Physics Constrained Multi-objective Bayesian Optimization to Accelerate 3D Printing of Thermoplastics

Kianoosh Sattari[1#], Yunchao Wu[1#], Zhenru Chen[1], Alireza Mahjoubnia[1], Changhua Su[3], Jian Lin[1,2*]

[1]Department of Mechanical and Aerospace Engineering
[2]Department of Electrical Engineering and Computer Science
[3]Food Science Program, Division of Food, Nutrition & Exercise Sciences
University of Missouri, Columbia, Missouri 65211, United States
[#]Equal contribution
[*]E-mail: LinJian@missouri.edu

Abstract

Vat photopolymerization (VPP) of thermoplastics faces an issue of unpredictable printability due to the involved competing photopolymerization and dissolution reactions. The printing outcome largely depends on physicochemical properties of monomers and their compositions in resins, which also greatly determine the material properties, e.g., strength/toughness and phase transition temperature (T_g). Thus, a methodology for optimizing the resin formulation is of paramount importance in realizing highly printable thermoplastics with balanced strength/toughness and target T_g while remaining largely underexplored. Herein, we introduce a multi-objective Bayesian optimization (MOBO) algorithm with two physics informed constraints (printability and T_g) to optimize two conflicting properties: tensile strength (σ_T) and toughness (U_T). The two constraints are formulated as two machine learning (ML) models, which are trained with weight ratios of the six input monomers and physics informed (PI) descriptors derived from their physiochemical parameters. Dimensional reduction analysis reveals that the algorithm avoids recommendation of the monomer ratios that do not pass the two constraints. The printing failure rate is reduced from 16% in the background experiments to 3% in the recommended experiments. Within only 36 iterations (72 samples), the MOBO algorithm successfully identifies five sets of ratios leading to Pareto optimal of σ_T and U_T. Over these iterations, the hypervolume indicator increases while the error decreases. The partial dependence analysis indicates that σ_T and U_T depend on both the ratios and physiochemical features of the monomers. These results underscore capability of such a smart decision-making algorithm—with constraints that are not fully understood but can be informed by prior knowledge—in planning the experiments from the vast design space, thus holding a great promise for broader applications in materials and manufacturing.

Keywords: 3D printing, Bayesian optimization, constraints, multi-objective, physics informed, thermoplastics, vat photopolymerization

1. Introduction

Three-dimensional (3D) printing is increasingly popular for fabricating objects with complex structures and desired functionalities. Several common 3D printing technologies including stereolithography (SLA), vat photopolymerization (VPP) (e.g., digital light processing (DLP) and liquid-crystal display (LCD) printing), selective laser sintering (SLS), direct ink writing (DIW), and fused deposition modeling (FDM) have been utilized for this purpose. Among different types of materials for 3D printing, thermoplastics have recently become a surge of interest.[1] This surge stems from their remarkable recyclability, positioning them as viable alternatives to the conventional thermosets.[2] While FDM is commonly employed for printing the thermoplastics, it is a relatively slow and low-precision process. VPP, in contrast, offers much better printing resolution and faster printing rates. When subjected to digital light exposure, monomers in a photo-curable resin are polymerized to a polymer. Depending on the reaction kinetics, the cured polymer could be dissolved into the surrounding liquid resin. Thus, a rapid solid-liquid phase separation is critical, which constitutes a fundamental and enabling characteristic particularly for VPP based 3D printing of thermosets. Because the thermosets are usually infusible and insoluble due to their networked structures.[3] Although photocurable thermoplastics are a known entity, meeting the unique requirements for VPP proves to be a challenge. Because in the process, competition of two concurrent reactions—polymerization or solidification of a resin to a polymer and diffusion/dissolution of the cured polymer into the resin—results in an unpredictable printability issue, particularly if multiple monomers are used in the resin. Achieving the sought-after rapid liquid-solid separation is contingent upon dominance of the first reaction, while minimizing the influence of the second. A key to success is to promote the polymerization rate to obtain a

thermoplastic with a sufficiently high molecular weight. Because if the polymerization kinetics is slow, it would afford ample time for the cured thermoplastic to diffuse and dissolve into the resin.

The polymerization and the dissolution rates are affected by a range of physicochemical parameters of the monomers, e.g., reactivity, viscosity, molecular weight, glass transition temperature (T_g), and intermolecular forces.[4-6] In many cases, to print a thermoplastic with the tunable mechanical properties a resin consisting of different types of monomers with various weight ratios is usually utilized.[7] A resin made of mainly soft monomers usually shows a high dissolution rate and a low polymerization rate, resulting in a thermoplastic with a high stretchability and but a low strength. A resin composed of mainly hard monomers will likely have a high polymerization rate, resulting in a thermoplastic with a high strength and but a low stretchability. Given numerous options for the monomers and spacious choices in their ratios, the design space for the resins becomes enormous and is very difficult to be exhaustively searched. Thus, a methodology for optimizing the resin formulation is of paramount importance in realizing printable thermoplastics with improved tensile strength (σ_T) and toughness (U_T) as well as target T_g from the vast design space, while it remains largely underexplored. In a traditional practice, establishing correlation of the monomers' ratios with the resulting printability and properties of the materials requires extensive experimental trials.[8] Decision is mainly made by scientists who are well trained with domain knowledge. This human-centered, labor- and time-consuming practice greatly limits new material development efficiency by 3D printing. Sometimes, if the number of input variables dramatically increases, it could become powerless in finding a solution to realize the optimum materials properties with this practice. This challenge will become even bigger if an application involves optimization of two or more competing property objectives such as σ_T and U_T.

Bayesian optimization (BO) algorithms present a cost-effective avenue to tackling this challenge.[9-12] In BO, an objective function is approximated by a surrogate model that is cheap to evaluate and can be updated with both successful and failed samples.[13] They show quite success in single objective optimization (SOO).[14-16] As an example, Gongora et al. employed a BO algorithm as a decision maker in an autonomous robot to search structures with increased compressive toughness.[15] Balachandran et al. applied a BO algorithm for SOO, named efficient global optimization (EGO),[17] to optimize curie temperature of ferroelectric perovskites.[14] Liu et al. developed a BO framework for optimizing efficiency of the perovskite solar cells. Using prior experimental data as a probabilistic constraint they increased the success rate to 47% compared to 12% resulting from the Latin hypercube sampling (LHS) method.[18] However, many real-world applications necessitate the achievement of multiple performance objectives.[19,20] Cao et al. developed a multi-objective Bayesian optimization (MOBO) algorithm to design formulation consisting of three surfactants, a polymer, and a thickener.[21] To lower the number of experiments for evaluation, they trained a separate classifier to screen out the recommended formulation that might lead to failed experiment. The limitation is that the classifier acts as a post-screener and is not integrated in the MOBO algorithm, thus the algorithm must run again if the proposed samples do not pass the screening, leading to low searching efficiency. In another study, Erps et al. applied a MOBO algorithm without constraints to formulate the ink compositions for 3D printing thermosets.[22] Their method, a purely data-driven algorithm, which would lack generalizability when applied to other materials or manufacturing systems, e.g., it is not directly applicable to VPP of thermoplastics. Because, as stated previously, the competition polymerization and dissolution reactions in VPP of thermoplastics would make many recommended resin formulations fail to meet the printability and property constraints. Therefore, how to incorporate the constraints and

the prior knowledge in physics and chemistry directly in a MOBO framework to tackle application scenarios with limited available data remains a knowledge gap in the field.

Herein, we introduce a MOBO algorithm with two physics informed (PI) constraints to recommend feasible resin formulations for VPP—more specifically, LCD—of thermoplastics with optimized σ_T and U_T (the two objectives of the MOBO algorithm). We developed two ML models to output two constraints. The first one is to predict if a recommended resin formulation leads to a printable thermoplastic. The second one is to predict if a recommended resin formulation realizes a thermoplastic with T_g in a target range (10 °C $\leq T_g \leq$ 60 °C). They were trained by using the weight ratios of six monomers for the resins and their physiochemical parameters as the PI features. In this study, the six monomers include two soft monomers of 2-hydroxy-3-phenoxypropyl acrylate (HA) and isooctyl acrylate (IA) and four hard monomers of 1-vinyl-2-pyrrolidone (NVP), acrylic acid (AA), N-(2-hydroxyethyl)acrylamide (HEAA), and isobornyl acrylate (IBOA). To train MOBO, two separate Gaussian Processes (GPs) for both objectives of σ_T and U_T were chosen as the surrogate models. To pretrain the two constraint models, 43 initial weight ratios of the monomers were given by the LHS principle to afford a database. In the training process, functions were sampled from each fitted GP. Subsequently, a non-dominated sorting genetic algorithm (NSGA-II) was employed to identify the Pareto front among the sampled functions that meet the criteria of the two established constraint functions. In each iteration, the MOBO algorithm suggests two sets of ratios, which are used to print new thermoplastics for experimental evaluation. Data from these newly evaluated samples are included in the database to update both the two constraint functions and the two GPs. Within 36 iterations (recommending 72 new sets of the ratios), the MOBO algorithm can successfully discover five compositions leading to thermoplastics with Parato optimal fronts of σ_T and U_T and with T_g in the target range. Among them, one exemplifies

a remarkable U_T, characterized by a high stretchability while one showcases a high σ_T. Performance of the two constraint models is evaluated by comparing the failure rates of the initial training samples and recommended samples. The failure rate is reduced from 16% in the initial LHS-guided experiments to 3% in the MOBO recommended experiments. The unsatisfactory rate regarding T_g is also reduced from 35% in the LHS-guided experiments to 17% in the recommended samples, indicating that integrating the constraint models within the MOBO algorithm greatly enhances the efficiency and reduces the amount of required training data. We expect that this physics-constrained MOBO algorithm can be applied to metals,[23] shape memory polymers,[24] and hydrogels[25] in other processes.

