

Highlights

A data-driven multiscale model for reactive wetting simulations

Jaideep Ray, Jeffrey S. Horner, Ian Winter, David J. Kemmenoe, Edward R. Arata, Michael Chandross, Scott A. Roberts, and Anne M. Grillet

- Developed a method to construct data-driven wetting model, from a small set of molecular dynamics simulations; studied wetting of molten silver-aluminum alloy on a Kovar™ substrate for brazing
- Performed molecular dynamics simulation to predict surface tension and wetting angle
- Calibrated probabilistic, data-driven surrogate model using limited training data and integrated wetting model into engineering-scale finite element brazing model
- Compared engineering model to experiment, identifying missing physics in simulations

A data-driven multiscale model for reactive wetting simulations

Jaideep Ray^{a,*}, Jeffrey S. Horner^b, Ian Winter^b, David J. Kemmenoe^b, Edward R. Arata^b,
Michael Chandross^b, Scott A. Roberts^b, and Anne M. Grillet^b

^a*Sandia National Laboratories, PO Box 969, Livermore, CA, 94550,*

^b*Sandia National Laboratories, PO Box 5800, Albuquerque, NM, 87185,*

Abstract

We describe a data-driven, multiscale technique to model reactive wetting of a silver-aluminum alloy on a KovarTM (Fe-Ni-Co alloy) surface. We employ molecular dynamics simulations to elucidate the dependence of surface tension and wetting angle on the drop's composition and temperature. A design of computational experiments is used to efficiently generate training data of surface tension and wetting angle from a limited number of molecular dynamics simulations. The simulation results are used to parameterize models of the material's wetting properties and compute the uncertainty in the models due to limited data. The data-driven models are incorporated into an engineering-scale (continuum) model of a silver-aluminum sessile drop on a KovarTM substrate. Model predictions of the wetting angle are compared with experiments of pure silver spreading on KovarTM to quantify the model-form errors introduced by the limited training data versus the simplifications inherent in the molecular dynamics simulations. The paper presents innovations in the determination of "convergence" of noisy MD simulations before they are used to extract the wetting angle and surface tension, and the construction of their models which approximate physio-chemical processes that are left unresolved by the engineering-scale model. Together, these constitute a multiscale approach that integrates molecular-scale information into continuum scale models.

Keywords: reactive wetting, brazing, molecular dynamics, Markov process, continuum modeling, sessile drops

PACS: 47.11.Mn, 47.10.ab

*Corresponding author

Email address: jairay@sandia.gov (Jaideep Ray)

1. Introduction

Liquids wetting (or not wetting) solids are ubiquitous in nature, with examples including ethanol on glass or water on super-hydrophobic lotus leaves, respectively. Many industrial applications, such as manufacture of photographic films and soldering/brazing, rely on reactive wetting where the contact angle is mediated by reactions at the interface of the liquid and solid [1]. In these cases, the wetting dynamics are affected by multiple interacting physical processes including reaction kinetics, diffusion of reactants or reaction products, flow of the liquid and thermal transport. Since it is an inherently multi-physics problem, reactive wetting is still not well understood and manufacturing processes that rely on reactive wetting are difficult to model [1, 2]. An exemplar that embodies the complexities of reactive wetting is brazing, where two surfaces are joined by a filler metal that is heated above its melting point (and notably below the melting point of the substrates to be joined), and bonds them together (both physically and chemically) on cooling. Understanding and modeling these processes is quite challenging [3] due to the multiple physical and chemical processes that must be considered. Wetting by the filler metal depends not only on the temperature but also the composition of both the filler metal and the substrates. The compositions change over time, as components are liberated from the substrates through dissolution or, in the case of active brazing, chemical reactions, and absorbed by the filler metal. While studies of single metal fillers are abundant in the brazing literature [4], data are scarce for high fidelity studies of wetting of multicomponent alloys as a function of temperature and composition. This makes brazing an ideal application for the development of reduced order, data-driven models that can accurately predict the material properties that control reactive wetting behavior.

The motivation for our research is to lay a foundation for engineering-scale (millimeter to centimeter) simulations of a brazing process, with the specific goal of studying issues such as run-out where the filler material spreads beyond the surfaces being joined. Specifically, we seek a method to construct wetting models, trained on data from high-fidelity (HF) simulations, that can be integrated into engineering simulators of manufacturing processes. The exemplar system of interest consists of a Ag-Al braze filler joining KovarTM (a speciality alloy composed of 29% Ni, 17% Co, 0.2% Si, 0.3% Mn, 0.01% C by mass, with the balance being Fe [5, 6]) surfaces, a system that has been previously investigated [7]. The primary technical challenge lies in modeling the reactive wetting at the filler-substrate interface in

33 a computationally efficient manner so that it may be reproduced within an engineering-
34 scale simulation that cannot resolve all the relevant physics. Explicit resolution of such
35 processes, using phase-field or molecular dynamics models, can only be performed for
36 nanometer-sized drops [8, 9] (or smaller), and the simulations are very computationally
37 expensive. Additional challenges lie in enforcing flow boundary conditions on the filler-
38 substrate interface as it evolves.

39 In this paper, we describe a data-driven augmentation of conventional finite element
40 models (FEM) that can be used to simulate active brazing at engineering scale. We begin
41 with the assumption that macroscale reactive wetting at the filler-substrate interface de-
42 pends on two fundamental interrelated phenomena, namely the surface tension of the liquid
43 filler alloy $\sigma(T, Y)$ and its wetting angle on the substrate $\theta(T, Y)$, where T is the absolute
44 temperature of the molten filler and Y is the composition of the filler alloy. Our hypothesis
45 is that data-driven models of $\theta = F(T, Y; \mathbf{w})$ and $\sigma = G(T, Y; \mathbf{v})$ can be constructed based
46 on HF simulations of a filler metal drop on a flat substrate surface. Here $F(T, Y; \mathbf{w})$ and
47 $G(T, Y; \mathbf{v})$ are the regression models and (\mathbf{w}, \mathbf{v}) are the regression models' parameters (or,
48 loosely speaking, the sensitivity of θ and σ to T and Y). These models can be constructed
49 by fitting to training data (TD) $(X = (T, Y)_i, Z = (\sigma, \theta)_i, i = 1 \dots N)$ by executing the
50 HF simulations at a sampled set of $X = (T, Y)_i$ inputs and extracting $Z = (\sigma, \theta)_i$ from the
51 simulation outputs. The FEM, augmented with the models $F(T, Y; \mathbf{w})$ and $G(T, Y; \mathbf{v})$, will
52 be demonstrated by simulating a millimeter-sized silver sessile drop relaxing to equilibrium
53 on a KovarTM surface. In doing so, we will also describe how the boundary conditions are
54 imposed in FEM at an evolving filler-substrate boundary.

55 The choice of the HF simulations is a difficult one. Ideally, the TD for the wetting
56 model should include reactive processes and the formation of intermetallic phases (IMP)
57 that are known to exist in the exemplar system [7]. However, the phase-field [8, 9] and re-
58 active molecular dynamics (MD) simulations [10, 11] that can simulate these processes are
59 computationally expensive and will not allow the assembly of TD in a timely fashion. Con-
60 sequently, we will employ MD simulations using the embedded atom method (EAM; [12]),
61 without reaction between the braze alloy and substrate, and in the process incur a modeling
62 error. The key quantities for inclusion in the FEM are the surface tension and the wetting
63 angle, and it is only the wetting angle that may be significantly affected by our choice of
64 HF dataset. However, the wetting angle has contributions from the surface energy at the

65 liquid-gas, solid-gas and the liquid-solid interfaces, all of which could, depending on the
66 extant reactions, incur a model-form error due to the absence of IMP in the TD simula-
67 tions. This paper, in essence, will investigate whether this data-driven methodology can
68 lead to stable and accurate engineering simulation, when embedded in a partial differential
69 equation (PDE) continuum model. It will also quantify the PDE’s model-form error (due
70 to the approximate wetting models) and provide a path forward for ameliorating them via
71 multi-fidelity modeling (discussed in Sec. 5.4).

72 The technical challenges presented here are (1) sufficient sampling of the (T, Y) space,
73 that will necessarily use a limited number of samples N due to the computational cost of MD
74 simulations (2) devising a proper functional form for $F(T, Y; \mathbf{w})$ and $G(T, Y; \mathbf{v})$ conditional
75 on limited TD, (3) deciding on when an MD simulation has converged sufficiently (given
76 limited computational resources/time) so that (σ, θ) may be extracted and (4) ensuring that
77 the FEM, with embedded $F(T, Y; \mathbf{w})$ and $G(T, Y; \mathbf{v})$, is stable and can reproduce, at least
78 qualitatively, the process of a sessile drop equilibrating to its equilibrium shape. The sources
79 of prediction error (of the data-driven models) are the fidelity of the MD simulations and
80 the limited nature of the TD (and the consequent simplicity of $F(T, Y; \mathbf{w})$ and $G(T, Y; \mathbf{v})$).
81 Identifying the relative importance of these sources of errors is a goal of this paper.

82 We will address the issue of constructing $F(T, Y; \mathbf{w})$ and $G(T, Y; \mathbf{v})$ using limited TD in
83 two ways. First, we will ensure, in a data-driven manner, that the complexity of $F(T, Y; \mathbf{w})$
84 and $G(T, Y; \mathbf{v})$ is consistent with the information in the TD. Second, we will estimate the
85 model parameters \mathbf{w} and \mathbf{v} as a joint probability density function (JPDF) so that we capture
86 the uncertainty due to limited TD. Finally, we will devise a way to detect the convergence of
87 fluctuating (i.e., time-dependent) MD data for θ , and extract a wetting angle *with a measure*
88 *of the “extraction uncertainty”* that is then incorporated into the uncertainty estimates of
89 $F(T, Y; \mathbf{w})$ and $G(T, Y; \mathbf{v})$ (specifically, the JPDF of \mathbf{w} and \mathbf{v}). The convergence detection
90 technique is based on approximating the noisy MD trace for $\theta(t)$ (where t represents time
91 in the MD simulations) as a Markov chain (MC). If the MD simulation is determined to
92 have not yet converged, our convergence detection method estimates how much longer the
93 simulation is required to run to reach equilibrium.

94 The paper is laid out as follows. In Sec. 2 we review previous work on reactive wet-
95 ting and the various techniques we use. In Sec. 3, we formulate the modeling problem
96 (data-driven and the FEM). In Sec. 4, we describe the generation of the MD training and

97 experimental data. In Sec. 5, we present results of modeling σ and θ and the performance of
98 the FEM. This section will also contain a comparison of FEM predictions to experimental
99 measurements of a nominally pure silver drop on a KovarTM surface, with the aim of esti-
100 mating the magnitude of the model-form error in $F(T, Y; \mathbf{w})$ and $G(T, Y; \mathbf{v})$. We draw our
101 conclusions in Sec. 6.

102 2. Literature review and background

103 In this section we review literature on experiments and MD simulations of reactive
104 wetting, with emphasis on the fundamental physical processes. We also discuss the existing
105 literature on FEM approaches (at engineering scales) that we employ, and the modifications
106 necessary to include data-driven models of σ and θ . Finally, we discuss TD generation
107 and data-driven modeling. Note that finite element studies of brazed joints, especially
108 their residual stresses (due to a mismatch of thermal expansion coefficients of the materials
109 involved), have been an active field of research (see Ref. [13] for a review), but they generally
110 rely on experiments and/or temperature-tabulated material properties for specific filler-
111 substrate combinations [14, 15, 16] rather than on the development of a generalized model
112 for interfacial properties as presented here.

113 2.1. Reactive wetting and its models

114 The wetting of a solid metallic substrate by a molten metal or alloy can be of two
115 types – inert (or passive) and reactive. In reactive wetting, the two phases (solid and
116 liquid) may simply dissolve into each other or form a new intermetallic phase (IMP) that
117 exists at the interface of the solid and the liquid [17]. Our system, Ag-Al on KovarTM,
118 undergoes reactive wetting [10, 7], with dissolution of elements and formation of an IMP
119 of the form $(\text{FeNiCo})_x\text{Al}$, where $1 < x < 3$, deposited in a reaction zone of about $10 \mu\text{m}$
120 thickness [7]. These wetting experiments were performed using millimeter-sized drops on
121 a KovarTM substrate.

122 Reactive wetting has been extensively studied but is not well-understood; see reviews in
123 Refs. [18, 19, 17]. During reactive wetting, the change in composition of the substrates can
124 be reaction-limited or transport-limited (i.e., diffusion of the active element – Al in our case
125 – to the liquid-solid interface where the reactions occur). During the spreading of the drop,
126 the solid-vapor-liquid interface (called the contact or triple line in 3D and triple point in 2D

127 descriptions) may initially travel along the bare substrate, but ultimately travels on a layer
128 of the reaction products (RP) until equilibrium is reached. While the initial evolution of the
129 interface shape is dominated by viscosity and inertial effects, the equilibrium configuration
130 is determined by the balance of capillary forces. Young’s equation, which relates the wetting
131 angle (also called contact angle) θ to the surface energy of the substrate, surface tension of
132 the liquid and the solid-liquid interfacial energy, should ideally be cast in terms of the liquid-
133 RP interface rather than the native substrate. Since this energy is usually not known, there
134 have been proposals to modify Young’s equation to include the change in the free energy
135 of the system caused by the production of the RPs [20, 21]. Thus by ignoring the RP
136 when generating our TD, we will incur an error in our data-driven wetting models, but the
137 magnitude is likely to be specific to our exemplar system.

138 Wetting can be improved (indicated by a smaller θ [18]) by the addition of trace elements
139 to the filler alloy or the use of higher brazing temperatures; higher temperatures reduce
140 viscosity and surface tension, and increase diffusion-driven transport, and thus accelerate the
141 path to equilibrium. Ref. [22] modeled this convergence to equilibrium using an exponential
142 in time whereas Ref. [9] observed more complex processes overlaid on the exponential. This
143 auto-correlated behavior in time will play a central role in how we select a θ from MD
144 simulations (discussed in Sec. 3.1). In addition, as the drop reacts with the substrate, it
145 loses height without an equivalent change in its basal area [17, 18] (i.e., it penetrates into the
146 solid as the solid substrate dissolves into the liquid filler material), and for nanometer-sized
147 drops, may disappear completely [8]. This same phenomenon also affects MD simulations
148 that are limited in the number of atoms that can be simulated. Additionally, as drop
149 geometries can be difficult to extract in reactive simulations (see review in Ref. [19]) we use
150 a fixed (i.e. non-reactive) substrate (see discussion below). In doing so, we incur model-form
151 errors in our data-driven model for θ (see Sec. 5.4). Allowing the substrate to be flexible
152 e.g., reactive, enhances wetting and reduces the wetting angle [10].

