



Combining automated microfluidic experimentation with machine learning for efficient polymerization design

Benjamin A. Rizkin¹ , Albert S. Shkolnik, Neil J. Ferraro and Ryan L. Hartman¹

Understanding polymerization reactions has challenges relating to the complexity of the systems, the hazards associated with the reagents, the environmental footprint of the operations and the highly nonlinear topologies of reaction spaces. In this work, we aim to present a new methodology for studying polymerization reactions using machine-learning-assisted automated microchemical reactors. A custom-designed rapidly prototyped microreactor is used in conjunction with automation and *in situ* infrared thermography for efficient, high-speed experimentation to map the reaction space of a zirconocene polymerization catalyst and obtain fundamental kinetic parameters. Chemical waste is decreased by two orders of magnitude and catalytic discovery is reduced from weeks to hours. Bayesian regularization backpropagation is used in conjunction with kinetic modelling to understand the reaction space and resultant technoeconomic topology. Here, we show that efficient microfluidic technology can be coupled with machine-learning algorithms to obtain high-fidelity datasets on a complex chemical reaction.

The advent of synthetic plastics over the past century has drastically reshaped the chemical industry and the world at large. Long-chain α -olefin polymers are most often synthesized using specialized catalysts that are carefully selected both for end-product performance and optimal process economics^{1–3}. The discovery and design of conditions related to these catalysts is a time-consuming, wasteful and expensive process. Single-site catalyst research alone is estimated to have cost five billion dollars by the year 2000, with the entire polymer industry being worth US\$611.9 billion in 2017^{4,5}. Traditional batch-based catalyst discovery consumes litres of solvent per trial and results in large quantities of chemical waste. Additionally, catalysts and the associated activators are expensive and may be hazardous, exemplified by the commonly used activator trimethylaluminium⁶. Homogeneous catalytic polymerizations are an active area of research due to the advantages in control of the polymer product as well as environmental benefits in the manufacturing process^{2,7}. However, problems with dissolved homogeneous catalyst recovery and recycling arise due to difficulties in separating them from the solvent under the mild conditions that would preserve catalyst activity. Major aims are thus to reduce the amount of catalyst, activator and solvent used per experimental trial as well as minimizing the actual number of trials needed to complete a study.

The current state of polymer manufacturing has an environmental impact from the vast amount of energy used to convert petrochemicals into commodity plastics such as poly(ethylene) and poly(propylene). It is estimated by the US Department of Energy that ~6% of all energy produced in the United States is used in the production of polymers⁸. More importantly, ~37% of all global greenhouse gases (GHGs) are created directly or indirectly by polymer manufacturing⁹. In an industry that has shaped the world as we know it and is continuing to expand rapidly, innovative solutions are needed to quickly engineer and understand new production pathways¹⁰. The design of laboratory-scale reactors for quick discovery and the integration of machine learning for robust process understanding are two possible ways to work towards developing

more efficient polymerization catalysts while minimizing environmental harm. Efficient and automated experimentation could reduce experimental times from weeks to hours and reduce chemical waste by several orders of magnitude.

Zirconocene-based alkene polymerization catalysts were derived from the knowledge gained on titanium-based metallocene catalysts originating from the works of Ziegler and Natta¹¹, with significant early advances made by Chien¹², Kaminsky¹³, Brintzinger¹ and Rieger¹⁴. Additional important studies have been carried out since then by various groups investigating both the catalysts and resultant polymers using numerous techniques^{15–17}. Many novel methods have been applied to the study of these catalysts, including quantitative structure–activity relationships¹⁸, mass spectroscopy^{19,20}, cyclic voltammetry²¹ and NMR^{22–24}. However, many of these laboratory techniques are inherently limited by throughput as they rely on batch experimentation, relatively large quantities of reagents, and bulky spectroscopic equipment, greatly limiting the time to market of promising new catalyst compounds. Recently, Reubens et al. have demonstrated, through the use of online size exclusion chromatography, that it is possible to autonomously optimize the molecular weight of photopolymerized acrylate and methacrylate solutions, a significant advancement in autonomous polymerization development²⁵. Other recent advances are summarized in the work of Knox and others²⁶. Overall, designing zirconocene catalytic reactions remains a challenge, as the ratios of reagents, activator composition and temperature play competing roles²⁷. A method in which large amounts of potential candidate reactions can be screened, using *in situ* thermography to quantify catalytic-activity-based exotherms, has not yet been developed.