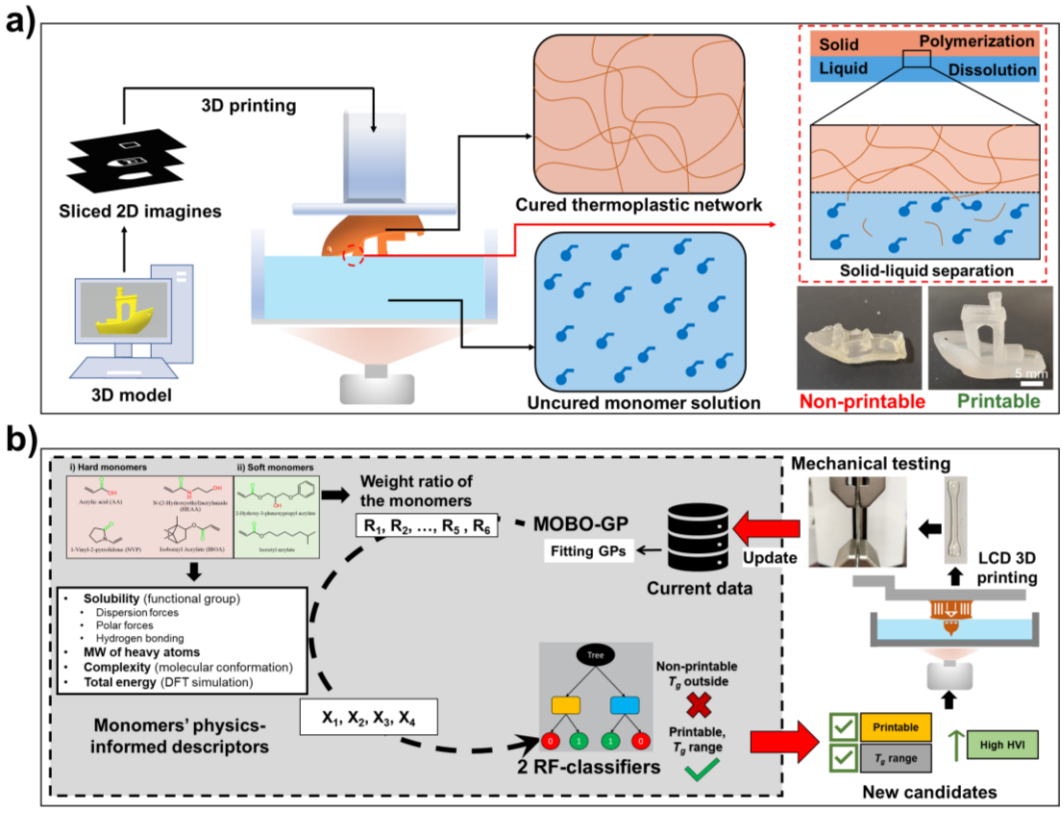

Figure 1. Development of a MOBO algorithm with PI constraints for 3D printing. (**a**) Schematic showing solid-liquid separation mechanism in LCD printing of thermoplastics and

photographs of two represented printable and nonprintable samples. (**b**) Workflow of the proposed MOBO framework with PI constraints for LCD thermoplastics printing.

2. Results

2.1. Architecture of the MOBO algorithm with the physics informed constraints

To formulate the resins, the six monomers as depicted in **Fig. S1** were mixed in various weight ratios. We first started with 3D printing of 43 initial samples with the ratios sampled by the LHS principle[18] with an aim to maximize the coverage of the design space. As illustrated in **Fig. 1a**, each set of the ratio was printed at least five times to obtain > five samples. Slow polymerization kinetics would provide the cured polymer ample time to diffuse and dissolve before reaching a desired molecular weight. If the resulting polymer's molecular weight remains low even with a satisfactory polymerization kinetics, it would still be susceptible to dissolution throughout the entire printing process, resulting in failed prints with obvious defects and incomplete shapes, as shown in **Fig. 1a**. The ones that were printable then underwent tensile testing to collect the stress-strain curves, from which two critical mechanical properties of σ_T and U_T were derived, as exemplified in **Fig. S2**. Due to the variances in the printed samples even if they are from the same set of monomer ratios, the derived σ_T and U_T were picked from a representative sample. T_g was measured from the differential scanning calorimetry (DSC) (**Fig. S3**). These 43 ratios along with the averaged σ_T, U_T, and T_g form an initial database. Then two random forest (RF) classifiers were pretrained for predicting if the resin is printable (yes/no) and if the resulting thermoplastic has T_g within a predefined range (10 °C $\leq T_g \leq$ 60 °C) (**Fig. S4**). The reason why we chose this T_g range is because the samples with T_g outside this range exhibit very poor mechanical properties, as shown

from the results of the initial 43 samples (**Fig. S5**). The samples with T_g in the range are classified as 1 otherwise classified as 0.

To make the RF models more generalizable and be better pretrained with limited experimental data, the PI descriptors of solubility, heavy atoms molecular weight, complexity, and total energy to represent the physiochemical properties of the monomers were serve as the inputs (**Fig. S4**). Solubility was calculated from three contributions of dispersion forces, polar forces, and hydrogen bonding.[6] Those contributions cannot be determined directly. One approach is to approximate them from the functional groups of the monomers. Complexity accounts for the atom size, symmetry, branches, rings, number of bonds, and heteroatoms characteristics of the input monomers.[26] The total energy was calculated by density function theory (DFT). Details about the calculations are explained in Materials and Methods. To make these additional features the variables in the RF model training, we normalized them to their respective weight ratios.

After the pretraining, the two RF classifiers were incorporated as the constraints in the MOBO algorithm (**Fig. 1b**). Two GPs that approximate the two objective functions were also pretrained on the initial 43 datasets. For training the GPs, only the weight ratios of the monomers were used as the input. In each iteration, the MOBO recommended two new sets of the weight ratios that were expected to lead to printable thermoplastics with the predicted T_g in the predefined range while showing improved σ_T and U_T. The recommended two sets of the ratios in each iteration were then used to make new resins to print thermoplastics which underwent tensile testing. The new datasets of the monomer ratios, printability, corresponding T_g, σ_T, and U_T were incorporated into the existing database to update both the two GPs and the two RF models. This loop continued until the predefined iteration number was reached or enough Pareto frontier points were obtained. More details are described in Materials and Methods.

2.2. Algorithm description

To optimize the two objectives concurrently, we adopted a MOBO strategy renowned for its efficacy in dealing with the black-box objective functions.[11] Our algorithm was built upon "Thompson sampling efficient multi-objective optimization" (TSEMO),[27] where Thompson sampling (TS) was used. As illustrated in **Fig. 2**, two independent GPs as the surrogate models were fitted to the initial 43 datasets to approximate the two objectives of σ_T and U_T. The essence of TS lies in the art of balancing the exploration of high-uncertainty regions and the exploitation of high-performing areas within the design space.[27] Regions with fewer data points are associated with higher uncertainty and therefore are more actively explored. TS chooses an action that matches the probability where the action leads to an improved reward in each iteration. It does so by drawing a sample based on the posterior distribution and returning an maximum index from the sampled vectors.[28] To sample a function from the posterior distribution of each GP, we employed spectral sampling, the details of which are outlined in the TSEMO paper.[27] Unlike SOO, which yields a single optimized outcome, MOO produces a region referred to as the Pareto front.[11] This region encompasses points that balance the objectives and cannot be surpassed by any other points. Consequently, enhancing one objective relative to the Pareto front necessitates compromising the other objective. Given that the sampled functions are computationally inexpensive to evaluate, we utilized NSGA-II, a genetic-based MOO method, to identify the Pareto front of the sampled functions. The Pareto set (sets of ratios) that pass the constraints of both printability and T_g were kept as the candidate set. From the candidate set, we then used the expected hypervolume improvement (EHVI) to select a set of the monomer ratios which leads to the largest improvement in the hypervolume. EHVI can be calculated based on a predicted Pareto frontier point and a reference point. The reference point is unknown and can be approximated by the anti-ideal point

from the candidate sets with the lowest σ_T and U_T. A schematic showing the hypervolume improvement and the reference point is shown in **Fig. S6**. Details about the EHVI can be found in the TSEMO paper.[27]

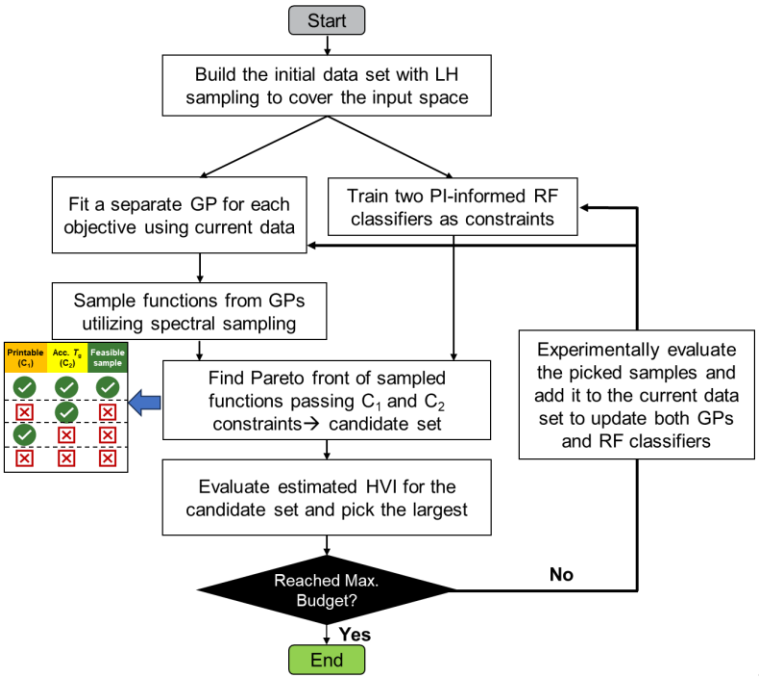

Figure 2. Flowchart showing the MOBO algorithm training. The newly evaluated samples are used to update GPs as well as the RF models which are the two constraints: C_1 and C_2.

2.3. Model Performance

Evaluation on printability and T_g. One innovation of the proposed MOBO algorithm is the PI constraints can greatly improve the success in both the printability and target T_g of the thermoplastics printed by the recommended sets of the monomer ratios. As presented in **Table 1**, the failure rate for printability is reduced to 3% for the ratios recommended by MOBO, compared to 16% for the initial samples by the LHS principle. The percentage of the printed samples with T_g outside the preassigned range (10-60 ℃) decreases from 35% to 17% for the ones printed using the sets of the weight ratios recommended by the MOBO algorithm. It's important to note that the

success rate regarding the printability is higher than that of T_g. This discrepancy can be attributed to closer correlation of the monomers' structural characteristics with the printability than that with T_g. This observation well agrees with our previous experimental findings, which suggest that the hydrogen interactions among the monomers play a crucial role in expediting polymerization and diminishing the solubility of the resulting polymers in liquid precursors.[4] This behavior promotes rapid solid-liquid phase separation, ultimately enhancing the printability. The behavior of T_g may more rely on other characteristics such as the molecule weights and chain-to-chain interactions, which could be informed by molecular dynamics (MD) simulations while it is out of current study.[29] The future direction of this research is to include more descriptors to inform the reactivity of the monomers as well as the polymer chain interactions for better informing T_g of the resulting polymers.

Table 1. Failure rates about printability and T_g for the initial samples recommended by LHS and the ones recommended by the proposed MOBO algorithm.