153 Detailed simulations of the dissolution and reactive wetting of a nanometer-sized drop on
154 a substrate are often performed using phase-field models [8, 9, 23]. These are multicompo-
155 nent (drop/substrate/RP), multiphase (solid/liquid/gas/RP) partial differential equation
156 (PDE) models that use the Navier-Stokes equations to capture fluid dynamics, and the
157 Allen-Cahn equations to track the evolution of the phases. The PDEs are solved using fi-
158 nite elements [8] or finite-differences [9], and due to the fine resolutions (and the consequent

159 computational cost) needed to resolve the reaction zone, are currently only possible for
160 nanometer-sized drops. These simulations reproduce the exponential convergence of θ to
161 equilibrium and the disappearance of the drop into the substrate. These simulations show
162 the filler material penetrating into the substrate, forming a curved interface on which the
163 IMP is deposited as it is formed. The simulations thus verified experimental findings [24]
164 that caution against the naïve use of Young’s equation as there is no fixed, planar substrate
165 to use when defining the wetting angle. Ref. [8] also found that reactions promote wetting
166 and the solid-liquid-vapor triple point moves faster in reactive systems (over the IMP) than
167 non-reactive ones. IMP formation is controlled by the kinetics as well as the transport of
168 the filler material and IMP to and from the filler-substrate interface. Similar simulations for
169 an Al droplet on a Au substrate (a solder; Ref. [9]) provided an example of the proper use of
170 Young’s equation in a reactive system by defining it with respect to the curved IMP surface
171 rather than the original, flat one. The simulation resolved the fluid mechanics inside the
172 droplet as well as the reactions at the filler-substrate interface that lead to the formation
173 of IMP; these kinetics were developed in Ref. [23]. They show the change in wetting angle
174 when the IMP is formed, and the difference is significant (at least for the Al-Au system).
175 The timescale of variation is in nanoseconds, implying that it is the equilibrium conditions
176 that are of relevance in engineering-scale simulations, (see Fig. 5). They also found that the
177 equilibrium wetting angle, at least for nanometer-sized droplets, depends on the size of the
178 droplet. In addition, Ref. [9] shows that the wetting angles computed with spherical drops
179 in 3D are larger than their counterparts computed using 2D simulations. Since we will
180 use quasi-2D simulations (see below) to generate our training data, we may incur another
181 model-form error, (likely an under-prediction of the wetting angle), in our approach (see
182 discussion in Sec. 5.4).

183 *2.2. Molecular dynamics simulations*

184 Previous molecular dynamics simulations of brazing have focused on determining the
185 degree of wetting and spreading of a liquid braze alloy on a solid substrate [25, 26, 27, 10].
186 These simulations have considered both passive wetting [25, 10], where the substrate is held
187 fixed, and reactive wetting where diffusion and reactions between the drop and substrate
188 are allowed [26, 10]. To simplify the development of these data driven models, this work has
189 initially focused on fixed substrates. To explicitly calculate the wetting angle of a drop on
190 a substrate, various approaches can be used. The wetting angle of a drop can be estimated

191 using Young’s equation if the interfacial free energies of the system are known or can be
192 calculated. Specifically, this requires the surface free energy of the solid substrate, surface
193 tension of the drop, and solid-liquid interfacial free energy between the drop and substrate
194 [28, 29].

195 The wetting angle can be computed from atomistic simulations, for instance, by taking
196 the angle between the solid/liquid interface and liquid/vacuum interface at the triple point
197 [10, 29]. The location of the triple point and the surface and drop profiles all fluctuate during
198 dynamic simulations and introduce errors in the direct calculation of the wetting angle. An
199 alternative approach to measure the wetting angle of a drop consists of fitting the contour
200 of a drop to a circular sector (2D) or spherical cap (3D). By knowing the distance between
201 the solid/liquid interface and the origin of the circle (2D) or sphere (3D) the wetting angle
202 can be found [30]. This fitting method is robust to local fluctuations, but incurs fitting
203 errors as we approximate a discrete boundary formed by molecules with a smooth curve or
204 surface.

205 *2.3. Data-driven modeling*

206 Many processes that are not explicitly resolved by continuum simulators are empirically
207 modeled using data-driven models, fit to experimental [31] or other high-fidelity [32] data.
208 The functional forms for the data-driven fits can be motivated by theory or by flexible data
209 approximators such as neural networks [32] or random forests [33]. Polynomials have also
210 been widely used for this purpose in surrogate modeling (Chapter 13, Ref. [34]), especially
211 in cases where the TD might be limited, as they allow their simplification (commensurate
212 with the limited TD) using shrinkage (or an L_1 penalty) [35, 36] or by backward-forward
213 stepwise elimination [37, 38]. Both these methods remove terms in the polynomial that
214 are not significantly correlated with $Z = \{\theta, \sigma\}$. However, in these polynomial models the
215 unknown model coefficients \mathbf{w} and \mathbf{v} are linearly related to Z , and if Z has uncertainties
216 that can be modeled as a Gaussian, there exists an analytical method for propagating them
217 back to the (Gaussian) JPDF for \mathbf{w} and \mathbf{v} (see Ref. [39]), as we will do. The method
218 requires us to quantify our prior belief regarding (\mathbf{w}, \mathbf{v}) as Gaussians also.

219 A challenging feature of our TD-generating model (MD) is that, even when converged,
220 the output consists of a value for $\theta(t_l)$, $l = 1 \dots L$ that strongly fluctuates around a steady-
221 state value. On convergence, it should be possible to subsample (or thin) $\theta(t_l)$, $l = 1 \dots L$
222 to resemble independent draws around a central value with no temporal trend. In contrast,

223 before convergence, $\theta(t_l)$ is auto-correlated (in fact, per the discussion above, it has an
 224 exponential temporal trend) and can be thought of as a p^{th} -order Markov process. The two
 225 can be distinguished by testing whether (windows of) $\theta(t_l)$ resemble a first-order Markov
 226 chain (henceforth, MC) more than independent draws around θ_m . Details are in Ref. [40]
 227 and are summarized in Sec. 3.1, where we adapt this method to determine the convergence
 228 of MD simulations. The method uses a specified quality (i.e., level of certainty) with which
 229 the wetting angle must be extracted from the simulation, to evaluate whether the time-series
 230 $\theta(t_l)$, $l = 1 \dots L$ is sufficiently long to allow such a computation. If the time-series is too
 231 short, the method predicts how much longer the simulation must be run. An alternative
 232 approach is to compute windowed averages of $\theta(t_l)$, $l = 1 \dots L$ (over Δt to smooth over
 233 statistical fluctuations) and check for an approach to a limiting θ_m . The latter approach,
 234 however, requires one to empirically determine Δt , and if convergence has not been reached,
 235 does not provide an estimate of how much longer to run.

236 *2.4. Continuum modeling*

237 The Cauchy momentum and continuity equations form the basis for continuum modeling
 238 of the fluid flow in brazing applications [41]. In the molten alloy (or filler material), the
 239 dynamic flow behavior is dictated by the density and viscosity. In addition, the method used
 240 to apply the boundary conditions is critical to both the dynamic and equilibrium behavior.
 241 The wetting speed at the solid/liquid contact line is dictated by a slip factor that overrides
 242 the traditional no-slip condition. The gas/liquid interface (top surface in Fig. 6) is a free
 243 surface that is allowed to move. At this boundary, the equilibrium behavior of the alloy
 244 is governed by the surface tension and the applied contact angle at the solid/gas/liquid
 245 interface.

246 Modeling free surfaces has been a historically challenging problem in computational
 247 fluid dynamics due to the discontinuity across the interface and the dynamic nature of
 248 moving surfaces. However, several advanced methods such as the volume of fluid [42], phase
 249 field [43], level set [44], and Arbitrary Lagrangian Eulerian (ALE) [45] have been shown to
 250 be accurate and stable in free surface simulation applications ranging from porous flow to
 251 propagating fluid surfaces. The popular level set method works by solving an independent
 252 function ϕ that is advected through an Eulerian grid. The phase interface can be defined
 253 as $\phi = 0$, with the positive and negative regions corresponding to the different phases. One
 254 difficulty with this approach is capturing the sharp interface location, as the mesh cannot be

255 specified with the interface in mind *a priori*. To resolve this, the conformal decomposition
 256 finite element method (CDFEM; [46, 47, 48]), which dynamically adds nodes along the
 257 interface into the background mesh rather than requiring interpolation of quantities between
 258 nodes, has been used in this work. This allows for sharper interface modeling as well as
 259 discontinuous representation of variables across the interface.

260 For wetting applications, such as brazing simulations, the traditional no-slip boundary
 261 condition is no longer applicable as the fluid needs to spread or contract along the solid sur-
 262 face. For finite element simulations, this requires refined modeling of the interface contact
 263 line along the solid surface. Typically, the slip condition is incorporated into the Navier-
 264 Stokes equations using a slip length, an extrapolated distance normal to the wall where the
 265 traditional no-slip condition could be theoretically applied [49]. For finite element simula-
 266 tions, this length can be recast as a dimensionless coefficient that is inversely dependent
 267 on the mesh size [50, 47]. In addition to slip modeling, at the solid/liquid contact line the
 268 surface energies of the respective phases must be accounted for through a contact angle.
 269 This has been achieved previously by incorporating a force at this contact line into the
 270 Navier-Stokes equations to drive the contact line toward the material-dependent wetting
 271 angle [51, 52]. While these methods can be used efficiently to model wetting applications,
 272 it is important to note that they still rely on accurate knowledge of the material properties;
 273 here we propose to incorporate these properties using data-driven models.

274 **3. Formulation**

275 In this section we describe a conservative method to gauge the convergence of a MD
 276 simulation generating wetting angles $\theta(t_l)$ (or the surface tension $\sigma(t_l)$). By conservative, we
 277 mean that it establishes a sufficient, but not necessary condition, i.e., if our technique indi-
 278 cates that the simulation has converged, then, with high probability, it has indeed done so;
 279 however, if the method fails to detect convergence, that does not imply that the MD simula-
 280 tion has not converged. This conservative method is needed to automatically track the gen-
 281 eration of TD, as manual checking is simply not feasible for the multitude of MD simulations
 282 that are required. We also describe the method used to construct data-driven models i.e.,
 283 $\sigma = F(Y_{Al}, T; \mathbf{w})$ and $\theta = G(Y_{Al}, T; \mathbf{v})$. The training data (TD) $(\{Y_{Al}, T, \sigma, \theta\}_i, i = 1 \dots N)$
 284 are generated using converged θ, σ . $F(;;:)$ and $G(;;:)$ are integrated into FEM simulations
 285 of sessile drops to ensure numerical stability and for model verification purposes.

286 *3.1. Diagnosing convergence of a fluctuating MD data series*

287 Consider a series of wetting angles $\theta(t_l) = \{\theta_l\}$, $l = 1 \dots L$ calculated from an MD simu-
 288 lation. The wetting angle θ_l is determined at time t_l by fitting the drop shape to a spherical
 289 cap as described section 4.1. The sequence may indicate a systematic evolution/trend of
 290 θ and thus a lack of convergence to a non-zero θ [22, 9]. The sequence might also show
 291 an oscillation as the MD simulation approaches equilibrium in $\theta(t_l)$ (e.g., Fig. 2b); esti-
 292 mating the final wetting angle as the empirical mean θ_m can be biased depending on the
 293 stopping point of the sequence, the number of periods included in the empirical mean and
 294 the size of the oscillations (which could depict physical effects like capillary waves or a lack
 295 of convergence). Ideally the samples used in the mean should resemble independent draws
 296 from a stationary distribution, which could be obtained by thinning (i.e., subsampling) the
 297 sequence by k i.e., only every k^{th} item of the sequence is retained for estimation purposes.
 298 Note that if θ_l oscillates periodically about a central value, the method would indicate (er-
 299 roneously) a lack of convergence. Note, too, that if there are many periods included in
 300 the median (or mean) and/or the oscillations are small, the bias may be sufficiently small
 301 to be acceptable. This makes the ideal case, of computing θ_m from independent draws, a
 302 conservative method for diagnosing convergence i.e., it is a sufficient, but not necessary,
 303 condition for convergence.

304 We assume that the sequence $\{\theta_l\}$ is a Markov chain (MC), and we wish to estimate a
 305 central value e.g., its median θ'_m , from it. We also desire that θ'_m lies between $\pm r$ of the
 306 true median q with confidence s i.e. $P(q - r \leq \theta'_m \leq q + r) = s$. If L is too short to allow
 307 this estimation, we desire an estimate of a length L' that will allow the estimation. The
 308 derivation below is adapted from our previous work on a generative model for uncorrelated
 309 complex networks [40], and is based on determining convergence of an MC [53, 54].

310 We compute the empirical median θ'_m from $\{\theta_l\}$ and convert it to a binary sequence
 311 $\{v_l\}$, where $v_l = 1$ if $\theta \geq \theta_m$ and $v_l = 0$ otherwise. Let $\{v_l^k\}$ be the k -thinned version of
 312 $\{v_l\}$, and let N_{mn} be the number of (m, n) transitions observed in $\{v_l^k\}$, $m, n \in \{0, 1\}$. N_{mn}
 313 is used to populate a 2×2 contingency table with entries normalized by the sequence length
 314 i.e. $(L/k - 1)$. Let \widehat{N}_{mn} and $\widehat{\pi}_{mn} = \widehat{N}_{mn}/(L/k - 1)$ be the expected values of the table
 315 entries depending on whether we assume the sequence $\{v_l^k\}$ contains independent draws, or
 316 are generated by a first-order Markov chain (and thus can display autocorrelation). The
 317 goodness-of-fit of the data to these models is given by the likelihood-ratio statistic (G^2 ; see

318 Chapter 4.2 in Ref. [55]) and the Bayesian Information Criterion

$$G^2 = -2 \sum_{m=0}^{m=1} \sum_{n=0}^{n=1} N_{mn} \log \left(\frac{\widehat{N}_{mn}}{N_{mn}} \right) \quad \text{or} \quad BIC = G^2 + N_p \log \left(\frac{L}{k} - 1 \right), \quad (1)$$

319 where N_p is the number of parameters in the model used to fit the table data. Log-linear
 320 models are generally used to model table data and estimate the transition probabilities $\pi_{mn}^{(I)}$
 321 and $\pi_{mn}^{(M)}$ (see Ref. [40] for derivation) where superscripts I, M indicate an independent and
 322 a Markov process respectively. Drastically abbreviating the derivation in Ref. [40], we get

$$\Delta BIC = BIC^{(I)} - BIC^{(M)} = -2 \sum_{m=0}^{n=1} \sum_{n=0}^{n=1} N_{mn} \log \left(\frac{\widehat{N}_{mn}^{(I)}}{N_{mn}} \right) - \log \left(\frac{L}{k} - 1 \right). \quad (2)$$

323 Here $\widehat{N}_{mn}^{(I)}$ is the expected number of (m, n) transitions if the binary sequence $\{v_l^k\}$ resembled
 324 independent draws from a distribution. A negative ΔBIC implies that the independent
 325 draws model fits the data better.