Microreaction engineering has evolved in the past few decades due to the benefits it offers for controlling chemical reactions. Decreasing the volume of the reactor enables careful manipulation of heat and mass transport physics for desired performance. Additionally, microreactors allow for the integration of novel non-interfering *in situ* and *in operando* spectroscopic methods. Microfluidics have been applied to numerous kinetic and

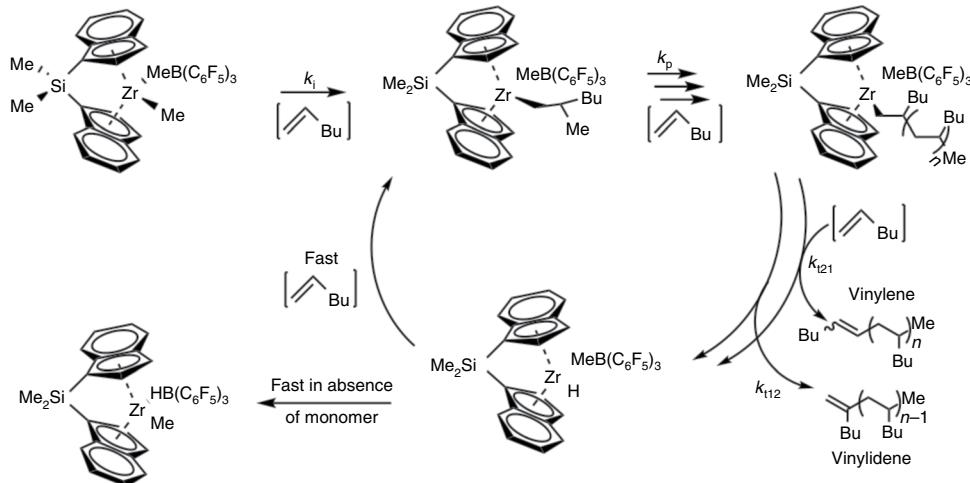


Fig. 1 | Overview of the chemistry. Reaction mechanism for the polymerization of 1-hexene, showing the initiation, propagation and chain transfer steps, based on the work of Moscato and others²². Here, k_i , k_p and k_t are the constant rates of initiation, propagation and termination, respectively.

first-principles research problems, including applications in pharmaceuticals, fine chemicals and petrochemicals^{28–32}. These recent developments were driven by a desire to perform experiments more quickly, maintain better control over the reaction environment, and reduce waste in chemical research. There have been a number of studies in recent years integrating spectroscopic techniques³³. Overall analysis and implementation of microsystems has emerged in pharmaceuticals and fine chemicals through robust understanding and reaction design^{34,35}. The rapid understanding and reaction design of metallocene catalysts has potentially large implications for the environmental footprint of polymer manufacturing. This acceleration in designing catalytic reactions can be accomplished using spectroscopic microreactors with machine intelligence.

A recently discovered catalyst of interest to academia and industry is (SBI)ZrMe₂ (**I**) in conjunction with a B(C₆F₅)₃ (**II**) activator. This combination is active with a broad range of α -olefins, creates polymers with desirable properties, and reduces the reliance on dangerous activators like trimethylaluminium^{22,36–40}. Recently, there have been several papers published about this and related catalyst systems aimed at understanding the kinetics and design of the reactions with the general reaction mechanism shown in Fig. 1^{19,22,37,41,42}. Our findings expand on these works through a semi-automated study of the catalyst's behaviour to quickly estimate the reaction space topology along with estimating technoeconomic operating parameters. Here, traditional data analysis and visualization are supplemented by using artificial neural networks (ANNs) as a non-linear fitting tool to model and predict catalytic behaviour without full knowledge of the model's underlying dependencies and degrees.

This work seeks to address two main challenges: (1) the design of flow-based microsystems for the quick and efficient screening of catalyst for exothermic chemical processes and (2) extraction of the most information possible out of a given set of experiments. The first challenge is addressed through the creation of an integrated continuous-flow microfluidic platform that incorporates pumps, manifolds, controls and analytics into a singular interface that is amenable to automation and integration with analytical techniques. Due to heat and mass transport considerations, flow-based microreactors are challenging to use for such polymerization reactions, necessitating careful design and selection of operating conditions to ensure high data fidelity with minimal transport limitations and safe operation. The second challenge is addressed through the use of rational experimental design to sample an entire experimental space quickly and efficiently while extracting information of interest from a non-invasive and real-time thermal camera. Neural

networks were used as a fitting tool to supplement the analysis by modelling and visualizing the behaviour of the experimental space. Overall the system and process presented address current trends spanning chemical engineering and computer science by integrating microscale reactions with automated experimentation and MI-enhanced process understanding.