	43 Initial Samples by LHS		**72 New Samples by MOBO**	
RF1: printability	Printable # of samples	Failure rate	Printable # of samples	Failure rate
	36	16%	70	3%
RF2: T_g range	# of samples with acceptable T_g	Failure rate	# of samples with acceptable T_g	Failure rate
	28	35%	60	17%

This considerable improvement motivates us to further explore how the input ratios impact their printability and T_g of the resulting thermoplastics. However, visualizing the correlation in a 6-dimensional space is too difficult. To address this challenge, we employed Principal Component

Analysis (PCA) to reduce the dimensions for easy visualization and comparison. **Fig. 3a** represents a two-dimensional (2D) PCA map of the ratios with respect to the printability. Notably, the nonprintable samples which are predominately from the initial experiments are clustered in a distinct region, while the printable ones from both the initial and evaluated experiments are grouped in a separate region. That shows that the constraint model for printability can avoid the nonprintable recommendations. The two sets of nonprintable ratios out of 72 recommended samples are located close to other printable samples and could not be recognized by the trained RF constraint model. This insight underscores the impracticality of simply restricting the design space solely based on observations of a few infeasible samples. Instead, the intricacies of the relationship between these ratios and the printability necessitate the use of an ensemble-based model such as RF as a constraint. **Fig. 3b** shows the PCA map of the ratios with respect to T_g of the printable samples. The marked region shows the initial samples with T_g outside the predefined range. The constraint model can largely avoid this region, but there are still evaluated samples that do not satisfy the T_g requirement but are outside this region. It suggests that as we stated previously T_g is less correlated to the monomers' structural characteristics than the printability is.

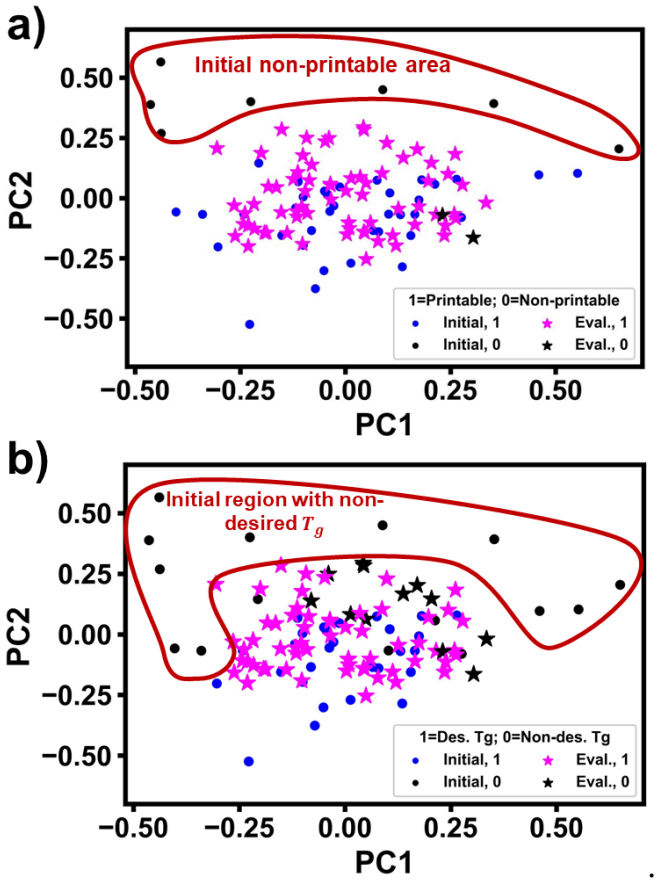

Figure 3. 2D PCA maps on the monomer weight ratios with respect to (**a**) printability and (**b**) T_g range labels.

Samples with the Parato frontier points of U_T and σ_T. Iteratively, as data from the newly evaluated samples was added, the two objective GPs were updated and started to converge. We examine how a relative error (RE), which is calculated as $|\text{predicted value} - \text{experimental value}|/\text{experimental value}$ for both σ_T and U_T, evolves as the number of the iterations (**Fig. S7**). The REs from the evaluated samples in each iteration are averaged to get average REs. They show relatively large fluctuations in the beginning of the iterations, which is expected, because the algorithm tends to explore different regions of the design space. As the projected design space is enormous, some of them can lead to poor recommendations.

Nevertheless, as shown by the trendlines, as the iteration increases the average REs of both U_T and σ_T decrease, indicating the model convergence.

Out of a total of 115 sets of ratios, comprising 43 initial and 72 evaluated ones—recommended by MOBO, 106 lead to printable thermoplastics. The σ_T and U_T values for these 106 printed thermoplastics are illustrated in **Fig. 4a**. From the 72 evaluated samples, we can identify five Pareto frontier points, each annotated with the corresponding sample number. For instance, Sample 15 means that it is the 15th evaluated sample. The five sets of the monomer ratios, T_g values, and σ_T and U_T are summarized in **Table S1**, from which we can see that Samples 15, 17, and 37 have high U_T and moderate σ_T while Samples 38 and 70 have high σ_T and moderate U_T. It is worth highlighting that the two soft monomers collectively for all the five samples occupy ~ 70% of the total weight, and the two hard monomers, NVP and IBOA, account for ≤ 5% each. **Fig. 4b** presents change of the hypervolume over the 72 evaluated samples. It increases from an initial value of 100 to 815 in the final iteration. It changes a lot at the beginning and then starts to saturate at the 40th sample. This indicates model convergence, signifying the effectiveness of the proposed algorithm in exploring the feasible design space. It is worth noting that within the realm of MOO featuring non-convex objective functions, the potential for further improvement remains, particularly with a larger experimental budget. To visually track the algorithm's progress in identifying the Pareto frontier points, **Fig. 4c** displays the cumulative number of Pareto points discovered over the course of evaluations. The data points in the graph indicate the count of Pareto points found at specific milestones for the 10th, 20th, 30th, 40th, 50th, 60th, and 70th samples.

Furthermore, we investigated evolution of the cumulative ratios of the soft and hard monomers (**Fig. 4d-e**). Notably, before the 45th sample, the recommended hard and soft monomer ratios substantially fluctuate, while in the later iterations they are stabilized at ~ 0.53 for the hard

monomers and 0.47 for the soft monomers, respectively. It is another indication for the model convergence. The recommended hard monomers and soft monomers ratios agree well with our intuition that the balanced soft and hard monomer ratios would lead to the thermoplastics with improved U_T and σ_T.

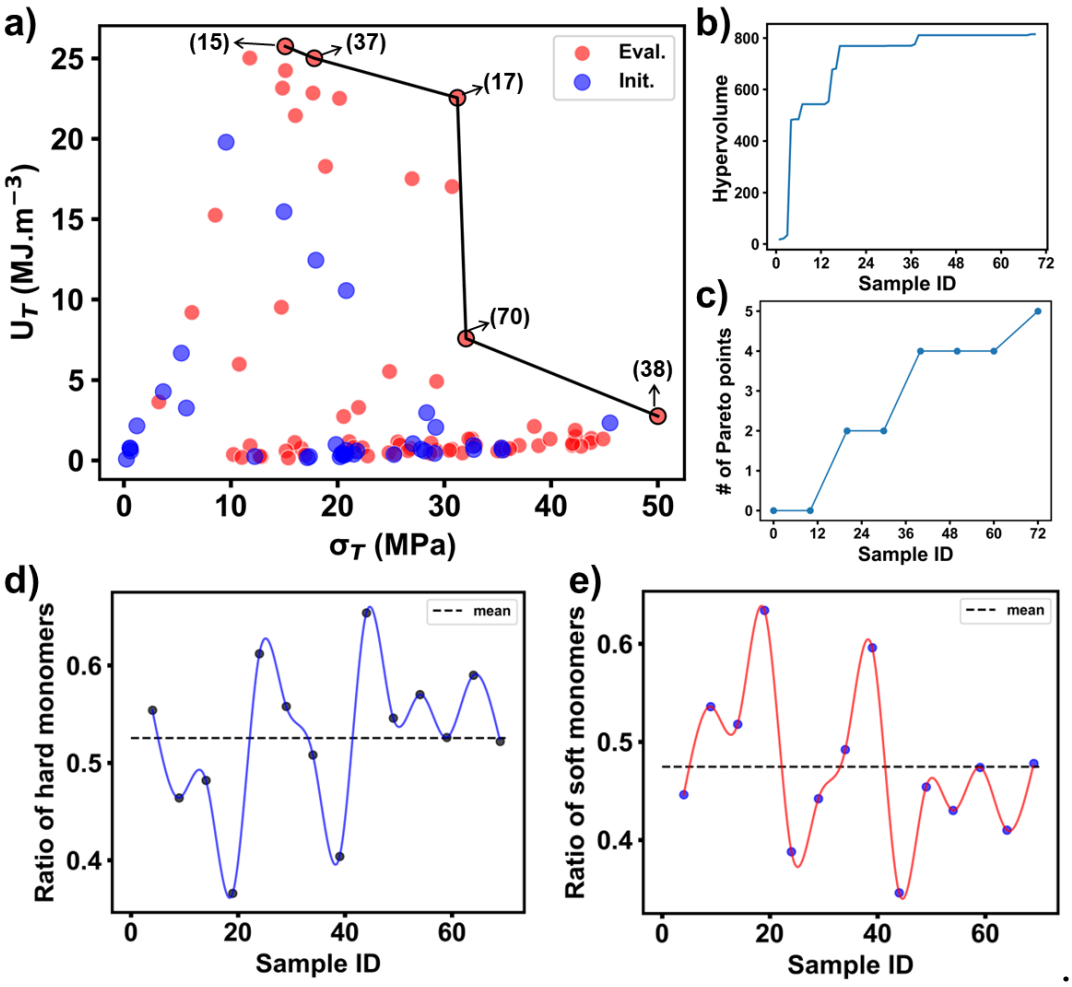

Figure 4. Performance evaluation of the model. (**a**) U_T versus σ_T of the initial samples (depicted in blue) and new evaluated ones from recommendations by MOBO (shown in red). The discovered five points form a Pareto front. (**b**) Hypervolume improvement in the evaluated samples. (**c**) Number of the Pareto frontier points discovered over the evaluated samples. Evolved cumulative hard monomer ratios (**d**) and soft monomer ratios (**e**) in the evaluated samples.

Partial dependance analysis (PDA) on the GPs' predictions. After the defined GPs were trained on all collected experimental data, all of their predictions are analyzed to disclose how the algorithm correlates the input ratios with the output properties. A series of plots, each depicting the partial dependence (PD) of U_T and σ_T of the thermoplastics on the input monomer ratios, are shown in **Fig. 5a** and **5b**, respectively.[30] One-way PD plots with two exmaples shown in **Fig. 5a-i** and **Fig. 5b-i** inform the correlation of the U_T and σ_T with the input ratio of a specific monomer. In these PD plots, we kept one ratio constant while allowing the other ratios to vary, and then averaged all the predictions made by the GPs. As shown in the blue plots, the ratios of the soft monomers (R_1 and R_2) mainly exhibit a positive correlation with U_T and a negative one with σ_T. In contrast, as indicated by the red plots, the ratios of the hard monomers (R_3, R_4, and R_5) mainly yield a negative influence on U_T but a positive one on σ_T. This trend agrees with our hypothesis that increase of the hard monomer ratios tend to increase the strength while the soft monomers favor toughness enhancement. As evident in **Fig. 5a-i**, increasing the ratio of R_2 up to 40% exhibits a positive correlation with U_T, which then reverses to a negative correlation with further increases. Notably, this threshold is significantly lower (~ 20%) when considering the influence of R_2 on σ_T, as shown in **Fig. 5b-i**. Additionally, the plots unveil that different hard monomer ratios exert differred effects. While NVP (R_3), a hard monomer, exhibits a linear, negative correlation with U_T, the effect of HEAA (R_5) on U_T is not quite so.