326 The log-linear models require that the sequence, of length $(L/k-1)$, is sufficiently long to
 327 provide good estimates and here we provide an estimate of how long is considered sufficient.
 328 Let θ_m^L be the empirical mean computed from the full sequence and θ_m^k from a k -thinned
 329 chain. Multiple θ_m^k can be computed for various levels of thinning (i.e., various k s) and will
 330 constitute draws from a normal distribution with mean q and variance ν^2 . To ensure that
 331 the k -thinned sequence is sufficiently long, i.e., the empirical mean θ_m^k lies near the true
 332 mean with sufficient confidence or $P(q - r \leq \theta_m^k \leq q + r) = s$, or

$$\left(\frac{r}{\Phi^{-1}\{0.5(1+s)\}} \right)^2 = \nu^2 \quad (3)$$

333 where Φ is the cumulative distribution function for a standard normal distribution and s, r
 334 have been defined above.

335 Consider an L -step $\{\theta_l\}$ sequence that has been thinned by a factor k' and has become
 336 a first-order MC. Let L' be its length. The contingency table entries provide the transition
 337 probabilities of the 2-state Markov sequence $\{v_l^k\}$. The transition probabilities α and β can
 338 be computed trivially and ν^2 can be written as

$$\nu^2 = \frac{\alpha\beta(2-\alpha-\beta)}{L'(\alpha+\beta)^3} \quad \text{or} \quad L' = \frac{\frac{\alpha\beta(2-\alpha-\beta)}{(\alpha+\beta)^3}}{\left(\frac{r}{\Phi^{-1}\{0.5(1+s)\}} \right)^2}, \quad (4)$$

339 using Eq. 3. Thus the k -thinned sequence $\{\theta_l^k\}$ must be at least L' long to provide an
 340 estimate of q with the specified tolerance r . An implementation of this method is in the R
 341 package `mccgibbist` [56].

342 *3.2. Constructing data-driven models*

343 In this section, we describe the method by which we construct polynomial models for
 344 θ and σ . In our case, the filler material is primarily silver (Ag) with some aluminum (Al)
 345 and the composition Y is defined in terms of mass fraction i.e. $Y = \{Y_{Ag} = 1 - Y_{Al}, Y_{Al}\}$.
 346 Thus the independent variables in the models for θ and σ are (T, Y_{Al}) . The convergence
 347 diagnostic described in Sec. 3.1 results in a k -thinned sequence $\{\theta_i^k\}$, a subset of which
 348 is used to compute the empirical mean and standard deviations (θ_m, ς) . The subset of
 349 the k -thinned sequence that we use is the usually the last half, as by doing so we ignore
 350 the initial transients in $\theta(t)$ and $\sigma(t)$. This is repeated for many $(T_i, Y_{Al,i}), i = \dots N$
 351 combinations (the features in our TD) to yield (θ_i, ς_i) responses (the labels in our TD).
 352 Consider the vector $\mathbf{T} = \{T_i\}$. We scale the variables $\bar{T} = (T_i - \mathbb{E}(\mathbf{T}))/\sqrt{\text{Var}(\mathbf{T})}$, where
 353 $\mathbb{E}()$ and $\text{Var}()$ denote the empirical mean and variance of TD values. The other variables
 354 are similarly scaled to get $\overline{Y_{Al,i}}$ and $\bar{\theta}_i$. We define the polynomial model simply as

$$\bar{\theta}_i = w_0 + w_1\bar{T}_i + w_2\overline{Y_{Al,i}} + w_3\overline{Y_{Al,i}}^2 + w_4\bar{T}_i^2 + w_5\overline{Y_{Al,i}} \times \bar{T}_i + \epsilon_i, \quad (5)$$

355 where ϵ_i is the fitting error, modeled as a Gaussian random variable with a zero mean and an
 356 unknown variance. The quadratic form is motivated by previous work [57] (which modeled
 357 the effect of temperature and surface roughness) and the tendency of trace elements to
 358 reduce the wetting angle [18]. Eq. 5 is fitted to the data via least-squares minimization,
 359 and simplified, via backward-forward stepwise elimination, to remove terms in the equation
 360 which do not contribute to $\bar{\theta}_i$ predictions. This results in a model

$$\bar{\theta}^{(pred)} = \mathbf{A}\mathbf{w}, \quad \mathbf{w} = \{w_j\}, \quad (6)$$

361 where j spans the terms in Eq. 5 that were retained after simplification and the superscript
 362 $(pred)$ denotes a model prediction. The columns of \mathbf{A} contain the linear and quadratic
 363 terms of \bar{T}_i and $\overline{Y_{Al,i}}$.

364 The uncertainty in the wetting angle estimate (ς_i) implies that there is a corresponding
 365 uncertainty in the estimates of \mathbf{w} in Eq. 6. Since the model is linear in w_j , the uncertainty
 366 can be computed analytically (see Chapter 2, Ref. [39]). We create a prior model $\mathbf{w} \sim$
 367 $\mathcal{N}(\mathbf{w}_a, \Gamma_a)$, where Γ_a is diagonal and non-informative (i.e., the diagonal entries are large).
 368 The posterior distribution for \mathbf{w} is given by $\mathbf{w} \sim \mathcal{N}(\hat{\mathbf{w}}, \hat{\Gamma})$ where

$$\begin{aligned} \hat{\mathbf{w}} &= \mathbf{w}_a + \Gamma_a \mathbf{A}^T (\mathbf{A} \Gamma_a \mathbf{A}^T + \Gamma_\theta)^{-1} (\bar{\theta} - \mathbf{A} \mathbf{w}) \quad \text{and} \\ \hat{\Gamma} &= \Gamma_a \mathbf{A}^T (\mathbf{A} \Gamma_a \mathbf{A}^T + \Gamma_\theta)^{-1} \mathbf{A} \Gamma_a. \end{aligned} \quad (7)$$

369 Here $\Gamma_\theta = \text{diag}(c_i/\sqrt{\text{Var}(\boldsymbol{\theta})})$, $\boldsymbol{\theta} = \{\theta_i\}$ is a diagonal matrix denoting the uncertainty in
 370 $\bar{\boldsymbol{\theta}} = \{\bar{\theta}_i\}$. An identical process is followed to develop a model for the surface tension
 371 $\boldsymbol{\sigma}^{(pred)} = \mathbf{B}\mathbf{v}$, where $\mathbf{v} \sim \mathcal{N}(\hat{\mathbf{v}}, \hat{\boldsymbol{\Sigma}})$.

372 3.3. The finite element model for sessile drops

373 To model the alloy deformation at engineering length scales, we use finite element mod-
 374 eling (FEM) as implemented in Sandia's Sierra/Aria Galerkin FEM code [58]. The three-
 375 dimensional domain ($8 \times 8 \times 8$) mm includes two blocks corresponding to a solid substrate on
 376 which the alloy is allowed to spread and a multiphase alloy/atmospheric block resting above
 377 it. Only the boundary of the solid block interacts with the simulation, and all equations
 378 are solved on the fluid blocks. The maximum time step is set to 0.02 ms.

379 The alloy/atmospheric interface is represented using the CDFEM method [46, 47, 48].
 380 With CDFEM, the interface is originally represented on the non-decomposed alloy/atmospheric
 381 block mesh using a level set field, (ϕ) . This block is then conformally decomposed into sep-
 382 arate alloy and atmosphere regions along the $\phi = 0$ isosurface.

383 Within the braze alloy and the atmosphere above it, the Cauchy momentum and conti-
 384 nuity equations:

$$\rho \frac{\partial \bar{v}}{\partial t} + \rho \bar{v} \cdot \nabla \bar{v} - \bar{g} - \nabla \cdot \bar{T} = 0 \quad (8)$$

$$\nabla \cdot \bar{v} = 0, \quad (9)$$

386 respectively, are solved to evaluate the pressure P and velocity vector \bar{v} . Here, ρ represents
 387 the density of the specific material phase (10,490 kg/m³ for the alloy and 1.225 kg/m³ for
 388 the atmosphere), \bar{v} is the gravity vector (pointed towards the substrate), and \bar{T} is the stress
 389 tensor, which is defined as:

$$\bar{T} = -P\bar{I} + \mu \left((\nabla \bar{v})^\top + \nabla \bar{v} \right), \quad (10)$$

390 where μ is the viscosity (2.91 cP for the alloy and 0.0181 cP for the atmosphere) and \bar{I}
 391 is the identity tensor. At all outermost surfaces of the domain, an open flow boundary
 392 condition is applied with $P = 0$ and $\bar{n} \cdot \nabla \bar{T} = 0$.

393 At the interface, surface tension is incorporated through a flux boundary condition:

$$\bar{n} \cdot \left(\bar{T}_l - \bar{T}_g \right) = -\sigma \bar{n} \nabla \cdot \bar{n}, \quad (11)$$

394 where \bar{n} is the unit normal vector pointing out of the alloy and the subscripts l and g
 395 correspond to the alloy and atmospheric phases, respectively. Here, we incorporate the

396 species and temperature dependent surface tension $\sigma = G(T, Y_{Al}; \mathbf{v})$, as established in
 397 the preceding sections. At the substrate interface, fluid motion is enabled through a slip
 398 boundary condition applied tangentially along the substrate:

$$\bar{t}_w \cdot \bar{T} \cdot \bar{n}_w = -\frac{\mu\beta}{\Delta x} \bar{v} \cdot \bar{t}_w, \quad (12)$$

399 where \bar{t}_w is the unit tangent vector at the substrate surface, \bar{n}_w is the unit normal vector at
 400 the substrate surface, Δx is the mesh size (0.15 mm background with 0.01 mm surface edge
 401 tolerance), and β is a non-dimensional slip factor (set as 0.01 for the alloy and 1.0 for the
 402 atmosphere). The influence of the slip parameter is analyzed in Appendix A. To prevent
 403 the drop from sliding around the substrate (due to numerical errors or asymmetrical mesh
 404 elements), a no-slip condition is applied to a small region of the alloy/substrate interface
 405 at the very center of the domain. The species and temperature dependent contact angle
 406 $\theta = F(T, Y_{Al}; \mathbf{w})$ is applied to the contact line (the intersection of the alloy/atmosphere
 407 interface and the substrate surface) through a force \bar{f}_θ term:

$$\bar{f}_\theta = \sigma (\bar{t}_w \cos \theta + \bar{n}_w \sin \theta). \quad (13)$$

408 The mean values ($\widehat{\mathbf{w}}, \widehat{\mathbf{v}}$) (see Eq. 7) are used to compute the nominal values of (θ, σ) used
 409 in the FEM.

410 We initialize the level set field as a hemispherical surface of radius 1 mm above the
 411 substrate. The level set is the advected through the non-decomposed mesh according to:

$$\frac{\partial \phi}{\partial t} + \bar{v} \cdot \nabla \phi = 0, \quad (14)$$

412 where \bar{v} is the fluid velocity from the momentum equation. CDFEM is then used to update
 413 the conformal decomposition at every time step.

414 Thermal transport is coupled to the momentum equations through a separate convection-
 415 diffusion equation for the temperature T , using Fourier's Law for the diffusive flux:

$$\rho c_p \frac{\partial T}{\partial t} + \rho c_p \bar{v} \cdot \nabla T = \nabla \cdot (\kappa \nabla T). \quad (15)$$

416 Here, c_p is the specific heat capacity (arbitrarily set at 100 J/kg-K), and κ is the thermal
 417 conductivity (arbitrarily set at 100 W/m-K). The above equation was solved in both the
 418 liquid and gas phases, and a source term is applied as a flux boundary condition (1.5
 419 W/mm²) on the interface of the liquid with the solid wall. No flux is assumed for all
 420 other surfaces. Similarly, another convection-diffusion equation was coupled to the system

421 to represent chemical species conservation for aluminum, where we assumed the chemical
 422 species obeys Fick’s Law for diffusion:

$$\frac{\partial C_{Al}}{\partial t} + \bar{v} \cdot \nabla C_{Al} = \nabla \cdot (D_{Al} \nabla C_{Al}). \quad (16)$$

423 Here, C_{Al} represents the concentration of aluminum in the liquid, and D_{Al} is the aluminum
 424 diffusivity (arbitrarily set at $50 \text{ mm}^2/\text{s}$). Aluminum transport was only solved in the liquid
 425 domain with a flux condition (arbitrarily set at $0.3 \text{ mol}/\text{mm}^2\text{-s}$) at the solid wall interface
 426 and a no flux condition applied to the gas/liquid surface. Since incorporation of thermal
 427 and species transport presented here is intended as an illustrative example, the physical
 428 properties and fluxes were defined as to allow temperatures and concentrations to change
 429 significantly in the relatively short simulation time.

430 4. Data generation

431 In this section, we describe the method by which TD of MD simulations were generated.
 432 We also describe the experimental method by which we acquired the measurements of θ to
 433 estimate the model-form error in our MD simulations and therefore $F(T, Y_{Al}; \mathbf{w})$.

434 4.1. Generating the training data

435 To generate training data of θ and σ , as a function of (T, Y_{Al}) , we first generate samples
 436 in $1150 \text{ K} \leq T \leq 1350 \text{ K}$ and $0 \leq Y_{Al} \leq 0.043$ mass fraction (corresponding to a maximum
 437 mole fraction of 0.15), based on the temperatures and compositions observed in experimental
 438 studies of brazed joints [59]. The sampling was performed using a space-filling Halton
 439 sequence. Some of the (T, Y_{Al}) combinations were such that the filler material was not a
 440 liquid, and consequently we discarded these samples. Since the sampling design was space-
 441 filling, the excision of a subset of samples did not materially degrade the sample set. Fig. 1a
 442 shows the (T, Y_{Al}) samples that were retained, along with the (experimental) liquidus line
 443 that denotes the boundary above which the filler material is liquid. These 45 samples were
 444 used in MD simulations to compute the corresponding θ .