Microreactor design

Microfluidic platforms are uniquely suited to the study of olefin polymerizations, as these exothermic reactions take place very quickly, have multiple reaction pathways, require precise control, generate large amounts of chemical waste and use expensive and difficult-to-synthesize catalysts. Microfluidics have already been successfully applied to the research of various polymerizations and other exothermic reactions^{43–46}.

The first step in our design of an intelligent microsystem is an order-of-magnitude estimation of heat and mass transport properties and dimensionless numbers, including the Damköhler number (defined as the ratio of reaction rate to mass transport rate, $Da = kC_0^{n-1} \tau$, where k is the reaction rate constant, C_0 is the initial concentration, n is the reaction order and τ is the residence time), the beta number (the ratio of heat generated to heat removed, $\beta = -(r\Delta H_{rxn}d_F^2)/(4\Delta T_{ad}\kappa)$, where r is the reaction rate, ΔH_{rxn} is the heat of reaction, d_F is the diameter of the channel, ΔT_{ad} is the adiabatic temperature change and κ is the thermal conductivity) and the Reynolds number (the ratio of inertial forces to viscous forces, $Re = \rho VL/\mu$, where ρ is the density of the fluid, V is the velocity, L is the characteristic length and μ is the viscosity). Knowledge of these quantities enables the design of microfluidic devices that offer scalable chemical data mimicking the physics found in industrial-scale processes and enabling visualization. Additionally, this provides context for analysing other properties of interest such as stream mixing, dispersion, heat transfer, mass transfer and the reaction kinetics. Here, the microreactor was designed such that the Damköhler number can be varied between 0.3 and 101, enabling the sampling of both reaction-rate-limited and mass-transport-limited regimes. The beta number varies between $O(10^{-3})$ and $O(10)$ based on the standard heat of polymerization of 1-hexene (**III**), indicating that more heat is generated than removed, enabling thermographic analysis⁴⁷. Finally, the Reynolds number varies between $O(10^{-3})$ and $O(1)$, meaning that the reaction is operating in a laminar flow regime. The final design for the reactor is shown in Fig. 2a. Overall, the analysis and understanding of heat and mass transport characteristics are important to the efficient functioning of the reactor and accurate data collection.