Two-way PD plots represent the dependence of U_T and σ_T on the two monomer ratios. They are shown in the color maps with brighter colors corresponding to higher values and vice versa (**Fig. 5**). The data points denoted with star markers correspond to the Pareto frontier points. It is important to note that not all Pareto frontier points are situated in regions with the highest values of the displayed objective. This observation underscores a fundamental aspect of this optimization

challenge: the two objectives largely conflict with each other. **Fig. 5a-ii** shows dependence of U_T on R_2 and R_4. We can see that with R_2 of < 20% U_T shows less dependence on R_4, while for R_2 > 20%, increasing R_4 up to ~ 20% increases U_T. But further increases in R_4 above 20% decreases U_T. **Fig. 5b-ii** shows dependence of σ_T on R_2 and R_4. It reveals that an increase in R_4 leads to higher σ_T when maintaining R_2 of < 20%. Overall, the PDA indicates that the correlation of U_T and σ_T with the monomer ratios is complicated and cannot be simply delineated. To reach optimized U_T and σ_T, the monomer ratios should be meticulously formulated, which is almost impossible by human. But the results show that the proposed MOBO algorithm would make this challenging problem possible.

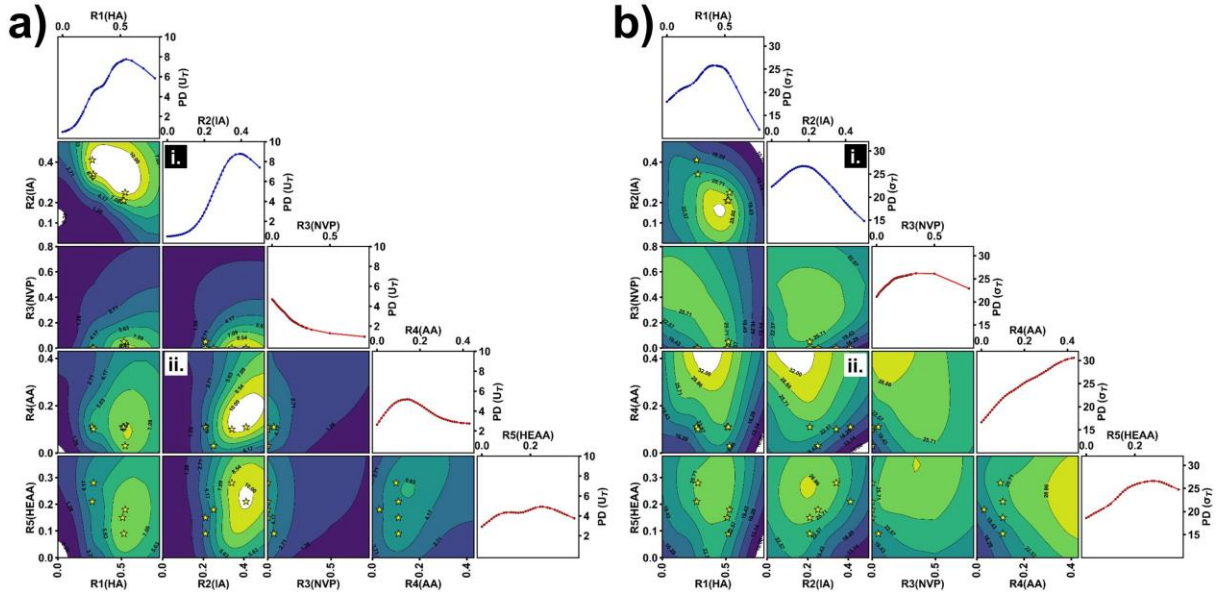

Figure 5. Partial dependance analysis of (a) U_T and (b) σ_T on ratios of the soft monomers (blue) and hard monomers (red). R_1: HA, R_2: IA; R_3: NVP, R_4: AA, R_5: HEAA. Panels (i) and (ii) are annotated for explanation in the text with (i) shows an example of a one-way PD while (ii) shows a two-way PD.

2.4. Applications of Pareto front samples with optimal trade-offs

The resin formulations that realize the Pareto set offer valuable materials for designing structures with tailored performance. Distribution of the five Pareto frontier points of U_T and σ_T are presented in **Fig. 6a**. Their representative stress-strain curves and respective T_g are illustrated in **Fig. 6b-f**. They show that these thermoplastics printed from the monomer ratios recommended by the MOBO algorithm can be classified into two categories. One category is characterized by high U_T with a large stretchability. The other one exhibits exceptional σ_T. It is worth noting that the constraint in T_g renders these materials with the shape memory properties in a suitable temperature range of 29—41 ℃, which is in the predefined range of 10-60 ℃ established from the initial samples, as illustrated in **Fig. S8**. Further analysis on the distribution of measured σ_T, U_T, and fractural strain (ε_f) with respective to T_g are presented in **Fig. S9**. As shown in **Fig. S9a**, elevating T_g up to 60 ℃ has a positive correlation to σ_T, while T_g above 60 ℃ results in a reduction f σ_T. The relationships between T_g and U_T, T_g and ε_f are depicted in **Fig. S9b** and **Fig. S9c**, respectively. They appear less straightforward but still highlight that samples falling outside the specified range exhibit notably lower toughness and strain.

With the Pareto frontier points' distinct performance attributes in U_T and σ_T, we illuminate two distinct applications. The first is to print a high-strength gripper from the set of the monomer ratios for Sample 38 (**Fig. 6g**). Due to phase transition above T_g, it shows a viscoelasticity property, thus can be easily programmed to a temporary, opened shape. It recovered to the original close shape upon heating, allowing it to securely grasp an object. Cooling down below T_g restores the stiffness of the gripper, ensuring a secure hold on the object (**Movie S1**). Additionally, a shape memory thermoplastic with high U_T finds application for printing a highly stretchable spring using the set of the monomer ratios for Sample 15. It was programmed to a fully stretched state when subjected to heat (**Fig. 6h**). This stretched spring recovered to its original shape when reheated

(**Movie S2**). Besides, among all the evaluated samples some show quite elastic performance with an example shown in **Fig. S10**. Although they do not show the properties in the Pareto front, they further prove the advantage of using the proposed MOBO algorithm for the resin formulation recommendation to print materials with the diverse mechanical behaviors.

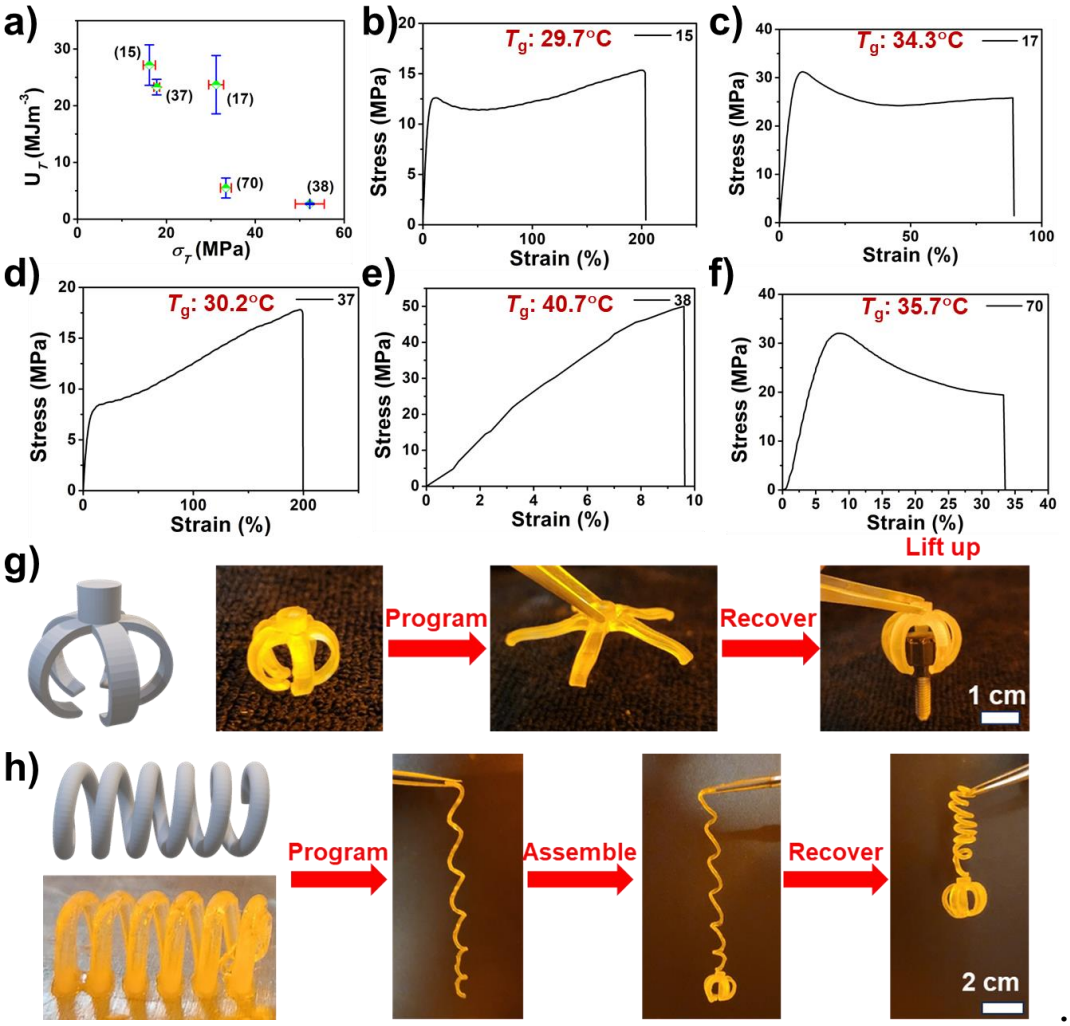

Figure 6. Shape memory behaviors of thermoplastics in the Pareto front of U_T and σ_T. (**a**) Distribution of U_T versus σ_T for the five Pareto front samples shown in Table S1. (**b-f**) Stress-strain plots of the five Pareto front samples along with their T_g. (**g, h**) Photographs showing shape memory behaviors for (**g**) a gripper printed from the set of the monomer ratios for Sample 38 and (**h**) a spring printed from the set of the monomer ratios for Sample 15.