445 MD simulations were run using the embedded atom method [12] potentials developed
 446 by Zhou et al. to approximate an Ag-Al braze alloy and Kovar substrate [60]. Quasi-2D
 447 cylindrical droplets were considered. Note that quasi-2D simulations will slightly over-
 448 predict the wetting angle and surface tension relative to 3D simulations; Ref. [9] shows a

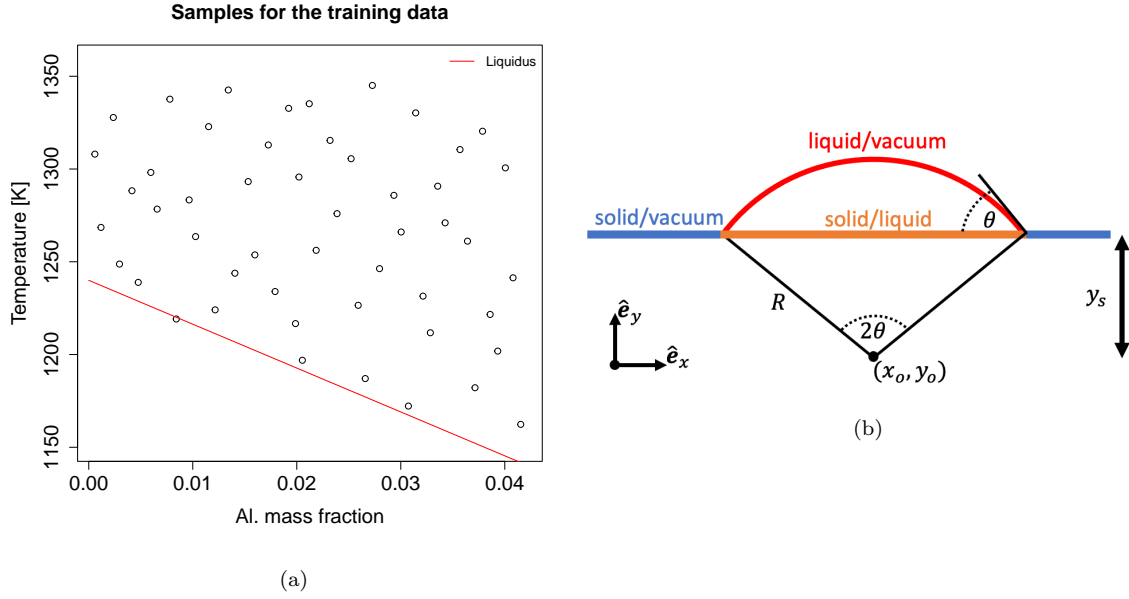


Figure 1: (a) Samples in the (T, Y_{Al}) -space which were used to generate the TD for (θ, σ) . The red line denotes an approximation of the experimental liquidus boundary; above it, the (T, Y_{Al}) combinations yield a liquid filler. (b) Relationship between the wetting angle of a droplet, θ , and the circle associated with the wetting angle defined by the radius of the circle, R , its origin, (x_0, y_0) , and the distance of the origin of the circle to the surface y_s .

449 5% discrepancy in wetting angle between the two, computed using a phase-field model, for
 450 a Al droplet on a Au substrate. The droplets were placed on a (010) face-centered cubic
 451 substrate with dimensions of $78a_0 \times 13a_0 \times 8a_0$ corresponding to the droplet spreading
 452 direction, interface plane, and droplet width. The substrate was given a lattice parameter
 453 of $a_0 = 3.6 \text{ \AA}$. The initial wetting angle of the droplet was set to 90° . A Nose-Hoover
 454 thermostat (NVT) was used with a 100 fs temperature damping parameter. Each simulation
 455 ran for 50 ns with the timestep set to 2 fs and atomic positions were output every 20 ps. The
 456 wetting angle at a given time $\theta(t_l)$ was measured using the approach outlined by Nijmeijer
 457 et al. [30]. The functional form of the droplet's liquid/vacuum interface is that of a circular
 458 sector. As such, by fitting the liquid/vacuum interface to that of a circle, the wetting angle
 459 can be shown to be

$$\theta(t_l) = \arccos\left(\frac{y_s(t_l)}{R(t_l)}\right), \quad (17)$$

460 with y_s and R defined in Fig. 1b. To determine the contour of the liquid/vacuum surface,
 461 bins of width 3 \AA were created along the \hat{e}_x direction. The position of the Ag atom with
 462 the highest y-component in a given bin was defined to be the height of the liquid/vacuum

463 interface for that bin. The solid/vacuum surface was defined as being 5 Å above the lowest
 464 value of all bins considered in the simulation to exclude surface states (i.e. the precursor
 465 foot that spreads before the main drop) along the solid/vacuum interface in the fitting of a
 466 circle to the surface.

467 The surface tension of the Ag-Al liquids was calculated from MD by creating a $10a_0 \times$
 468 $10a_0 \times 20a_0$, $a_0 = 4.22$ Å, block of FCC atoms with periodic boundary conditions in the
 469 x- and y-directions and free surfaces for the z-direction. The block of atoms was made into
 470 a liquid by randomly displacing each atom -0.4 to 0.4 Å in each of the three Cartesian
 471 directions, and then running the simulation at 2500 K for 2 ps in an NVT ensemble with a
 472 Langevin thermostat. After the liquid was formed, the simulation was run for 4 ns at the
 473 desired temperature. The surface tension was calculated from the equation [61]

$$\sigma = \frac{V}{4A} (\langle \sigma_{xx} \rangle + \langle \sigma_{yy} \rangle - 2p), \quad (18)$$

474 where V is the volume of the system, $\langle \sigma_{xx} \rangle$ and $\langle \sigma_{yy} \rangle$ are the time-averages of the xx and
 475 yy components of the system’s stress tensor, averaged over the final 2 ns of the simulation.
 476 A is the surface area between the liquid and vacuum. For mechanical equilibrium to be
 477 achieved for the planar interface between two fluids, the pressure in each fluid must equal p .
 478 In the case considered here, the interface between a vacuum and a liquid, We take $p = 0$. We
 479 note that within molecular dynamics simulations, averaged small, but finite, stresses normal
 480 to the vacuum/liquid interface can occur [62]. Substituting $p = \langle \sigma_{zz} \rangle$ into Eq. 18 changes
 481 the results of our surface tension calculations on average by 4 mJ/m². These finite values
 482 are within the average error of the calculation of $\langle \sigma_{xx} \rangle$ and $\langle \sigma_{yy} \rangle$ across all compositions
 483 and temperatures considered, 80 mJ/m².

484 *4.2. Experimental methodology*

485 Silver sessile drops on Kovar were formed by the following processes. Kovar sheet metal
 486 was sheared into 25×25 mm squares, degreased with acetone and isopropyl alcohol, and
 487 fired in a Thermal Technologies Inc. Astro AVF 430-SPL vacuum-hydrogen furnace for 2
 488 hours at 1000°C in a dry hydrogen atmosphere to reduce oxide on the surface. Brazing filler
 489 metal discs were cut from a pure silver rod. The discs weighed $0.42\text{g} \pm 0.06$. The discs were
 490 degreased in acetone and isopropyl alcohol. Sessile drops were formed by placing a silver
 491 disc on a Kovar square, placing a cylindrical Ti sheet metal “top-hat” on top of the sample
 492 to act as a getter and increase cleanliness, and brazed with the following parameters. The

493 furnace temperature was ramped up at a rate of $10^\circ\text{C}/\text{min}$, with holds at 500°C (30 min.),
494 900°C (10 min.), and the brazing temperature (10 min.). Brazing temperatures of 1243,
495 1263, 1273, 1283, 1293, 1303 and 1323 K (970, 990, 1000, 1010, 1020, 1030, and 1050°C)
496 were investigated. The initial heating was conducted under high-vacuum, approximately
497 10^{-6} Torr, generated by a cryo-pump which allowed the molybdenum oxide on the heating
498 elements to decompose and the O_2 be removed from the chamber. After 15 min. under
499 high-vacuum at 500°C , 9 Torr of gettered Ar (house liquid Ar source passed through a Ti
500 gettering furnace, yielding Ar with approximately 0.3 PPB O_2) was added to the chamber
501 to suppress Ag evaporation at brazing temperatures. With the Ar in place, temperature
502 was ramped to 900°C where a 10 min. hold increased temperature uniformity across the
503 chamber, and then to the brazing temperature where it was held for 10 minutes. The
504 furnace was then allowed to cool passively to room temperature.

505 The contact angle of the sessile drops was determined as follows. A “Depth-up” or “Z-
506 stack” series of images focused at different heights was taken at $500\times$ optical magnification
507 on a Keyence VHX-6000 digital microscope and stitched into a 3D surface. Height profile
508 data was then plotted along a line versus position and the flat plate and slope of the sessile
509 drop were fitted with lines by eye using the Keyence software. The angle between these lines
510 was taken as the contact angle. This process was repeated three times every 90° around
511 the circumference of the drop, for a total of 12 contact angle measurements. The average
512 of these 12 measurements is the reported contact angle.

513 **5. Results**

514 In this section we develop models for θ and σ and embed them in an FEM to simulate
515 the spreading of sessile drops.

516 *5.1. Determining convergence and extracting θ and σ*

517 In Fig. 2a, we plot the sequence $\{\theta_i\}$ from an MD simulation, of $L = 2501$ time-points
518 where θ was computed (henceforth, called “ticks”, and equal to a duration of 20 ps). The
519 figure also shows a window-averaged version (with a window width of 200 ticks) that shows
520 the convergence trend and the low-amplitude oscillations. We apply the convergence di-
521 agnostic of Sec. 3.1 with the quality requirement that the true median θ of the trace be
522 computed, with 95% confidence, between the 40th and 60th percentile of the θ data. While

Run	L	L'	Correlation length	(θ, ς)
Well-behaved	2501	1709	17.6	$(22.81^\circ, 0.91^\circ)$
Ill-behaved	2502	4802	49.7	$(26.54^\circ, 1.70^\circ)$

Table 1: Statistics of the two runs shown in Fig. 2 (top). The “well-behaved” run attains its quality requirements within the number of timesteps simulated, primarily because of its smaller correlation length. The “ill-behaved” (unconverged) run does not, and results in a larger ς .

523 these lower and upper bounds may seem excessively generous, the actual variation of θ_l is
524 small. The diagnostic yields that for the specified quality requirement, a subset of the run
525 $L' = 1709$ is sufficient; further, it computes a correlation timescale of 17.6 ticks. Thinning
526 the $\{\theta_l\}$ by 18 (the nearest integer greater than the correlation length), we get the samples
527 that are plotted in red. We compute the mean and standard deviation of the last half of
528 the red symbols to obtain (θ, ς) for this MD simulation. For the purposes of the data-driven
529 model, θ can be thought of as an observation/measurement and ς , its measurement error.
530 Per Sec. 3.2, only the last half of the samples, right of the vertical line in Fig. 2a, are used
531 to compute (θ, ς) . This information is summarized in Table 1. In Fig. 2b, we show an
532 non-converged run (per the quality metrics specified to our conservative diagnostics). The
533 time-averaged trace shows clear oscillations and the ideal run-length (L') indicates that
534 the simulation ought to be run twice as long and thinned more aggressively to yield $\{\theta_l\}$
535 samples that are independent (see Table 1). Due to the computational cost, this was not
536 done and the last half of the samples in Fig. 2b are used to compute (θ, ς) . The mean
537 and standard deviation computed for this run leads to ς that is almost twice as large as the
538 previous one. These ς are explicitly captured in our data-driven model and its predictions.
539 In Fig. 2c we plot the histogram of all θ in the TD; the median value is plotted with a
540 solid vertical line and the first and third quartiles with dashed lines. We see that about
541 50% of the wetting angles lie between 21° and 27° . In Fig. 2d we see that the coefficient of
542 variation ς/θ is small, with a median of about 0.045 i.e., despite the occasional inability of
543 our MD simulations to meet the conservative quality requirements, the uncertainty in the
544 simulated θ is rather small. The ill-behaved run is at the extreme right of the figure. Note
545 that the net effect of “ill-behavior” is a larger uncertainty ς which is incorporated into the
546 estimates of \mathbf{w} via Γ .

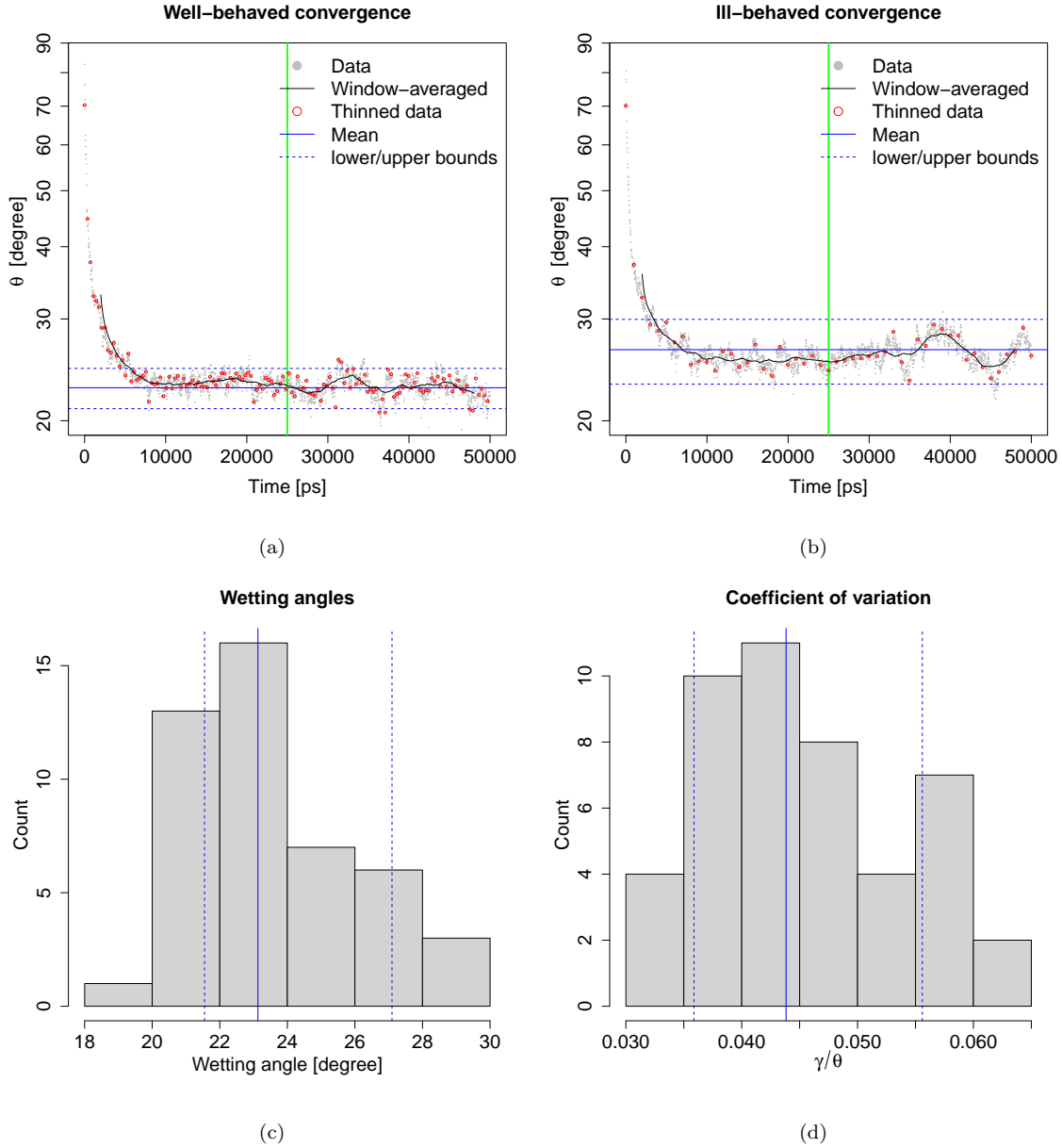


Figure 2: (a) A converged time series of θ for a run with ($T = 1250$ K, $Y_{Al} = 0.042$). The dots are the θ computed at the timesteps that were saved (the “ticks”), the black solid line is the window-averaged version of the same and the red symbols are the θ that constitute independent draws. The final/converged values of the wetting angle and twice its uncertainty i.e., $(\theta, 2\varsigma)$ are also shown in blue. The last half of the samples, right of the vertical line, are used to compute (θ, ς) . (b) The same, but for an “ill-behaved” (or unconverged) simulation time series (corresponding to a run with $T = 1337$ K, $Y_{Al} = 0.0078$). The net effect of using this “ill-behaved” MD simulation is a larger ς when constructing data-driven models. (c) The distribution of the 45 θ in the TD. (d) The distribution of the coefficient of variation ς/θ . In both the figures, the median is plotted with a solid vertical line, and the first and third quartiles are plotted using dashed lines.

