendix

EoS-G^E mixing rules

Huron-Vidal and Wong-Sandler

Not lectured. Not examined. Read the note if this is your problem.

A.1 · Two worlds that do not meet

Equations of state

Handle high pressure and supercritical components. One model for both phases, no reference vapour pressure required. Fail on polar and hydrogen-bonding mixtures, because the quadratic mixing rule assumes random mixing.

Activity coefficient models

Handle polar and associating liquids, through local-composition terms fitted to data. Cannot cross the critical point, and need a reference fugacity that a supercritical component does not have.

The engineering problem is the mixture that is both — CO_2 with ethanol, water with a light gas — and neither description can be pushed into the other's territory. The mixing rule is the bridge.

A.2 · Huron-Vidal

Match the excess Gibbs energy of the equation of state to that of an activity model in the infinite-pressure limit, where the cubic collapses to $V = b$ and the expression closes.

$$\frac{a_{\text{mix}}}{b_{\text{mix}}} = \sum_i x_i \frac{a_i}{b_i} - \frac{g^E_\infty}{C}, \qquad C = \frac{1}{\delta_1 - \delta_2} \ln \frac{1+\delta_1}{1+\delta_2}$$

The constant

$C = \ln 2 = 0.6931$ for Soave-Redlich-Kwong; $C = \ln(1+\sqrt{2})/\sqrt{2} = 0.6232$ for Peng-Robinson.

The practical cost

Published low-pressure NRTL parameters cannot be reused. The reference state is the infinite-pressure limit, not the real liquid, so the parameters must be refitted — and they are not transferable between SRK and PR either, because C differs.

A.3 · Wong-Sandler

Keep the infinite-pressure condition and add the low-density one: the second virial coefficient must be quadratic in composition.

$$B = b - \frac{a}{RT}, \qquad Q = \sum_i \sum_j x_i x_j B_{ij}, \qquad D = \sum_i x_i \frac{a_i}{b_i RT} + \frac{g^E_\infty}{CRT}$$

$$b_{\text{mix}} = \frac{Q}{1 - D}, \qquad a_{\text{mix}} = RT \, \frac{Q\,D}{1 - D}$$

What is gained

Correct at both ends of the density range. b_{mix} stops being linear in composition, which is the structural difference from Huron-Vidal.

What it costs

One extra binary parameter in the B_{ij} combining rule, on top of the g^E model parameters. Pure-component reduction ($b_{\text{mix}} = b_i$, $a_{\text{mix}} = a_i$) is the unit test to run before trusting an implementation.

A.4 · What the names mean

PROPERTY METHOD	IS	PARAMETERS
PRWS, RKSWS	Wong-Sandler on PR or SRK	Refit; not the published low-pressure set
PRMHV2, RKSMHV2	Modified Huron-Vidal, second order	Zero-pressure reference — published UNIFAC/UNIQUAC parameters **can** be reused
PSRK	Predictive SRK: MHV1 with UNIFAC	Predictive; no binary VLE data needed

MHV1 and MHV2 (Michelsen, 1990) use a zero-pressure rather than an infinite-pressure reference, which is precisely what makes published parameters transferable. That is their entire practical advantage.

Full derivations, the sign-convention warning, and the references: `vle-mixing-rules-note.pdf`.