3. Discussion

In this study, we designed and implemented a multi-objective Bayesian optimization algorithm with physics-informed constraints to accelerate the LCD 3D printing of thermoplastics with optimized σ_T, U_T and with T_g in a target range. To develop the algorithm, we designed two RF classifiers as the constraint models that are directly integrated in the framework. In each iteration, they screen the Pareto sets of the ratios which are predicted to be printable and lead to thermoplastics with T_g in a target range before their hypervolume indicators are calculated. As a result, the percentage of nonprintable sets of the monomer ratios was reduced from 16% in the initial samples to 3% in the recommended samples. The percentage of the printed samples with T_g in the target range increased from 65% to 83%. The developed MOBO algorithm afforded five Pareto frontier points of σ_T, U_T within only 36 iterations (72 recommendations). These thermoplastics exhibit T_g in a range of 29—41 ℃, which is suitable for use as shape memory polymers. While this MOBO algorithm with physics informed constraints was used for optimizing σ_T and U_T of the thermoplastics, it is expected to be applicable to diverse performance objective optimization in different manufacturing processes.

4. Materials and Methods

Materials. 2-Hydroxy-3-phenoxypropyl acrylate (HA), isooctyl acrylate (IA, > 90%), and acrylic acid (AA, 98%), were purchased from Sigma Aldrich (St. Louis, MO, U.S.). Diphenyl(2,4,6-trimethylbenzoyl)phosphine oxide (TPO, >97%), isobornyl acrylate (IBOA, > 90%), N-Vinylpyrrolidone (NVP, > 99%), and N-(2-hydroxyethyl)acrylamide (HEAA, >98%) were purchased from Fisher Scientific (Pittsburgh, PA, U.S.).

LCD 3D printing. The six monomers HA, IA, AA, IBOA, NVP, and HEAA with a certain weight ratio were mixed in a beaker to get a resin, which was added with a photoinitiator, diphenyl(2,4,6-trimethylbenzoyl)phosphine oxide (TPO) (2 wt%) and then magnetically stirred for 1 minute. The uniformly miscible resin was used for printing by a LCD 3D printer (Anycubic Photon Mono 4K) with an irradiation wavelength of 405 nm. In each experiment, a constant power density of ~5 mW/cm^2, a layer thickness of 50 μm, and an exposure of 15 s were used. After the printing, the samples were post-cured by 405-nm UV light for 60 s.

Mechanical testing of the printed samples. The tensile testing was conducted on a Mark-10 universal testing machine at a loading rate of 50 mm/min^{-1}. For each set of the monomer ratios, at least five samples were printed and tested for statistical analysis.

Physics informed input descriptors for RF models. The extracted descriptors are solubility, total energy, molecular weight of heavy atoms, and complexity. The solubility parameters were predicted from group contributions of the monomers.[6] The total solubility was calculated from Eq. 1 and Eq. 2.

$$\delta^2 = \delta_d^2 + \delta_p^2 + \delta_h^2 \quad (1)$$

$$\delta_d = \frac{\sum \boldsymbol{F}_{di}}{V};\ \delta_p = \frac{\sqrt{\sum \boldsymbol{F}_{pi}^2}}{V};\ \delta_h = \sqrt{\frac{\sum \boldsymbol{E}_{hi}}{V}} \quad (2)$$

where δ_d, δ_p, and δ_h are the contribution of dispersion forces, the contribution of polar forces, and the contribution of hydrogen bonding, respectively. $\mathbf{F}_{di}$, $\mathbf{F}_{pi}$, and $\mathbf{E}_{hi}$ for different functional groups were extracted from Table 7.10 in the book by Krevelen.[6] V is the molar volume of the monomers. The total energy of each monomer was calculated by DFT. First, we used Open Babel, an open chemical toolbox,[31] to construct the 3-dimensional coordinates from the simplified molecular-input line-entry system (SMILES) of the monomers. Then, 50 steps of local optimization were carried out by the MMFF94 force field. The DFT calculations for all monomers were carried out

using Gaussian-16-A.03, licensed under the University of Missouri. Geometry optimization and frequency calculations were carried out using the B3LYP functional with the basis 6-31G(2df, p) to calculate the total energy. Complexity accounts for the atom size, symmetry, branches, rings, number of bonds, and heteroatoms characteristics of the input monomers.[26] We extracted the complexity values of the monomers from the PubChem database,[32] which were calculated by the Bertz's approach.[26] We also extracted molecular weight of heavy atoms of the monomers from the PubChem database. All the calculated descriptors are listed in **Table S2**.

Model parameters. In each iteration, the GPs for both objective were updated as new datasets from the evaluated experiments were added into the database. To tune the hyperparameters of the GPs, we employed a maximum posteriori estimate method proposed by Bradford et al.[27] The kernel function is the Matérn 5/2, which is versatile and supports a wide range of function types.[33] To solve MOO of the sampled GPs, picked by TS sampling, a standard NSGA-II solver was applied with a population size of 100 and a total number of 100 generations.[34] To handle the constraints inside the solver, we used the "feasibility first" approach introduced in the *pymoo* python library.[35] To evaluate a solution, the solver first checks whether the solution satisfies the constraints. In this case, the solver does not calculate the GP values for the solutions that do not pass the constraints. Thus, we defined the values of ≤ 0 in Eq. 3 and Eq. 4 as satisfaction with the constraints.

$$g_1 : (-1) \times \mathbb{P}(\mathit{Printable}) + 0.7 \leq 0 \quad (3)$$

$$g_2 : (-1) \times \mathbb{P}\left(\mathit{Acceptable}\ T_g\right) + 0.7 \leq 0 \quad (4)$$

Where, $\mathbb{P}$ is the predicted probability from the RF models. Numbers closer to 1 indicate higher confidence in predictions. We multiplied (-1) to the predicted values from the RF models since we considered positive values to be the infeasible recommendations. We added 0.7 to the predicted

values to bias the recommendations that are highly probable being feasible for printability and satisfaction of T_g. The hyperparameters of the fitted RF models are listed in **Table S3**.

Design space. The ratio of each monomer can take continuous values from 0 to 1. Given the resolution limitation, we only considered two decimals, which resulted in 101 possible values for each ratio. To ensure that the sum of all proposed monomer ratios equates to 1, we considered only five ratios as the design space in the algorithm, as shown in Eq. 5.

$$g_3 : R_1 + R_2 + R_3 + R_4 + R_5 - 1 \leq 0 \quad (5)$$

with $0 \leq R_i \leq 1 \; for \; i = 1, 2, 3, 4, 5$

Where R_1 to R_5 are ratios of HA, IA, NVP, AA, and HEAA, respectively. **Fig. S12** displays the distribution of two ratios in each plot for all the initial and evaluated 115 samples. The ratio of the last monomer, IBOA, was determined by Eq. 6.

$$R_6 = 1 - (R_1 + R_2 + R_3 + R_4 + R_5) \quad (6)$$

References and Notes

1 Maines, E. M., Porwal, M. K., Ellison, C. J. & Reineke, T. M. Sustainable advances in SLA/DLP 3D printing materials and processes. *Green Chem.* **23**, 6863-6897 (2021).

2 Zhu, G., Hou, Y., Xu, J. & Zhao, N. Reprintable Polymers for Digital Light Processing 3D Printing. *Adv. Funct. Mater.* **31**, 2007173 (2021).

3 Deng, S., Wu, J., Dickey, M. D., Zhao, Q. & Xie, T. Rapid Open-Air Digital Light 3D Printing of Thermoplastic Polymer. *Adv. Mater.* **31**, 1903970 (2019).

4 Wu, Y. *et al.* Photocuring Three-Dimensional Printing of Thermoplastic Polymers Enabled by Hydrogen Bonds. *ACS Appl. Mater. Interfaces* **13**, 22946-22954 (2021).

5 Wu, Y. *et al.* H-bonds and metal-ligand coordination-enabled manufacture of palm oil-based thermoplastic elastomers by photocuring 3D printing. *Addit. Manuf.* **47**, 102268 (2021).

6 Van Krevelen, D. W. & Te Nijenhuis, K. in *Properties of Polymers (Fourth Edition)* (eds D. W. Van Krevelen & K. Te Nijenhuis) 189-227 (Elsevier, 2009).

7 Chin, K. C. H., Cui, J., O'Dea, R. M., Epps, T. H., III & Boydston, A. J. Vat 3D Printing of Bioderivable Photoresins – Toward Sustainable and Robust Thermoplastic Parts. *ACS Sustain. Chem. Eng.* **11**, 1867-1874 (2023).

8 Zeng, Y. *et al.* Highly Stretchable Fatty Acid Chain-Dangled Thermoplastic Polyurethane Elastomers Enabled by H-Bonds and Molecular Chain Entanglements. *ACS Sustain. Chem. Eng.* **10**, 11524-11532 (2022).
9 Xie, Y., Sattari, K., Zhang, C. & Lin, J. Toward autonomous laboratories: Convergence of artificial intelligence and experimental automation. *Prog. Mater Sci.* **132**, 101043 (2023).
10 Sattari, K., Xie, Y. & Lin, J. Data-driven algorithms for inverse design of polymers. *Soft Matter* **17**, 7607-7622 (2021).
11 Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & Freitas, N. d. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE CVPR* **104**, 148-175 (2016).
12 Rickman, J. M., Lookman, T. & Kalinin, S. V. Materials informatics: From the atomic-level to the continuum. *Acta Mater.* **168**, 473-510 (2019).
13 Raccuglia, P. *et al.* Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73-76 (2016).
14 Balachandran, P. V., Kowalski, B., Sehirlioglu, A. & Lookman, T. Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning. *Nat. Commun.* **9**, 1668 (2018).
15 Gongora, A. E. *et al.* A Bayesian experimental autonomous researcher for mechanical design. *Sci. Adv.* **6**, eaaz1708 (2020).
16 Liu, Z. *et al.* Machine learning with knowledge constraints for process optimization of open-air perovskite solar cell manufacturing. *Joule* **6**, 834-849 (2022).
17 Jones, D. R., Schonlau, M. & Welch, W. J. Efficient Global Optimization of Expensive Black-Box Functions. *J. Glob. Optim.* **13**, 455-492 (1998).
18 Stein, M. Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics* **29**, 143-151 (1987).
19 Torres, J. A. G. *et al.* A Multi-Objective Active Learning Platform and Web App for Reaction Optimization. *J. Am. Chem. Soc.* **144**, 19999-20007 (2022).
20 Gopakumar, A. M., Balachandran, P. V., Xue, D., Gubernatis, J. E. & Lookman, T. Multi-objective Optimization for Materials Discovery via Adaptive Design. *Sci. Rep.* **8**, 3738 (2018).
21 Cao, L. *et al.* Optimization of Formulations Using Robotic Experiments Driven by Machine Learning DoE. *Cell Rep. Phys. Sci.* **2**, 100295 (2021).
22 Erps, T. *et al.* Accelerated discovery of 3D printing materials using data-driven multiobjective optimization. *Sci. Adv.* **7**, eabf7435 (2021).
23 Saccone, M. A., Gallivan, R. A., Narita, K., Yee, D. W. & Greer, J. R. Additive manufacturing of micro-architected metals via hydrogel infusion. *Nature* **612**, 685-690 (2022).
24 Zhang, C. *et al.* 4D Printing of shape-memory polymeric scaffolds for adaptive biomedical implantation. *Acta Biomater.* **122**, 101-110 (2021).
25 Sun, J.-Y. *et al.* Highly stretchable and tough hydrogels. *Nature* **489**, 133-136 (2012).
26 Bertz, S. H. The first general index of molecular complexity. *J. Am. Chem. Soc.* **103**, 3599-3601 (1981).
27 Bradford, E., Schweidtmann, A. M. & Lapkin, A. A. Efficient multiobjective optimization employing Gaussian processes, spectral sampling and a genetic algorithm. *J. Glob. Optim.* **71**, 407-438 (2018).