547 *5.2. Data-driven models*

548 The TD generated in Sec. 4.1 have features $(T_i, Y_{Al,i})$ and labels $(\theta_i, \varsigma_i), i = 1 \dots N$.
 549 The first step in making a model for $\bar{\theta}$, per Eq. 6, is to determine the prior model i.e., the
 550 structure of \mathbf{A} . We propose a model of the form of Eq. 5 and fit it to the scaled TD i.e.,
 551 to $(\overline{T}_i, \overline{Y}_{Al,i}, \overline{\theta}_i)$, while ignoring ς_i . The centering values $\mathbb{E}(\cdot)$ for (T, Y_{Al}, θ) are $(1.268 \times$
 552 $10^3, 2.222 \times 10^{-2}, 23.71^\circ)$ and the scaling ones $(\sqrt{\text{Var}(\cdot)})$ are $(47.35, 1.269 \times 10^{-2}, 2.280^\circ)$.
 553 The model is simplified via backward-forward stepwise elimination, which removes the w_4
 554 term. Thus the scaled wetting angle $\bar{\theta}^{(pred)}$ is linear in (scaled) temperature, while being
 555 quadratic in \overline{Y}_{Al} . It yields \mathbf{w}_a , that is used in the prior distribution of \mathbf{w} (see Eq. 7). The
 556 adjusted R^2 of the fit is 0.946 and a 20-way cross-validation resulted in a prediction error
 557 of 0.29° . Note that per Fig. 2c, the median θ is about 23° and the “measurement error”
 558 ς (Fig. 2d) is about 5%. Thus the prior model has an prediction error of about 1.2%,
 559 about a quarter of the measurement error. This is an estimate of one component of the
 560 model-form error in our data-driven model, and is due to the limited nature of the TD. The
 561 other component is due to the limitations of the MD simulations and will be quantified in
 562 Sec. 5.4.

563 Having determined the structure of \mathbf{A} we proceed to compute the coefficients \mathbf{w} via
 564 Eq. 7. The prior uncertainty on \mathbf{w} is modeled as $\Gamma_a = \text{diag}((3 * \mathbf{w}_a)^2)$ (i.e., we assume that
 565 the coefficient of variation is 3). Using Eq. 7, we compute

$$\begin{aligned} \widehat{\mathbf{w}} &= \{-0.57, -0.13, -0.75, 0.56, -0.17\} \\ \widehat{\Gamma} &= \begin{bmatrix} 10.1 & -1.2 & -0.6 & -5.8 & 5.4 \\ -1.2 & 6.0 & 1.3 & 0.8 & -2.0 \\ -0.6 & 1.3 & 6.1 & 0.4 & 0.9 \\ -5.8 & 0.8 & 0.4 & 7.0 & 1.5 \\ 0.54 & -2.0 & 0.9 & 1.5 & 7.9 \end{bmatrix} \times 10^{-3}, \end{aligned} \quad (19)$$

566 where $\widehat{\mathbf{w}}$ contains the intercept w_0 and coefficients of the $\overline{T}_i, \overline{Y}_{Al,i}, \overline{Y}_{Al,i}^2$ and $\overline{Y}_{Al,i} \times \overline{T}_i$
 567 terms. $\widehat{\Gamma}$ contains the uncertainty in the \mathbf{w} estimate. While it is diagonally dominant,
 568 it does display significant off-diagonal terms, denoting correlations between the various
 569 elements of \mathbf{w} .

570 In Fig. 3a we plot the $\hat{\theta}^{(pred)}$. It is computed using $\widehat{\mathbf{w}}$ from Eq. 7, Eq. 6 and Eq. 19.
 571 All variables have been restored to their unscaled, physical values. The linear dependence
 572 on T and the quadratic dependence on Y_{Al} are clearly evident. Also evident is the muted

573 sensitivity of θ to T compared to Y_{Al} (see Eq. 5 and 19). We also plot the data points
 574 whose fitting errors are in the top 10th percentile using symbols; two such points are below
 575 the shaded surface and not visible. These “badly modeled” points are the same ones that
 576 had the largest uncertainties in θ in the TD, as plotted in Fig. 2d. Note that the two
 577 axes are plotted in reverse order to illustrate the surface. The same figure is plotted as
 578 contours in the (T, Y_{Al}) -plane in Fig. 3c; the liquidus is clearly visible and demarcates the
 579 (T, Y_{Al}) -space where the filler material is not a liquid. In Fig. 3b, we plot the uncertainty
 580 in $\theta^{(pred)}$ due to uncertain \mathbf{w} . We do so by taking 10,000 samples from $\mathbf{w} \sim \mathcal{N}(\hat{\mathbf{w}}, \hat{\Gamma})$ (and
 581 Eq. 19), generating realizations of $\theta^{(pred)}$ using Eq. 6, and computing the standard deviation
 582 as a function of (T, Y_{Al}) . We see that the uncertainty in θ is largest at the boundaries; this
 583 is expected as the boundaries have data (to constrain the data-driven model) only on one
 584 side. In addition, the prediction uncertainty in the vicinity of the liquidus is also large.
 585 We also plot the ς of the same poorly fit points in Fig. 3a. Finally, in Fig. 3d, we plot a
 586 posterior predictive test. We use 1,000 samples of $\mathbf{w} \sim \mathcal{N}(\hat{\mathbf{w}}, \hat{\Gamma})$ and generate realizations
 587 $\theta^{(pred)}$ using Eq. 5 for $(T_i, Y_{Al,i}), i = 1 \dots 4$ examples that were held back from the TD.
 588 The box-and-whisker plots summarize the predictions. The values of θ from the TD are
 589 plotted as red symbols, with the 2ς bounds as whiskers. We see that in three out of four
 590 cases, the θ from the TD is contained within the first and third quartiles of the predictions,
 591 and the measurements, including the error bars, are contained within the whiskers of the
 592 predictions. However, the effect of the model-form errors (due to the limited TD) are also
 593 evident here - had these errors been negligible, the MD simulations’ θ would have coincided
 594 with the median of the predictions.

595 An identical process was followed to construct the model $\sigma = G(T, Y_{Al}; \mathbf{v})$ for the surface
 596 tension. The behavior of σ in (T, Y_{Al}) -space is simpler than θ and a less sophisticated
 597 sampling was used. The centering values $\mathbb{E}(\cdot)$ for (T, Y_{Al}, σ) are $(1500, 4.5 \times 10^{-2}, 717.7)$
 598 and the scaling ones $(\sqrt{\text{Var}(\cdot)})$ are $(2.02 \times 10^2, 3.26 \times 10^{-2}, 31.35)$. The units of σ are
 599 mN/m. A prior model was fitted to the TD, which revealed that $v_3 = v_4 = 0$ i.e, the only
 600 relevant quadratic term in the model was $\overline{Y_{Al}} \times \overline{T}$. The adjusted R^2 of the fit is 0.98 and a
 601 20-way cross-validation resulted in a prediction error of 4.75 mN/m. Note that the median
 602 σ in the training data was 717.7 mN/m, making the prediction error rather small. Having

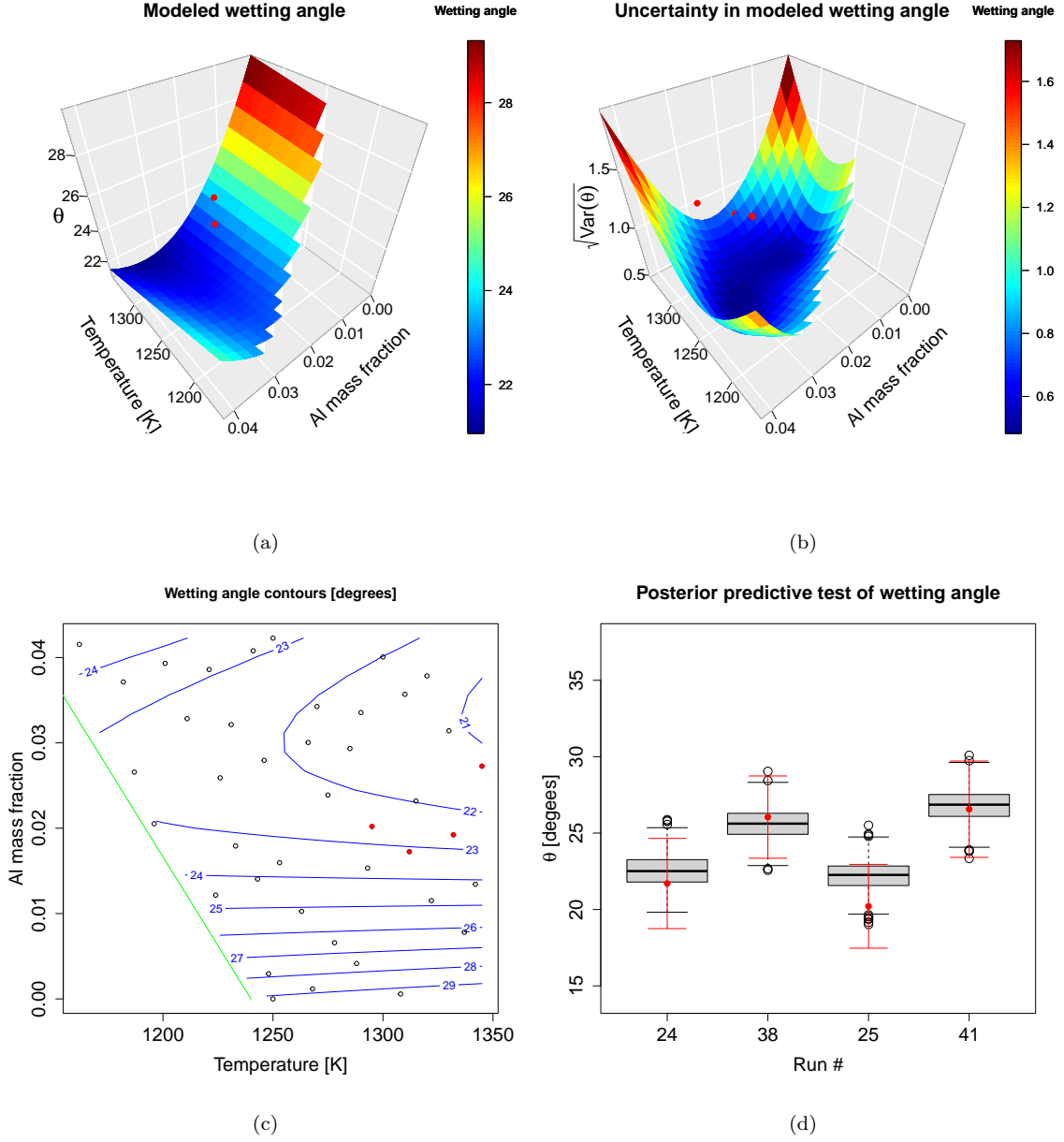


Figure 3: (a) Surface plot of $\hat{\theta}^{(pred)}$. TD points that could not be modeled well are plotted in red. The unit of θ is degrees. (b) The uncertainty (standard deviation) in $\theta^{(pred)}$. (c) Contour plot of mean $\hat{\theta}^{(pred)}$ in the (T, Y_{Al}) -plane. (d) Posterior predictive test of four held-back θ , showing uncertainties in θ predictions versus the “measurement uncertainty”. The red whiskers are 2σ bounds. The run numbers on the horizontal axis are insignificant and only denote the indices of the held-out data in the TD.