28 Shahriari, B., Wang, Z., Hoffman, M. W., Bouchard-Côté, A. & de Freitas, N. An entropy search portfolio for Bayesian optimization. Preprint at https://arxiv.org/abs/1406.4625 (2014).
29 Banerjee, A., Iscen, A., Kremer, K. & Kukharenko, O. Determining glass transition in all-atom acrylic polymeric melt simulations using machine learning. *J. Chem. Phys.* **159** (2023).
30 Hastie, T., Friedman, J. & Tibshirani, R. in *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (eds Trevor Hastie, Jerome Friedman, & Robert Tibshirani) 299-345 (Springer New York, 2001).
31 O'Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *J. Cheminf.* **3**, 33 (2011).
32 Kim, S. *et al.* PubChem 2023 update. *Nucleic Acids Res.* **51**, D1373-D1380 (2022).
33 Rasmussen, C. E. & Williams, C. K. I. *Gaussian Processes for Machine Learning*. (The MIT Press, 2005).
34 Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **6**, 182-197 (2002).
35 Blank, J. & Deb, K. Pymoo: Multi-Objective Optimization in Python. *IEEE Access* **8**, 89497-89509 (2020).

Acknowledgements

Funding: J. L. thanks the financial support from National Science Foundation (award number: 2154428), U.S. Army Corps of Engineers, ERDC (grant number: W912HZ-21-2-0050). The DFT computation in this work was performed in the University of Missouri high-performance computing infrastructure provided by Research Support Solutions and in part by the National Science Foundation under the grant number of CNS-1429294. **Author Contributions:** J.L conceived the project. K.S. developed the MOBO framework. Y.W. chose the initial monomers and conducted all the experiments. Z.C. designed the 3D printing models and studied the shape memory property. A.M. helped K.S. to run the optimization iterations. C.S. draw the schematic figures and helped Y.W. in the experiments. J.L. oversaw all research phases and provided regular guidance. K.S. and Y.W. wrote the first draft, which was thoroughly corrected and finalized by J.L.. All authors discussed and commented on the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Code and source data are

available in “https://github.com/linresearchgroup/PIMOBO_TP3DPrinting.git”. Additional information can be requested from the corresponding author.

Supplementary Materials

Figures S1-S11

Movies S1-S2

Supporting Information for

Physics Constrained Multi-objective Bayesian Optimization to Accelerate 3D Printing of Thermoplastics

Kianoosh Sattari[1#], Yunchao Wu[1#], Zhenru Chen[1], Alireza Mahjoubnia[1], Changhua Su[3], Jian Lin[1,2*]

[1]Department of Mechanical and Aerospace Engineering

[2]Department of Electrical Engineering and Computer Science

[3]Food Science Program, Division of Food, Nutrition & Exercise Sciences

University of Missouri, Columbia, Missouri 65211, United States

[#]Equal contribution

[*]E-mail: LinJian@missouri.edu

Supplementary Tables:

Table S1. Ratios, mechanical properties, and transition temperature of the discovered five Pareto frontier points.

Sample ID	R1 (HA)	R2 (IA)	R3 (NVP)	R4 (AA)	R5 (HEAA)	R6 (IBOA)	σ_T (MPa)	U_T (MJ.m^{-3})	T_g (°C)
15	0.27	0.41	0.00	0.11	0.21	0.00	15.10	25.76	29.7
17	0.28	0.34	0.00	0.10	0.28	0.00	31.22	22.56	34.3
37	0.53	0.25	0.00	0.03	0.18	0.01	17.82	25.02	30.2
38	0.51	0.21	0.00	0.11	0.15	0.02	49.99	2.75	40.7
70	0.52	0.21	0.05	0.11	0.09	0.02	32.03	7.58	35.7

Table S2. Extracted and calculated descriptors for the six monomers.

Features / Monomer	Complexity	Energy (Hartree)	MW of HAs (g/mol)	δ_d	δ_p	δ_h	δ
HA	221.00	-766.60	208.13	17.49	5.66	12.09	22.00
IA	55.90	-581.70	164.12	15.57	4.14	4.89	16.84
NVP	120.00	-364.00	102.07	17.87	10.39	8.09	22.20
AA	55.90	-267.20	68.03	16.04	13.39	17.91	27.52
HEAA	90.40	-401.20	106.06	17.34	9.07	15.55	25.00
IBOA	306.00	-657.90	188.14	17.99	4.11	4.86	19.08

Note: δ is the total solubility that was calculated from three components of δ_d, δ_p, and δ_h, which are the dispersion force, polar force, and hydrogen bonding, respectively. The unit for all solubility components is $(MJ.m^{-3})^{0.5}$. MW stands for molecular weight. HAs stands for heavy atoms.

Table S3. Hyperparameters of the trained random forest (RF) models used as constraints.

Hyperparameter	Value
max_depth	5
min_sample_split	2
n_estimators	50

Supplementary Figures:

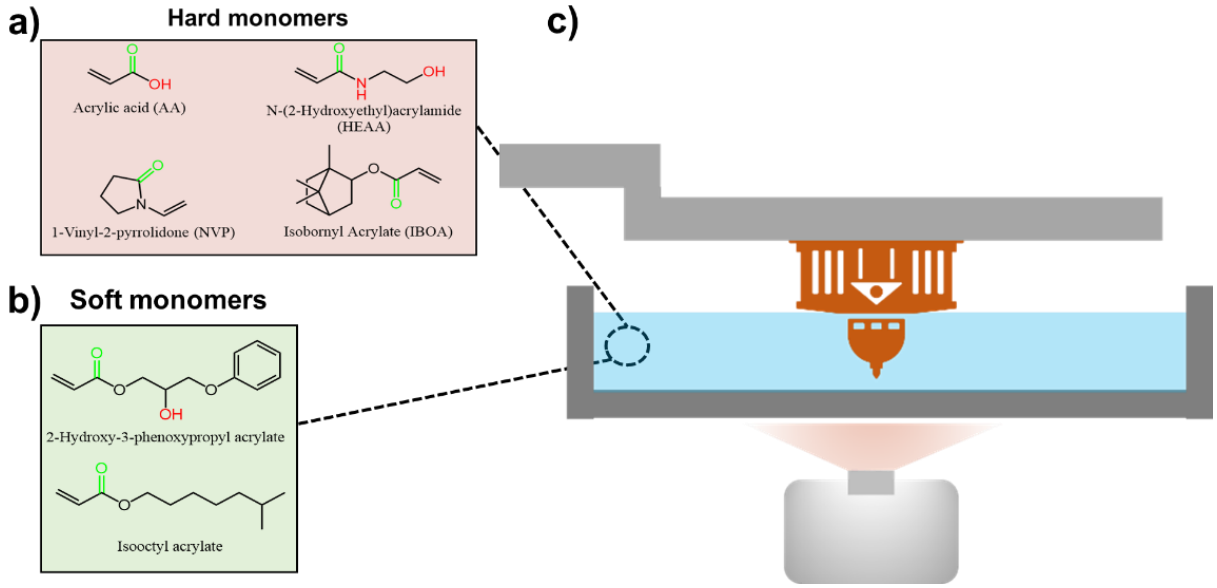

Figure S1. LCD 3D printing of thermoplastics consisting of hard and soft monomers. Pool of (**a**) hard and (**b**) soft monomers. (**c**) Schematic of LCD 3D printing.

SMILES of the six monomers shown in Fig. S1a and Fig. S1b:

1. 2-Hydroxy-3-phenoxypropyl acrylate (HA): “C=CC(=O)OCC(COC1=CC=CC=C1)O”
2. Isooctyl acrylate (IA): “CC(C)CCCCCOC(=O)C=C”
3. 1-Vinyl-2-pyrrolidone(NVP): “C=CN1CCCC1=O”
4. acrylic acid (AA): “C=CC(=O)O”
5. N-(2-Hydroxyethyl)acrylamide (HEAA): “C=CC(=O)NCCO”
6. Isobornyl acrylate (IBOA): “CC1(C2CCC1(C(C2)OC(=O)C=C)C)C”

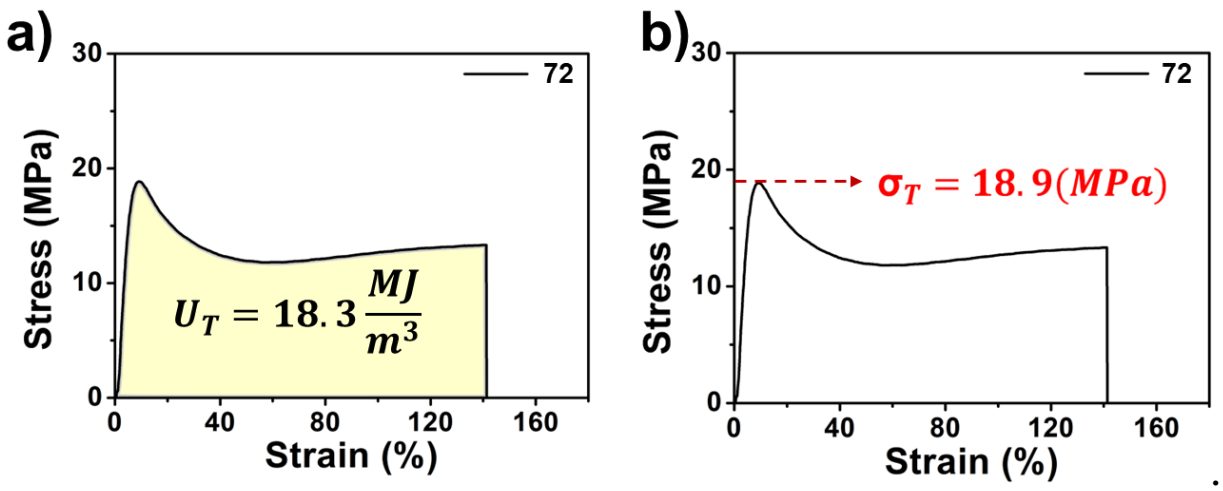

Figure S2. Stress-strain curves for deriving the two performance objectives for optimization by the MOBO algorithm: (**a**) **Tensile toughness; (b) Tensile strength**.

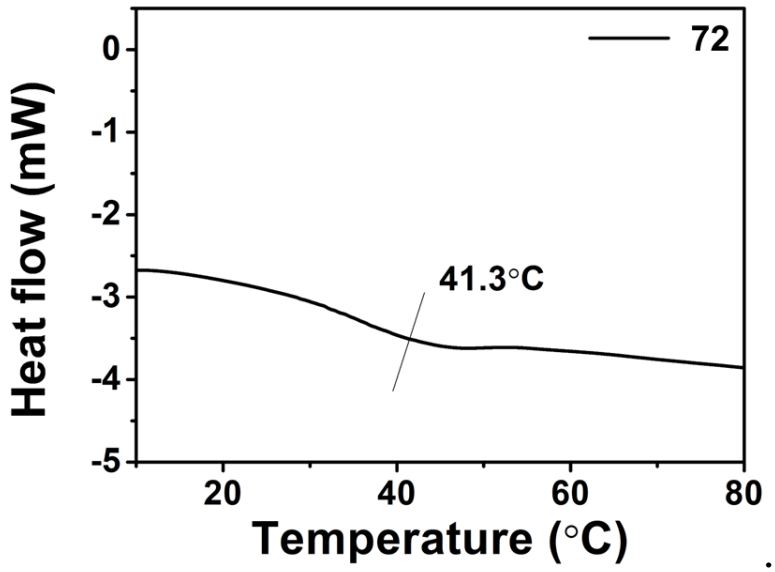

Figure S3. A differential scanning calorimetry (DSC) curve of a thermoplastic (Sample 72) from which T_g is calculated. Sample 72 was printed from HA, IA, NVP, AA, HEAA, and IBOA with monomer ratios of 16, 40, 21, 10, 3, and 10%, respectively.

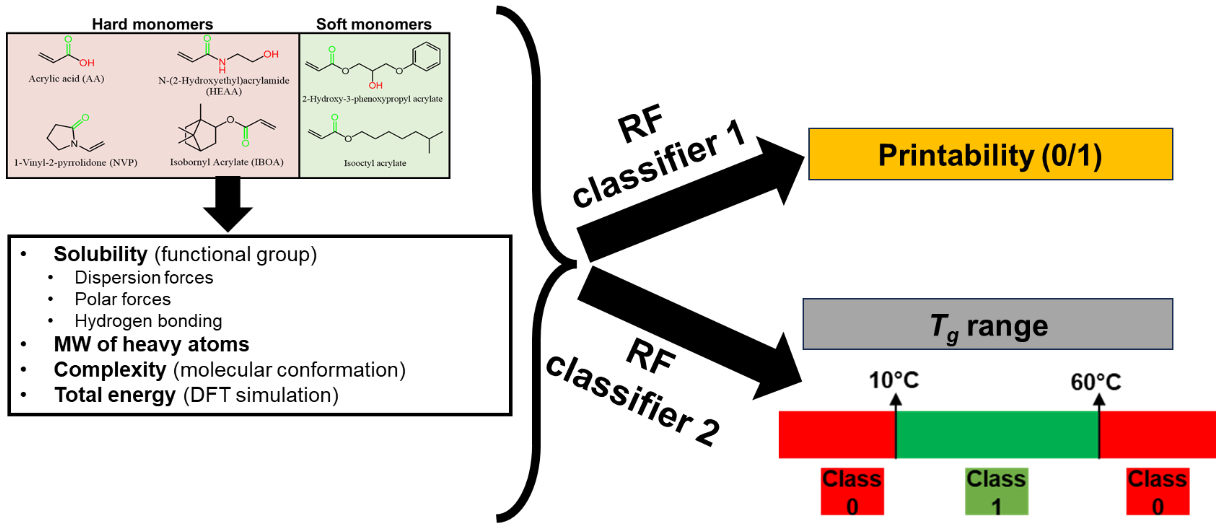

Figure S4. Two RF classifiers employed as constraint models. On the left side are the structures of the monomers along with the extracted descriptors. Those were included with the monomer ratios as the input to train the two RF classifiers. The first one is to predict if a recommended resin formulation leads to a printable thermoplastic. The second one is to predict if the recommended resin formulation realizes a thermoplastic with T_g in a target range (10 °C ≤ T_g ≤ 60 °C).

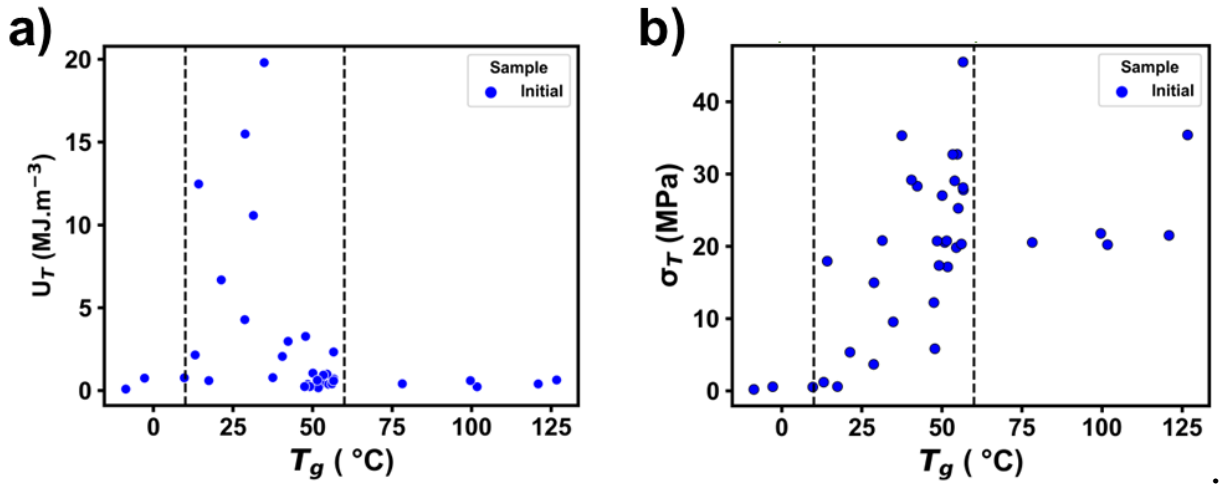

Figure S5. Distribution of **(a)** tensile toughness and **(b)** tensile strength in relation to transition temperature (T_g). Only successfully printed samples are represented.

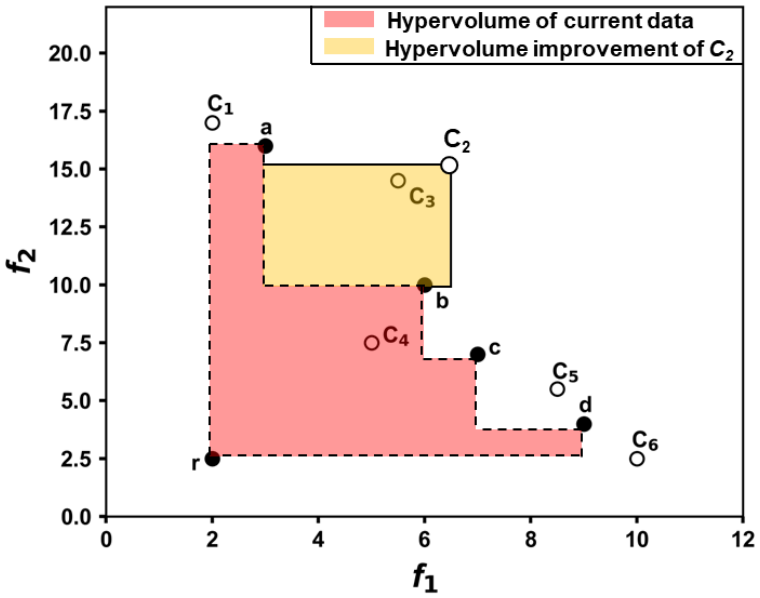

Figure S6. Visualization of the hypervolume improvement. The hypervolume of the existing Pareto front consisting of {**a, b, c, d**} is highlighted in red. In this figure, **r** represents the reference point, chosen as the anti-ideal point relative to the approximate Pareto front of the GP samples. **r** has the lowest $\mathbf{f_1}$ and $\mathbf{f_2}$ with $\mathbf{f_1}$ value from $\mathbf{C_1}$ and $\mathbf{f_2}$ value from $\mathbf{C_6}$. The yellow area indicates the potential contribution of point $\mathbf{C_2}$ added to the new Pareto front. From candidate set ($\mathbf{C_1}$-$\mathbf{C_6}$), $\mathbf{C_2}$ has the highest hypervolume improvement.

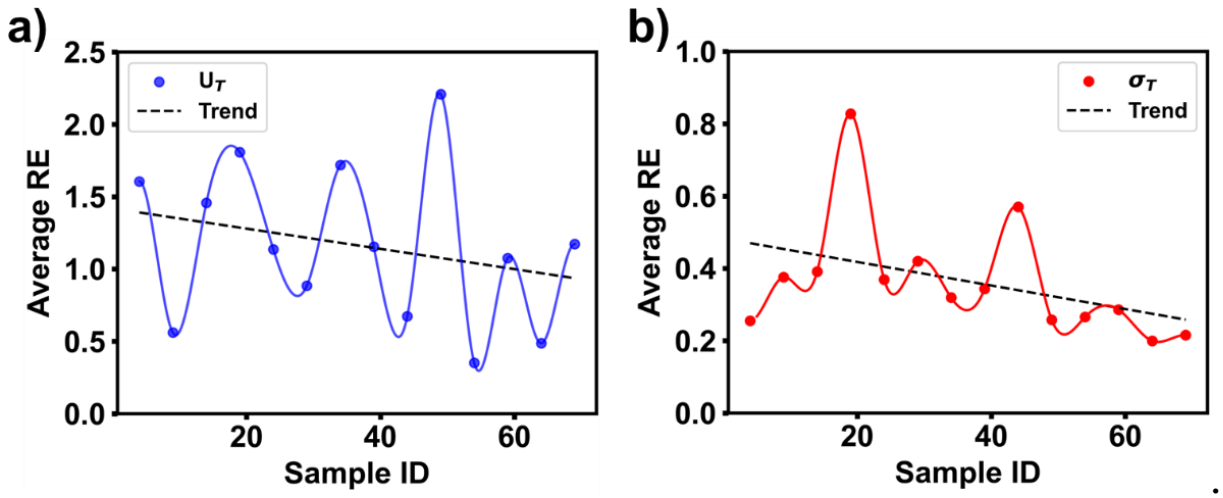

Figure S7. Evolved average relative errors of U_T and σ_T of the evaluated samples as function of the iteration numbers: (a) tensile toughness and (b) tensile strength.