603 determined the structure of \mathbf{B} , we computed the probabilistic model per Eq. 7 to give

$$\hat{\mathbf{v}} = \{-0.0019, -0.8863, 0.4418, 0.0432\}$$

$$\hat{\Sigma} = \begin{bmatrix} 0.319 & -0.0140 & 0.0246 & -0.00135 \\ -0.0140 & 220.9 & -1.83 & 2.30 \\ 0.0246 & -1.83 & 228.0 & -1.07 \\ -0.00135 & 2.30 & -1.07 & 95.1 \end{bmatrix} \times 10^{-4} \quad (20)$$

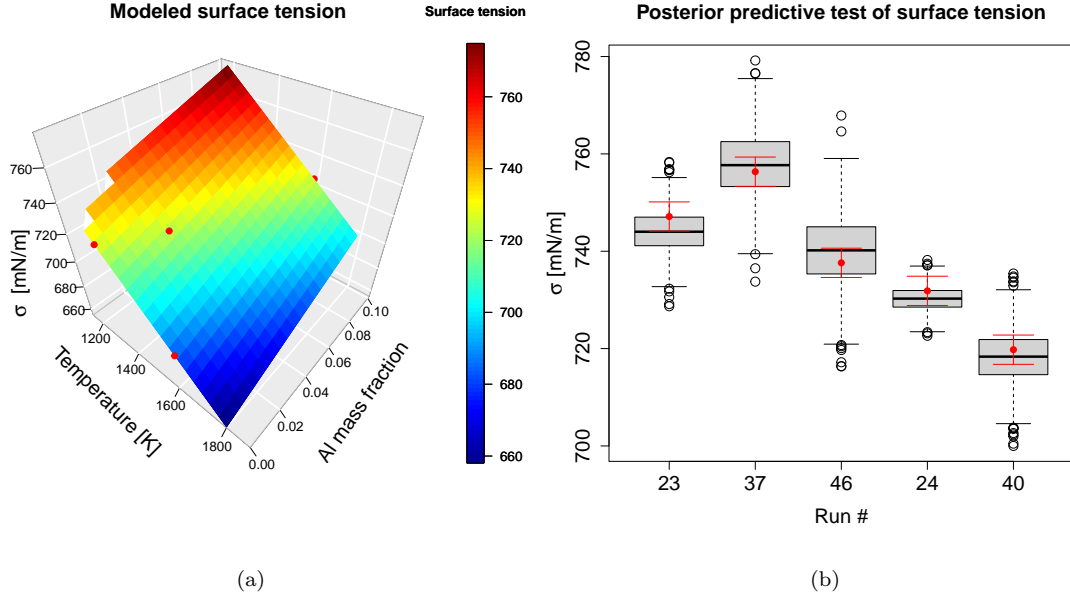


Figure 4: (a) Surface plot of $\hat{\sigma}^{(pred)}$ as a function of (T, Y_{Al}) . The red dots are the worst modeled points. (b) Posterior predictive test of held-back σ , showing uncertainties in $\sigma^{(pred)}$ predictions versus the “measurement uncertainty” ς . The red whiskers are twice-standard-deviation bounds.

604 where $\hat{\mathbf{v}}$ contains the intercept v_0 and coefficients of the \overline{T}_i , $\overline{Y_{Al,i}}$ and $\overline{Y_{Al,i}} \times \overline{T}_i$ terms.
605 $\hat{\Sigma}$ contains the uncertainty in the \mathbf{v} estimate. \mathbf{v} is seen, from $\hat{\Sigma}$, to have insignificant
606 correlations between themselves (off-diagonal terms are two orders of magnitude smaller
607 than the diagonal ones). In Fig. 4a we plot the most probable prediction of $\sigma^{(pred)}$ i.e.,
608 $\hat{\sigma}^{(pred)}$, which shows a strong linear trend. This is also evident from Eq. 20, where the
609 coefficients for \overline{T} and $\overline{Y_{Al}}$ are an order of magnitude larger than the quadratic term. The
610 TD points with errors in the top 10 percentile are also plotted. In Fig. 4b we plot the
611 posterior predictive results for 5 TD examples that were held-out when training the model.
612 We see that the prediction uncertainties are larger than ς from the MD simulations, though
613 all the data points (red symbols) are contained within the first and third quartiles of the
614 predictions. In addition, the uncertainty in the predictions, as quantified by the inter-
615 quartile range (the shaded box) as a proportion of the median prediction is quite small,
616 around 1%. These plots show that the data-driven model for σ is more predictive than
617 that of θ . Also note that these data-driven models are proxies for the MD simulation data;
618 errors latent in the MD simulations themselves have yet to be quantified.

619 *5.3. Sessile drop simulations with data-driven models*

620 To demonstrate that these data-driven models can be successfully incorporated into
621 finite element simulations, we simulated a hemispherical, Ag drop of radius 1 mm resting
622 on a substrate. Initially, the drop is at 1150 K and contains no aluminum. Throughout
623 the first half of the simulation, the drop is allowed to relax to an equilibrium state on the
624 substrate that is dictated by the contact angle and surface tension. At 0.02 s, we apply
625 a source term for the temperature and aluminum species at the substrate/drop interface,
626 mimicking a chemical reaction. The aluminum and energy are allowed to spread within
627 the drop via convection/diffusion. In typical brazing applications the heat transport
628 occurs several orders of magnitude faster than the mass transport. However, for illustrative
629 purposes we assigned fluxes that would allow the temperature and aluminum mass fraction
630 along the contact line to vary on similar time scales. As this mimicked reaction progresses,
631 the chemical composition and the temperature of the wetting surface change with time
632 leading to an evolution of the contact angle as shown in Fig. 5. Note that the contact angle
633 is post-processed based on the spatial mean of the unit normal vector along a reconstructed
634 sliver of interface 0.001 mm above the substrate. We found the post-processed contact angle
635 at this edge tolerance to be both not strongly influenced by the included area above the
636 substrate and not strongly susceptible to numerical fluctuations associated with a smaller
637 tolerance. The corresponding shapes of the drop at various times are shown in Fig. 6.

638 The FEM simulation demonstrates the ability to reproduce the wetting behavior of the
639 alloy at a length scale commensurate with engineering applications. As shown in Fig. 5,
640 the contact angle approaches the mean prediction from Eq. 6 and Eq. 19 that depends on
641 temperature and aluminum mass fraction. The deviation between the two at the start of
642 the simulation corresponds to the time needed for the drop to relax from its initial state,
643 which is still occurring slightly at 0.02 s. The fluctuations in the FEM simulations can be
644 attributed to discretization errors leading to localized errors in the post-processed contact
645 angle and can be minimized by using a smaller time step as shown in Appendix B. As
646 the aluminum mass fraction and temperature increase, additional deviation (about 0.6°)
647 between the simulation and the mean prediction from Eq. 7 persists, indicating there is a
648 slight lag for the simulations to respond to the new conditions. This deviation between
649 the simulation and the analytical model can be minimized by using a finer mesh as shown
650 in Appendix B.

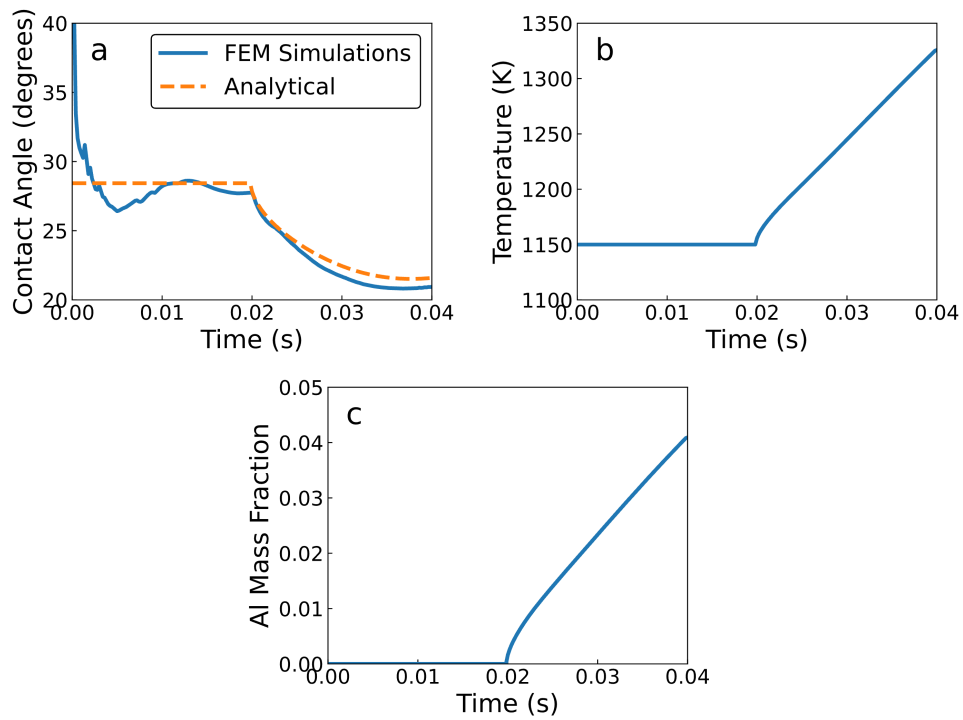


Figure 5: Results from FEM sessile drop simulations. (a) Post-processed contact angle as compared to the mean prediction (denoted analytical) from Eq. 6 and Eq. 19 for the prescribed temperature and aluminum mass fraction as plotted in (b) and (c), respectively. Note that the temperature and aluminum mass fraction are spatial means taken around alloy-atmosphere-substrate contact line.

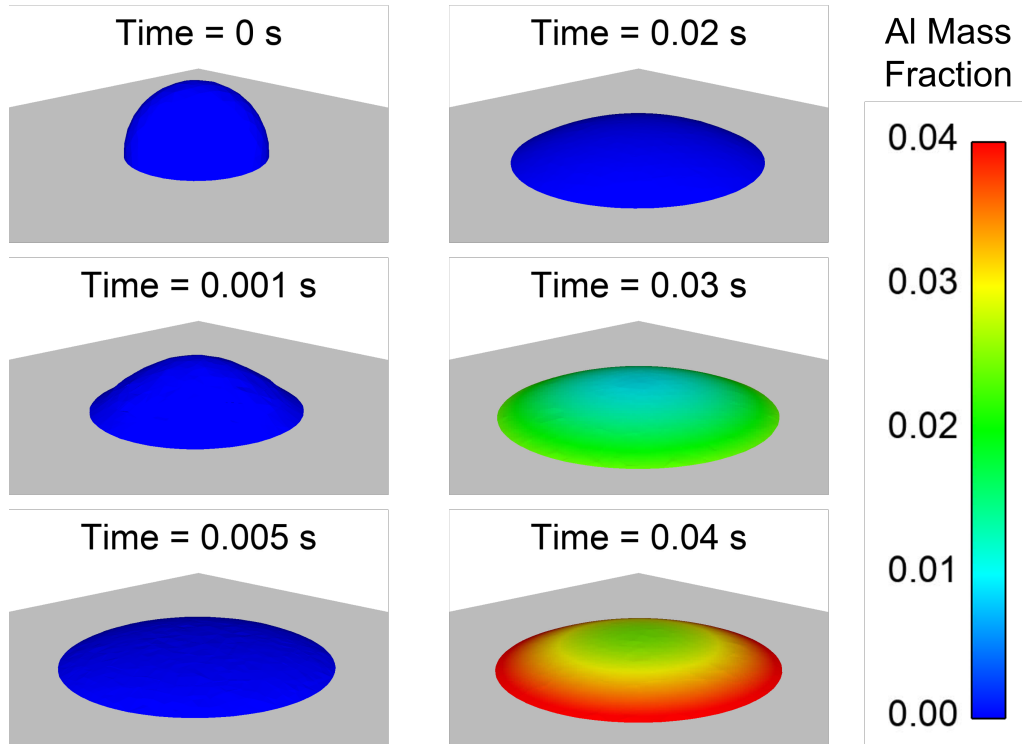


Figure 6: Depiction of FEM sessile drop shapes at different simulation times. The drop coloring corresponds to the aluminum mass fraction as indicated by the colorbar on the right.

651 5.4. Discussion

652 The discussion above focused on the development of the models $\theta = F(T, Y_{Al}; \mathbf{w})$ and
 653 $\sigma = G(T, Y_{Al}; \mathbf{v})$ and the shortcomings in their predictive skill versus the MD data. This
 654 first component of their model-form error is due to the simplicity of the data-driven model,
 655 which, in turn, is a consequence of the limited TD. However, these models have a second
 656 source for error from the MD simulations used to generate the TD.

657 The MD simulations generate θ using a quasi-2D model, rather than a 3D model. This
 658 was necessary for computational speed when assembling the TD and will result in smaller
 659 wetting angles compared to 3D simulations and experiments, as seen in Ref. [9], where the
 660 difference was about 5% for their Al-Au system. Further, the EAM potential is known
 661 to have difficulty predicting some multicomponent metallic alloy interactions [63, 64]. We
 662 consider this error acceptable to ensure the computational tractability of the TD; any errors
 663 in the resulting data-driven model can be corrected by assimilating experimental data via
 664 multi-fidelity modeling (see discussion below). In addition, our simulations assume a fixed
 665 substrate even though in reality the substrate dissolves into the molten filler, a phenomenon

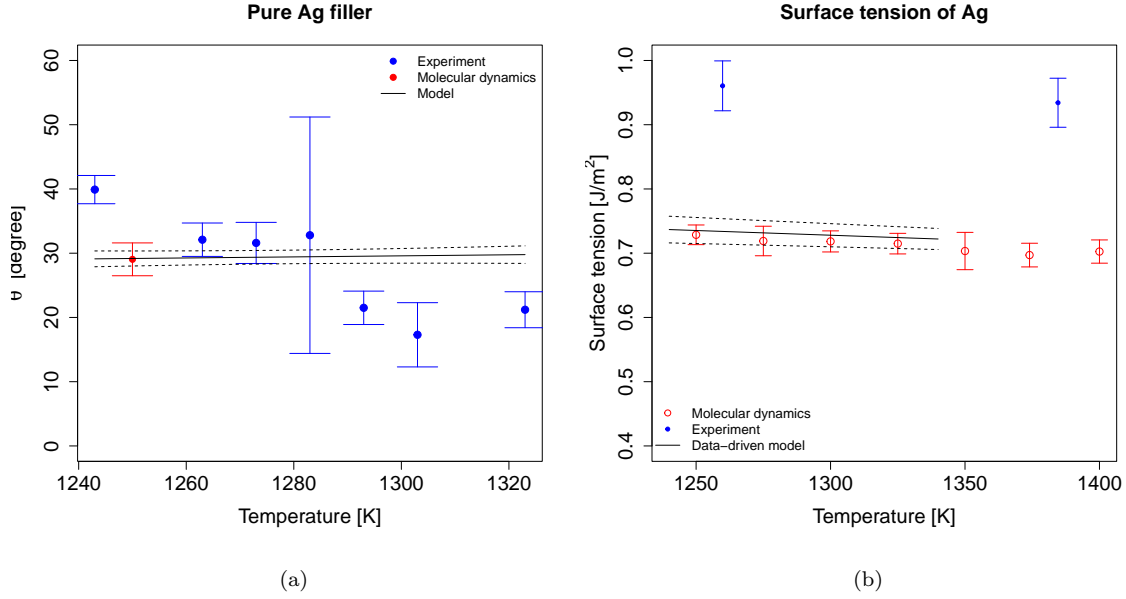


Figure 7: (a) Comparison of experimental data for pure Ag on Kovar with model predictions. A molecular dynamics simulation of the wetting angle at 1250 K, taken from our training dataset, is also plotted. (b) Comparison of experimental data for pure Ag surface tension (from Ref. [65]), MD simulations and our data-driven model’s predictions for σ . Error-bars are ± 2 standard deviation bounds. The solid line is the mean prediction using Eq. 7 and the dashed lines are the ± 2 standard deviation bounds computed using Eq. 7 and Eq. 19.