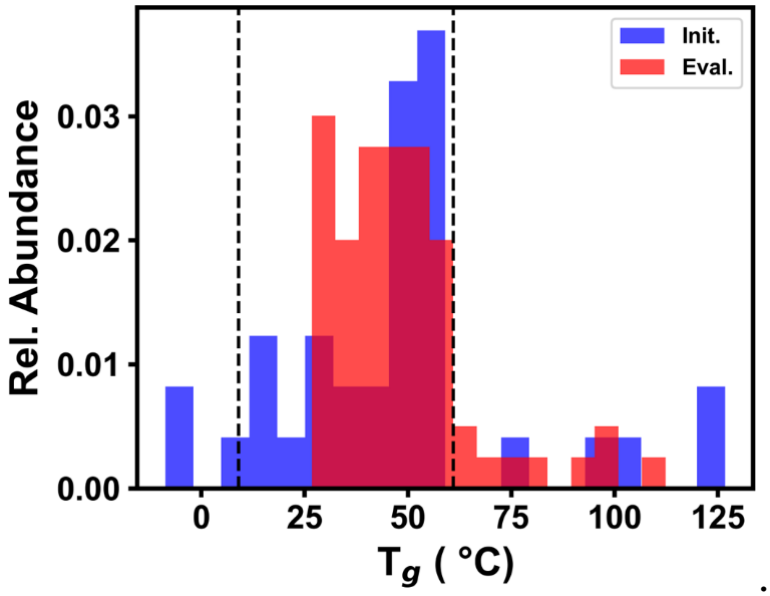

Figure S8. The histogram of T_g. The dashed lines show the acceptable T_g range of 10 and 60 °C, used in the constraint model.

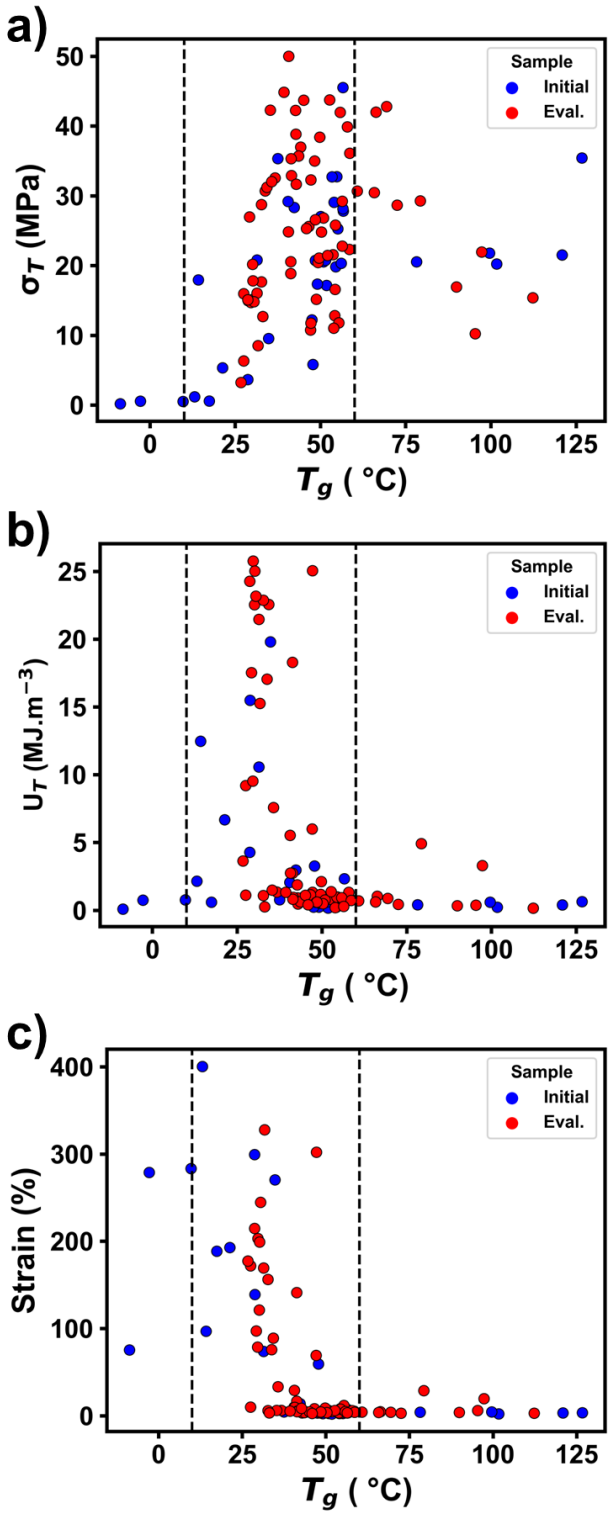

Figure S9. Correlation of mechanical properties over T_g for the initial and evaluated samples.

The dashed lines show the predefined T_g range of 10°C-60 °C.

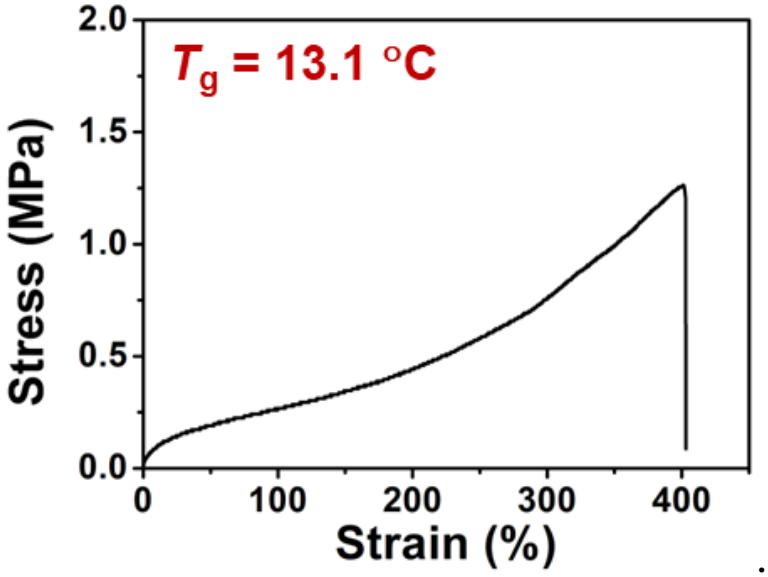

Figure S10. Tensile testing curve of a discovered thermoplastic showing an elastomer behavior.

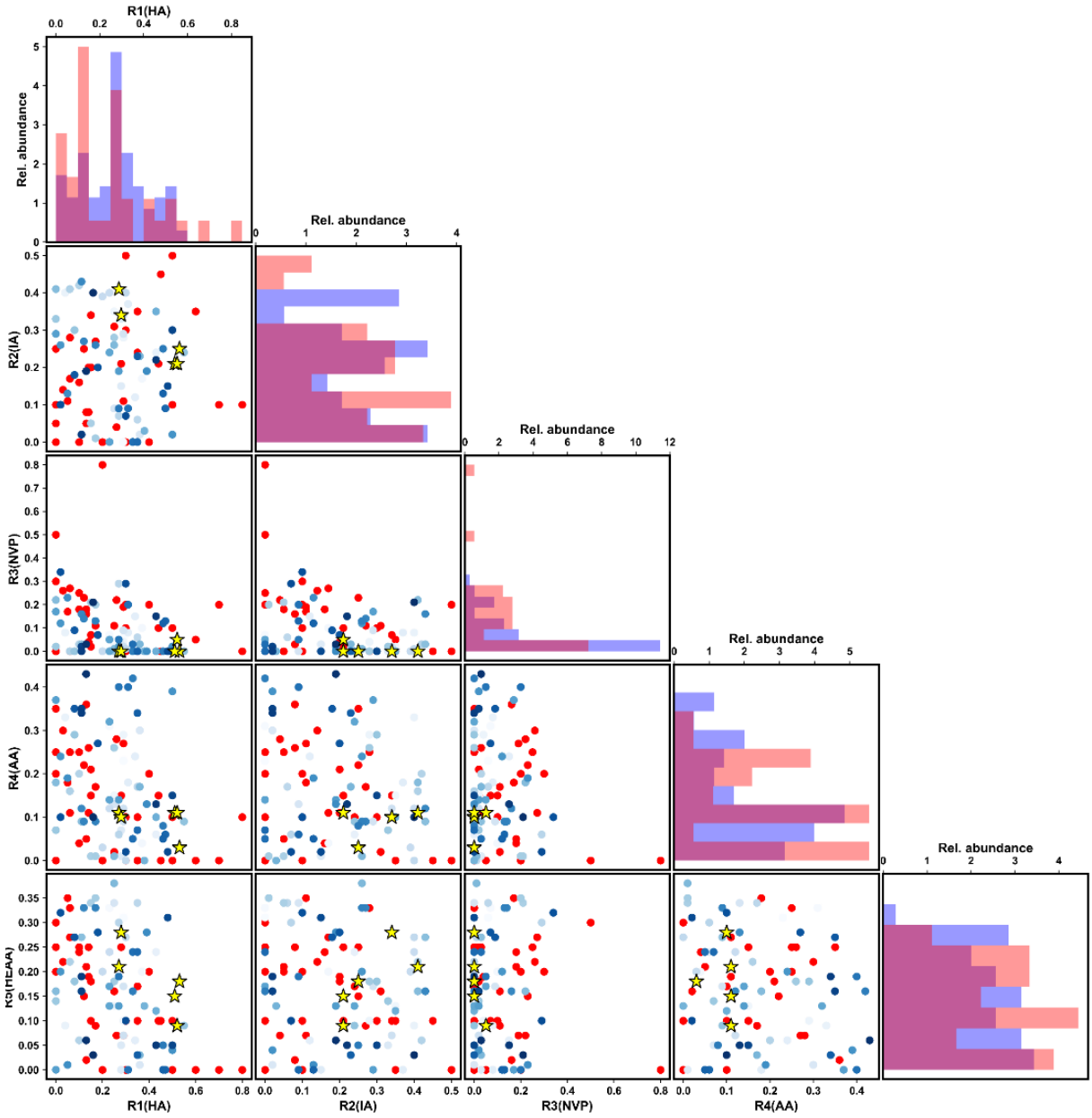

Figure S11. Distribution of ratios for the monomers employed in the optimization algorithm. The initial samples are represented by red points, while the evaluated samples are denoted by blue points. The intensity of the blue points corresponds to the iteration at which the samples were collected, with darker colors indicating later iterations in the algorithm. The samples highlighted with stars are the discovered Pareto frontier points.