666 that our MD model does not currently capture. A fixed substrate keeps the braze/substrate
667 interfacial free energy artificially high by preventing mixing of the substrate and braze. As a
668 result, the fixed substrate condition overestimates the wetting angle[10]. The use of a fixed
669 substrate implies that our θ could be larger than experiments. Finally, the experiments
670 were conducted under 9 Torr pressure whereas the MD simulations were done in vacuum.
671 We now investigate these errors by comparing against experimental data acquired with a
672 pure Ag filler i.e., $Y_{Al} = 0$ wetting Kovar. Fig. 7a plots the experimental measurements
673 of θ versus the model predictions (i.e., means, using Eq. 7 and Eq. 19). For the modeled
674 results, the dashed lines show the ± 2 standard deviation bounds for the $\theta^{(pred)}$, computed
675 using 100 samples drawn from Eq. 19 and Eq. 7. For the experiments, the average value
676 of 12 measurements on each drop is plotted with symbols and the error-bars denote ± 2
677 standard deviation. The error bound for the MD data denotes the variation of θ over
678 a time-window where it was computed. The experiment at 1280 K was repeated twice,
679 and yielded somewhat different values of θ , leading to much larger uncertainty bounds.

680 Fig. 7a shows that the uncertainty in $\theta^{(pred)}$ is insignificant compared to the uncertainty of
681 experimental measurements, as one would expect from Fig. 3b where the uncertainties are
682 seen to be around 1° . Further, the model predicts hardly any change in the θ across the
683 temperature range, also expected from Fig. 3a and the weak linear dependence of $\theta^{(pred)}$
684 on temperature. In contrast, the experimental data shows a clear temperature dependence.
685 Note, also, that the data-driven model’s prediction agree with the MD simulation results
686 very well. In Fig. 7b we perform a similar comparison for σ compared to experimental
687 measurements (obtained from Ref. [65]). Again, the MD and data-driven models agree
688 in their predictions, but there is a distinct bias in the MD predictions when compared to
689 measurements.

690 **Model shortcomings and improvements:** Fig. 7 exposes the shortcomings of adopt-
691 ing an approximate MD model that lacks reactive processes to generate the TD. It also raises
692 the question of how to improve the $\theta = F(T, Y_{AI}; \mathbf{w})$ and $\sigma = G(T, Y_{AI}; \mathbf{v})$ models, under
693 the condition that high-fidelity MD simulations (i.e., higher fidelity than the EAM model)
694 are not possible e.g., due to resource constraints. The answer may lie in the structure of
695 the discrepancy between the modeled and measured data seen in Fig. 7. Fig. 7a shows
696 that the measurements vary around $\theta^{(pred)}$ as a smooth function of T . If the same struc-
697 ture holds true as Y_{AI} is varied, then it may be possible to add a correction $\delta\theta$ to $\theta^{(pred)}$,
698 perhaps as a Gaussian Process. For σ , plotted in Fig. 7b, the discrepancy between the
699 MD simulations and experimental data follows a simpler, almost linear trend in tempera-
700 ture, and an approach based on adding $\delta\sigma$ to $\sigma^{(pred)}$ may also improve the predictive
701 skill of the data-driven model. Improving the predictive skill of inaccurate, but computa-
702 tionally tractable, process-based models (our MD) by including a correction learned from
703 sparse high-fidelity data is a type of multi-fidelity modeling used in many engineering prob-
704 lems [66]. Multi-fidelity wetting models, e.g., based on co-kriging, are feasible in principle
705 but introduce new practical challenges. Primarily this requires the design and execution of
706 a new, and far sparser, design of experiments, conducted via laboratory measurements or
707 MD simulations that encompass reactions and can simulate the creation of IMP; these are
708 under way. Secondly, data-driven multi-fidelity models are often Gaussian Processes or neu-
709 ral networks, which are far more difficult (from a solver and numerical stability viewpoint)
710 and computationally expensive to integrate into a partial differential equation simulator
711 vis-à-vis our polynomial fits; this is left to future work.

712 6. Conclusions

713 In this paper, we have developed a general technique for modeling the surface tension (σ)
714 and wetting angle (θ) of a braze filler alloy on a substrate. These quantities are motivated
715 by a desire for engineering-scale simulations of brazed joints, and in particular to investigate
716 run-out i.e., the spread of the filler material beyond the surfaces being joined. In our case,
717 the filler material is a silver-aluminum (Ag, Al) alloy and the substrate is KovarTM. The
718 dependence of σ and θ on the filler temperature (T) and composition (Y_{Al}) is captured by
719 the model. We first summarize our findings and then discuss their ramification on being
720 able to model brazing processes.

721 **Summary of findings:** The modeling process fundamentally consists of generating
722 a training dataset of σ and θ for 50 (T, Y_{Al}) combinations by performing molecular dy-
723 namics simulations with EAM potentials, and capturing their dependencies via polynomial
724 (nominally quadratic) fits $\theta = F(T, Y_{Al}; \mathbf{w})$ and $\sigma = G(T, Y_{Al}; \mathbf{v})$. The molecular dynamics
725 simulations here do not capture all the physics occurring at the filler-substrate interface.
726 We chose to begin with this simplified model due to its tractable computational cost, a
727 necessity when a large (T, Y_{Al}) -space has to be covered to generate a training dataset. In
728 particular, we plan to extend this work to use chemically accurate (and far more compu-
729 tationally expensive) simulations, and therefore made the decision to limit computational
730 cost as much as possible even with these simpler simulation techniques. We show how the
731 (T, Y_{Al}) combinations may be generated in a space-filling manner, while removing some
732 combinations may represent filler material that was not a liquid. We also discuss the sim-
733 plification of the quadratic fits, commensurate with the information content of the limited
734 training datasets. The simplified structure of the polynomials incur a model-form error,
735 and consequently an uncertainty in the estimates of model parameters (\mathbf{w}, \mathbf{v}) when fitted
736 to data. These parameter estimates are computed in a Bayesian manner i.e., as a joint
737 probability density function. The uncertainty in the model predictions, and their disagree-
738 ment with MD simulations are quantified. The models for θ and σ were integrated with
739 an engineering-scale FEM, and used to simulate a drop initialized on a Kovar surface. The
740 FEM reproduces the dynamics of the sessile drop reaching its equilibrium configuration in a
741 qualitatively correct manner indicating that the integration of $F(T, Y_{Al}; \mathbf{w})$ and $G(T, Y_{Al}; \mathbf{v})$
742 did not destabilize the FEM or add noticeable stiffness to time-evolving simulations as they
743 approach equilibrium. This was fortunate as these data-driven models interact with the slip

744 model on the solid surface as well as the force that drives the contact line to the equilibrium
745 $\theta = F(T, Y_{Al}; \mathbf{w})$ angle.

746 Finally, we investigated the predictive skill of the data-driven model by comparing with
747 experiments. The discrepancies seem to be almost entirely due to the missing physics
748 in the simplified molecular dynamics simulations that are quasi-2D and use a fixed (i.e.,
749 non-reactive) substrate. We find that these discrepancies are far larger than the errors
750 introduced by the simplicity of $F(T, Y_{Al}; \mathbf{w})$ and $G(T, Y_{Al}; \mathbf{v})$ arising from the limited nature
751 of the training dataset. However, we find that the mismatch between experiment and model
752 predictions shows autocorrelation (when $Y_{Al} = 0$) i.e., a smooth variation as a function of
753 temperature T . If the same characteristic holds true for $Y_{Al} > 0$, the current models for σ
754 and θ could be augmented in a multi-fidelity manner with a small number of experiments
755 conducted at carefully chosen (T, Y_{Al}) combinations.

756 **Ramifications on modeling brazing processes:** The method described above is,
757 in essence, a data-driven means of distilling results from MD simulations into a form that
758 can be used in multicomponent, multiphase PDE models that can be used in engineering-
759 scale simulations of reactive wetting e.g., brazing. As shown in Fig. 7, the model may
760 need to be improved, perhaps by assimilating experimental data in a multi-fidelity manner
761 to be predictive, but its inclusion into FEM appears to be straightforward and robust.
762 The method can be extended to other configurations e.g., metal-ceramic brazed joints, by
763 generating a training dataset using appropriate MD simulations. The process is rigorous,
764 as it allows us to quantify the errors that arise due to modeling decisions. It therefore
765 constitutes a multiscale simulation framework for brazing processes as well as more generally
766 to wetting processes.

767 Active brazing is sometimes used to join metallic surfaces with ceramic ones, many of
768 which contain aluminum. In these cases, the filler does not contain Al; rather, it diffuses
769 out of the ceramic surface following a reaction with the braze alloy. We attempted to
770 model this effect in our FEM simulation of a sessile drop with a point source. However,
771 unlike our example where we were free to introduce an arbitrary amount of Al, a true
772 brazing simulation would have to not only introduce the correct amount, but also validate
773 that its transport inside the filler material is realistic. In addition, temperature and any
774 components/elements that diffuse/dissolve out of the metallic surface, and are transported
775 to the ceramic one, will doubtless affect the amount of Al that is released into the filler. Thus

776 a general modeling framework for active brazing is not feasible. However, the framework
777 described in this paper may be sufficient for modeling the brazing of two metallic surfaces.
778 Augmented with experimental data, via multifidelity models for θ and σ , it holds the promise
779 of being predictive.

780 **Acknowledgments**

781 This article has been authored by an employee of National Technology & Engineering
782 Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of
783 Energy (DOE). The employee owns all right, title and interest in and to the article and is
784 solely responsible for its contents. The United States Government retains and the publisher,
785 by accepting the article for publication, acknowledges that the United States Government
786 retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the
787 published form of this article or allow others to do so, for United States Government pur-
788 poses. The DOE will provide public access to these results of federally sponsored research
789 in accordance with the DOE Public Access Plan ([https://www.energy.gov/downloads/doe-](https://www.energy.gov/downloads/doe-public-access-plan)
790 [public-access-plan](https://www.energy.gov/downloads/doe-public-access-plan)). This paper describes objective technical results and analysis. Any
791 subjective views or opinions that might be expressed in the paper do not necessarily repre-
792 sent the views of the U.S. Department of Energy or the United States Government.

793 **References**

- 794 [1] Yichuan Zhang, Hanqi Zhang, Mingming Guo, Joël De Coninck, and David Seveno.
795 Reactive spreading dynamics of molten polymer liquids. Macromolecules, 56(3):1111–
796 1121, 2023.
- 797 [2] G. Kumar and K.N. Prabhu. Review of non-reactive and reactive wetting of liquids on
798 surfaces. Advances in Colloid and Interface Science, 133(2):61–89, 2007.
- 799 [3] S. Liu, D. L. Olson, G. P. Martin, and G. R. Edwards. Modeling of brazing processes
800 that use coatings and interlayers. Welding J., 70:207–215, 1991.
- 801 [4] T. Iida and R.I.L Guthrie. The Physical Properties of Liquid Metals. Clarendon Press,
802 1999.
- 803 [5] Wikipedia. Kovar page. <https://en.wikipedia.org/wiki/Kovar>; accessed June
804 2023.

- 805 [6] Walter H. Kohl. Materials and Techniques for Electron Tubes. Reinhold Publishing
806 Corporation, New York, NY, 1960.
- 807 [7] P.T. Vianco, C.A. Walker, D. De Smet, A. Kilgo, B.M. McKenzie, and R.L. Grant.
808 Interface reactions responsible for run-out in active brazing: Part 1. Welding J., 97:35–
809 S, 2018.
- 810 [8] W. Villanueva, W.J. Boettinger, G.B. McFadden, and J.A. Warren. A diffuse-interface
811 model of reactive wetting with intermetallic formation. Acta Materialia, 60(9):3799–
812 3814, 2012.
- 813 [9] Fei Wang, Andreas Reiter, Michael Kellner, Jürgen Brillo, Michael Selzer, and Britta
814 Nestler. Phase-field modeling of reactive wetting and growth of the intermetallic Al₂Au
815 phase in the Al-Au system. Acta Materialia, 146:106–118, 2018.
- 816 [10] M. Chandross. Competitive wetting in active brazes. Welding J., 94(5):169–175, 2015.
- 817 [11] E.B. Webb III, J.J. Hoyt, G.S. Grest, and D.R. Heine. Atomistic simulations of reactive
818 wetting in metallic systems. J. Mat. Sci., 40:2281, 2005.
- 819 [12] S. M. Foiles, M. I. Baskes, and M. S. Daw. Embedded-atom-method functions for the
820 fcc metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys. Phys. Rev. B, 33:7983–7991, June
821 1986.
- 822 [13] Jaroslav Mackerle. Finite element analysis and simulation of adhesive bonding, sol-
823 dering and brazing — An addendum: A bibliography (1996–2002). Modelling and
824 Simulation in Materials Science and Engineering, 10(6):637, September 2002.
- 825 [14] M.I. Barrena, J.M. Gomez de Salazar, and M. Gomez-Vacas. Numerical simulation
826 and experimental analysis of vacuum brazing for steel/cermet. Ceramics International,
827 40(7, Part B):10557–10563, 2014.
- 828 [15] W.C. Jiang, J.M. Gong, S.D. Tu, and H. Chen. Three-dimensional numerical simulation
829 of brazed residual stress and its high-temperature redistribution for stainless steel plate-
830 fin structure. Materials Science and Engineering: A, 499(1):293–298, 2009. Fifth
831 International Conference on Physical and Numerical Simulations of Material Processing
832 (ICPNS 2007) held at Zhengzhou, China, October 23-27, 2007.

- 833 [16] F. Meng, A. Liu, H. Sun, and M. Guo. Finite element analysis of thermal effects on
834 brazed cold diamond grit. Transactions of the Indian Institute of Metals, 68:829–838,
835 2015.
- 836 [17] Nicolas Eustathopoulos. Wetting by liquid metals—application in materials processing:
837 The contribution of the Grenoble Group. Metals, 5(1):350–370, 2015.
- 838 [18] Girish Kumar and K. Narayan Prabhu. Review of non-reactive and reactive wetting
839 of liquids on surfaces. Advances in Colloid and Interface Science, 133(2):61–89, 2007.
- 840 [19] Qiaoli Lin, Kaibin Xie, Ran Sui, Dekui Mu, Rui Cao, Jinghuan Chang, and Feng Qiu.
841 Kinetic analysis of wetting and spreading at high temperatures: A review. Advances
842 in Colloid and Interface Science, 305:102698, 2022.
- 843 [20] Ilhan A Aksay, Carl E Hoge, and Joseph A Pask. Wetting under chemical equilibrium
844 and nonequilibrium conditions. The Journal of Physical Chemistry, 78(12):1178–1183,
845 1974.
- 846 [21] V. Laurent, D. Chatain, and N. Eustathopoulos. Wettability of SiO₂ and oxidized
847 SiC by aluminium. Materials Science and Engineering: A, 135:89–94, 1991. European
848 Materials Research Society 1990 Spring Meeting on Metal Matrix Composites.
- 849 [22] J. C. Ambrose, M. G. Nicholas, and A. M. Stoneham. Dynamics of braze spreading.
850 Acta Metallurgica et Materialia, 40(10):2483–2488, 1992.
- 851 [23] Fei Wang and Britta Nestler. A phase-field study on the formation of the intermetallic
852 al₂au phase in the al–au system. Acta Materialia, 95:65–73, 2015.
- 853 [24] E. Saiz, R.M. Cannon, and A.P. Tomsia. Reactive spreading: adsorption, ridging and
854 compound formation. Acta Materialia, 48(18):4449–4462, 2000.
- 855 [25] M. Benhassine, E. Saiz, A. P. Tomsia, and J. De Coninck. Nonreactive wetting kinet-
856 ics of binary alloys: A molecular dynamics study. Acta Materialia, 59(3):1087–1094,
857 February 2011.
- 858 [26] Edmund B Webb III and Gary S Grest. Molecular dynamics simulations of reactive
859 wetting. Scripta Materialia, 47(6):393–398, September 2002.

- 860 [27] Edmund B. Webb, Gary S. Grest, David R. Heine, and J. J. Hoyt. Dissolutive wetting
861 of Ag on Cu: A molecular dynamics simulation study. Acta Materialia, 53(11):3163–
862 3177, June 2005.
- 863 [28] Hao Jiang, Florian Müller-Plathe, and Athanassios Z. Panagiotopoulos. Contact angles
864 from Young’s equation in molecular dynamics simulations. The Journal of Chemical
865 Physics, 147(8):084708, August 2017.
- 866 [29] Graham Saville. Computer simulation of the liquid–solid–vapour contact angle. Journal
867 of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics,
868 73(7):1122–1132, January 1977.
- 869 [30] M.J.P. Nijmeijer, C. Bruin, A.F. Bakker, and J.M.J. Van Leeuwen. A visual mea-
870 surement of contact angles in a molecular-dynamics simulation. Physica A: Statistical
871 Mechanics and its Applications, 160(2):166–180, 1989.
- 872 [31] Ricardo Macias-Salinas, Fernando García-Sánchez, and Otilio Hernandez-Garduza.
873 Viscosity model for pure liquids based on Eyring theory and cubic EOS. American
874 Institute of Chemical Engineers. AIChE Journal, 49(3):799, 2003.
- 875 [32] Julia Ling, Andrew Kurzawski, and Jeremy Templeton. Reynolds averaged turbulence
876 modelling using deep neural networks with embedded invariance. Journal of Fluid
877 Mechanics, 80:155–166, 2016.
- 878 [33] Pedro M Milani, Julia Ling, Gonzalo Saez-Mischlich, Julien Bodart, and John K Eaton.
879 A machine learning approach for determining the turbulent diffusivity in film cooling
880 flows. Journal of Turbomachinery, 140(2):021006, 2018.
- 881 [34] Ralph Smith. Uncertainty quantification: Theory, implementation and application.
882 SIAM Computational Science & Engineering, 2014.
- 883 [35] J. Ray, Z. Hou, M. Huang, K. Sargsyan, and L. Swiler. Bayesian calibration of
884 the community land model using surrogates. SIAM/ASA Journal on Uncertainty
885 Quantification, 3(1):199–233, 2015.
- 886 [36] Jaideep Ray, Sophia Lefantzi, Srinivasan Arunajatesan, and Lawrence Dechant. Learn-
887 ing an eddy viscosity model using shrinkage and bayesian calibration: A jet-in-crossflow

- 888 case study. ASCE-ASME Journal on Risk and Uncertainty in Engineering Systems,
889 Part B: Mechanical Engineering, 4(1), 2018.
- 890 [37] John H McDonald. Handbook of Biological Statistics, volume 2. Sparky House Pub-
891 lishing, Baltimore, MD, 2009.
- 892 [38] Jaideep Ray, Lawrence Dechant, Sophia Lefantzi, Julia Ling, and Srinivasan Aruna-
893 jatesan. Robust Bayesian calibration of a k - ε model for compressible jet-in-crossflow
894 simulations. AIAA Journal, 56(12):4893–4909, 2018.
- 895 [39] Clive D. Rodgers. Inverse Methods for Atmospheric Soundings: Theory and Practice.
896 World Scientific Publishing Co. Pte Ltd, Singapore, 2000.
- 897 [40] Jaideep Ray, Ali Pinar, and C Seshadhri. A stopping criterion for Markov chains when
898 generating independent random graphs. Journal of Complex Networks, 3(2):204–220,
899 2015.
- 900 [41] R. Byron Bird, Warren E. Stewart, and Edwin N. Lightfoot. Transport phenomena,
901 Revised 2nd edition. John Wiley & Sons, Inc., 2006.
- 902 [42] C.W Hirt and B.D Nichols. Volume of fluid (VOF) method for the dynamics of free
903 boundaries. Journal of Computational Physics, 39(1):201–225, jan 1981.
- 904 [43] Gunduz Caginalp. An analysis of a phase field model of a free boundary. Archive for
905 Rational Mechanics and Analysis, 92(3):205–245, sep 1986.
- 906 [44] Rekha R. Rao, Lisa A. Mondy, David R. Noble, Harry K. Moffat, Douglas B. Adolf,
907 and P.K. Notz. A level set method to study foam processing: A validation study.
908 International Journal for Numerical Methods in Fluids, 68(11):1362–1392, September
909 2011.
- 910 [45] Henning Braess and Peter Wriggers. Arbitrary lagrangian eulerian finite element anal-
911 ysis of free surface flow. Computer Methods in Applied Mechanics and Engineering,
912 190(1-2):95–109, October 2000.
- 913 [46] Richard M. J. Kramer and David R. Noble. A conformal decomposition finite element
914 method for arbitrary discontinuities on moving interfaces. International Journal for
915 Numerical Methods in Engineering, 100(2):87–110, June 2014.

- 916 [47] David R. Noble, Alec Kucala, and Mario J. Martinez. A conformal decomposition
917 finite element method for dynamic wetting applications. In Volume 1B, Symposia:
918 Fluid Measurement and Instrumentation; Fluid Dynamics of Wind Energy; Renewable
919 and Sustainable Energy Conversion; Energy and Process Engineering; Microfluidics
920 and Nanofluidics; Development and Applications in Computational Fluid Dynamics;
921 DNS/LES and Hybrid RANS/LES Methods. American Society of Mechanical Engi-
922 neers, July 2017.
- 923 [48] Scott A. Roberts, Hector Mendoza, Victor E. Brunini, and David R. Noble. A verified
924 conformal decomposition finite element method for implicit, many-material geometries.
925 Journal of Computational Physics, 375:352–367, December 2018.
- 926 [49] Pierre Joseph and Patrick Tabeling. Direct measurement of the apparent slip length.
927 Physical Review E, 71(3):035303, March 2005.
- 928 [50] Sashikumaar Ganesan, Sangeetha Rajasekaran, and Lutz Tobiska. Numerical modeling
929 of the non-isothermal liquid droplet impact on a hot solid substrate. International
930 Journal of Heat and Mass Transfer, 78:670–687, November 2014.
- 931 [51] V. E. B. Dussan. The moving contact line: the slip boundary condition. Journal of
932 Fluid Mechanics, 77(4):665–684, October 1976.
- 933 [52] R. G. Cox. The dynamics of the spreading of liquids on a solid surface. Part 1: Viscous
934 flow. Journal of Fluid Mechanics, 168(-1):169, July 1986.
- 935 [53] A. Raftery and Steven M. Lewis. Implementing MCMC. In W. R. Gilks, S. Richardson,
936 and D. J. Spiegelhalter, editors, Markov Chain Monte Carlo in Practice, pages 115–130.
937 Chapman and Hall, 1996.
- 938 [54] A. E. Raftery and S. M. Lewis. How many iterations in the Gibbs sampler? In J. M.
939 Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, editors, Bayesian Statistics,
940 volume 4, pages 765–766. Oxford University Press, 1992.
- 941 [55] Y. M. Bishop, S. E. Fienberg, and P. W. Holland. Discrete multivariate analysis:
942 Theory and practice. Springer-Verlag, New York, NY, 2007.
- 943 [56] Gregory R. Warnes and Robert Burrows. mcgibbsit: Warnes and Raftery’s MCGibbsit
944 MCMC diagnostic, 2013. R package version 1.1.0.

- 945 [57] W. Tillmann, J. Pfeiffer, and L. Wojarski. The apparent contact angle and wetted
946 area of active alloys on silicon carbide as a function of the temperature and the sur-
947 face roughness: A multivariate approach. Metallurgical and Material Transactions A,
948 46:3592—3600, 2015.
- 949 [58] Sierra/SolidMechanics 5.10 ITAR users' guide. Technical report, sep 2022.
- 950 [59] P.T. Vianco, C.A. Walker, D. De Smet, A. Kilgo, B.M. McKenzie, P.M. Kotula, and
951 R.L. Grant. Understanding the run-out behavior of a Ag-Cu-Zr braze alloy when used
952 to join alumina to an Fe-Ni-Co alloy. In R. Gourley and C. Walker, editors, Proc. 6th
953 International Brazing and Soldering Conference, 2015.
- 954 [60] X. W. Zhou, R. A. Johnson, and H. N. G. Wadley. Misfit-energy-increasing dislocations
955 in vapor-deposited coFe/nife multilayers. Phys. Rev. B, 69:144113, Apr 2004.
- 956 [61] John G. Kirkwood and Frank P. Buff. The Statistical Mechanical Theory of Surface
957 Tension. The Journal of Chemical Physics, 17(3):338–343, 12 2004.
- 958 [62] Wenbin Yu and D. Stroud. Molecular-dynamics study of surface segregation in liquid
959 semiconductor alloys. Phys. Rev. B, 56:12243–12249, Nov 1997.
- 960 [63] R. A. Johnson. Phase stability of fcc alloys with the embedded-atom method. Phys.
961 Rev. B, 41:9717–9720, May 1990.
- 962 [64] X. W. Zhou, R. A. Johnson, and H. N. G. Wadley. Misfit-energy-increasing dislocations
963 in vapor-deposited coFe/nife multilayers. Phys. Rev. B, 69:144113, Apr 2004.
- 964 [65] I. Egry, E. Ricci, R. Novakovic, and S. Ozawa. Surface tension of liquid metals and
965 alloys — recent developments. Advances in Colloid and Interface Science, 159(2):198–
966 212, 2010.
- 967 [66] M Giselle Fernández-Godino, Chanyoung Park, Nam-Ho Kim, and Raphael T Haftka.
968 Review of multi-fidelity models. arXiv preprint arXiv:1609.07196, 2016.

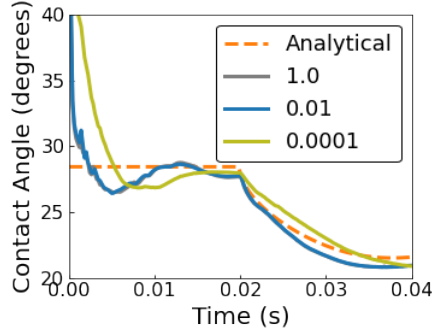


Figure A.8: Effects of varying the slip parameter β as compared to the analytical solution. The lines in the legend correspond to different values of β .

969 Appendix A. Slip parameter influence

970 To assess the influence of the semi-empirical slip parameter β on the FEM simulations,
 971 we conducted additional simulations with varying values of β . The results are shown in
 972 Fig. A.8 as compared to the analytical solution. Our results indicate that increasing β
 973 does not significantly affect the simulation as the dynamic behavior is restricted by the
 974 overall momentum balance. However, we found that decreasing β will result in delayed
 975 wetting behavior but may help to eliminate some numerical instabilities that arise during
 976 fast deformation of the interface.

977 Appendix B. Grid resolution study

978 The influence of mesh and time step sizing were assessed by independently varying the
 979 background mesh size from 0.25 mm to 0.15 mm and maximum time step from 0.1 ms to
 980 0.02 ms, as shown in Fig. B.9. Increasing the mesh resolution leads to the simulation more
 981 closely aligning with the analytical solution. This is likely due to the fact as the background
 982 mesh sizing increases, the facets along the alloy-atmosphere-substrate contact line are more
 983 dependent on the capillary effects above the contact line, resulting in a lower post-processed
 984 contact angle. Despite the improved accuracy, simulations with increased mesh resolution
 985 are more susceptible to numerical instabilities and result in fluctuations of the contact
 986 angle. However, as the maximum time step decreases, these numerical fluctuations can be
 987 minimized.

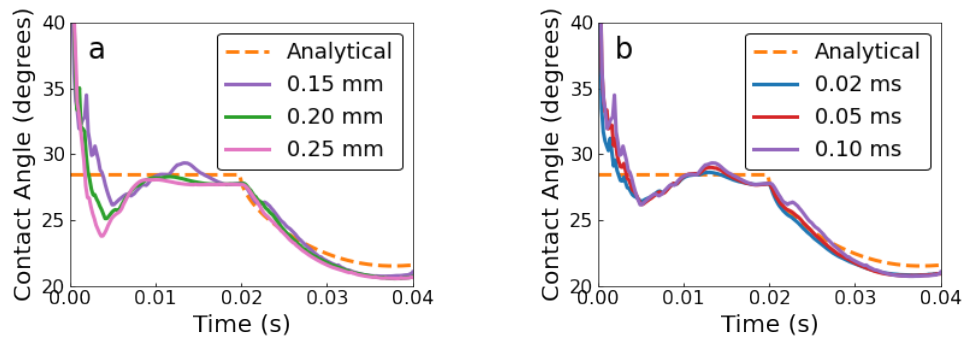


Figure B.9: Grid resolution study. (a) Effect of decreasing background mesh size for a fixed maximum time step of 0.1 ms. (b) Effect of decreasing maximum time step for a fixed background mesh size of 0.15 mm. The different lines in the legend correspond to mesh resolution (left figure) and time step size (right figure).