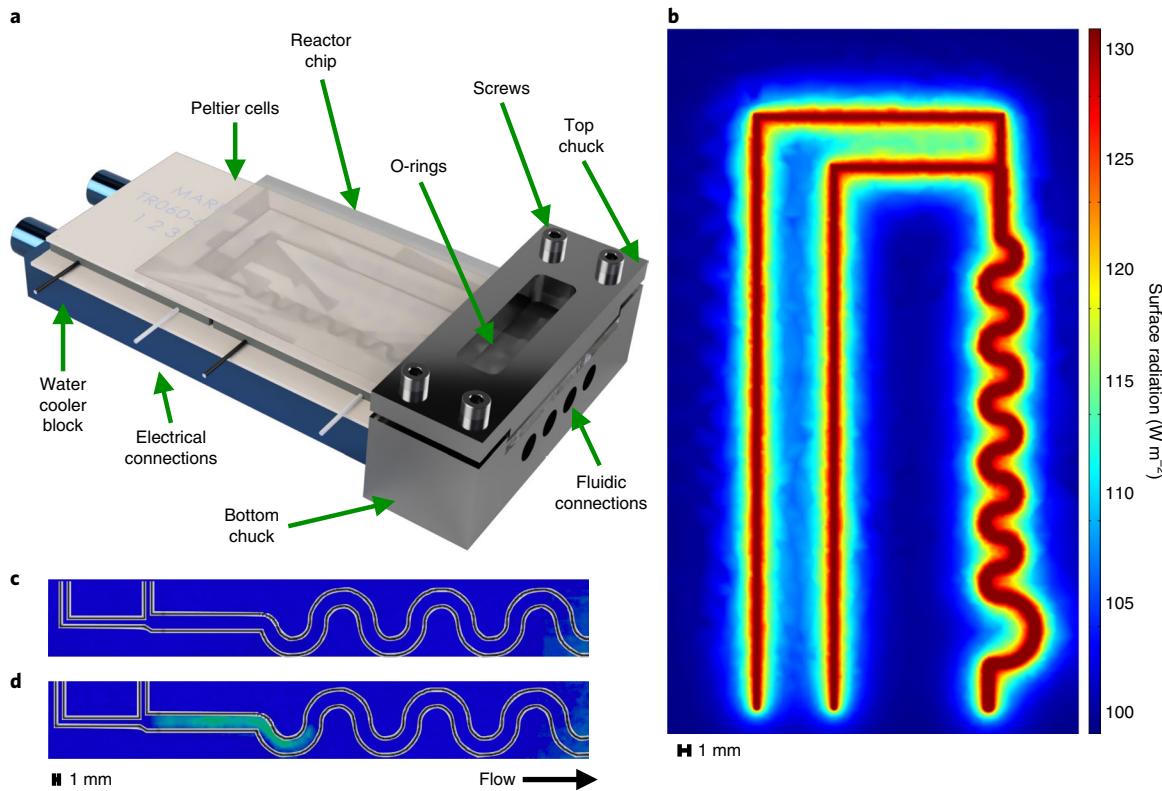


Fig. 2 | Reactor schematic and computational and experimental verification of performance. **a**, Computer-aided design rendering of the assembled reactor with thermal management and fluid delivery systems. The different components include the reactor, compression O-rings, fluid interface connections, Peltier cells for heating/cooling and a liquid cooling block to ensure stable thermal performance. **b**, FEA simulation of the reactor surface measuring the infrared irradiance (W m^{-2}) with a 20 K temperature gradient between the fluid and reactor. **c**, Image of the reactor channel with no flow, taken with the infrared camera. **d**, Image of the reactor with a fully developed reactive flow, taken with the same camera.

System characterization

The design of an intelligent microreactor involves characterization of the desired operation, both computationally using finite element analysis (FEA) and analytically with methods such as residence time distribution studies with chemical tracers. These analyses help with fully characterizing the system and ensuring that analytical methods will yield reliable results. FEA simulations were carried out, verifying a perceptible difference in infrared radiative flux between a warm reactor channel and the ambient reactor. The results of this simulation are presented in Fig. 2b, showing an $\sim 30 \text{ W m}^{-2}$ difference in flux with a 20 K temperature gradient. Experiments were then carried out to verify the thermal performance of the microreactor. Reagent concentrations taken from previous literature were injected into the reactor and the exotherm was observed²². Figure 2c,d presents thermographic images of the reactor, highlighting the development of reactive flow as the reagents mix. The catalyst flows through the far feed channel and the activator flows through the closer one. On contact, a change in radiation is observed, as indicated by the lighter blue colour. It is further observed that the reaction is initiated almost instantaneously when the reagents mix, with the strongest thermal signature in the first centimetre of the channel.

Finally, the microsystem design involves its integration with analytical methods including both in and ex situ measurement techniques. Accurate and precise control along with a diversity of measurable variables are critical to creating an accurate digital twin computer model of the complex polymerization. A challenge presented by a flow-based testing platform is the inclusion of calibrated pumps and manifolds not present in a traditional batch system. Other critical considerations are temperature mapping and homogeneity,

which are uniquely difficult in typical flow reactors but are simplified in microsystems. The successful analysis, testing and integration of all these aspects results in a system that provides relevant data quickly and accurately. An overview of the experimental system is summarized in Fig. 3.

Algorithmic concept

The current work seeks to develop a methodology for the testing, design and general understanding of one class of zirconocene-based catalyst by using an ANN fitting. The algorithm, used as a form of supervised machine learning (ML), helps aid in our understanding of the complex kinetics and reaction design for a homogeneous catalysed polymerization. Fitting of traditional numerical models involves understanding the degree of dependence of various independent variables (linear, polynomial, natural exponential and so on), while an ANN can be trained and adapted without complete knowledge of the system. It is, however, still necessary to have a rough understanding of the physical phenomenon to choose activation functions and a number of hidden-layer neurons that are plausible. Recently, we demonstrated that ANNs can be used as a tool for modelling and understanding complex catalytic pathways for polymerization reactions from a first-principles *in silico* approach⁴⁸. A design of experiments (DoE) strategy employing a Latin hypercube design is used to accurately and randomly sample the multidimensional experimental space⁴⁹. The integration of these steps together compounds the energy, cost and environmental impact savings of using a smaller lab space, with fewer trials, with a fraction of the time and labour. This enables faster data collection and development of models, while reducing the time and amount of chemical waste generated. This methodology serves as a proof of concept for

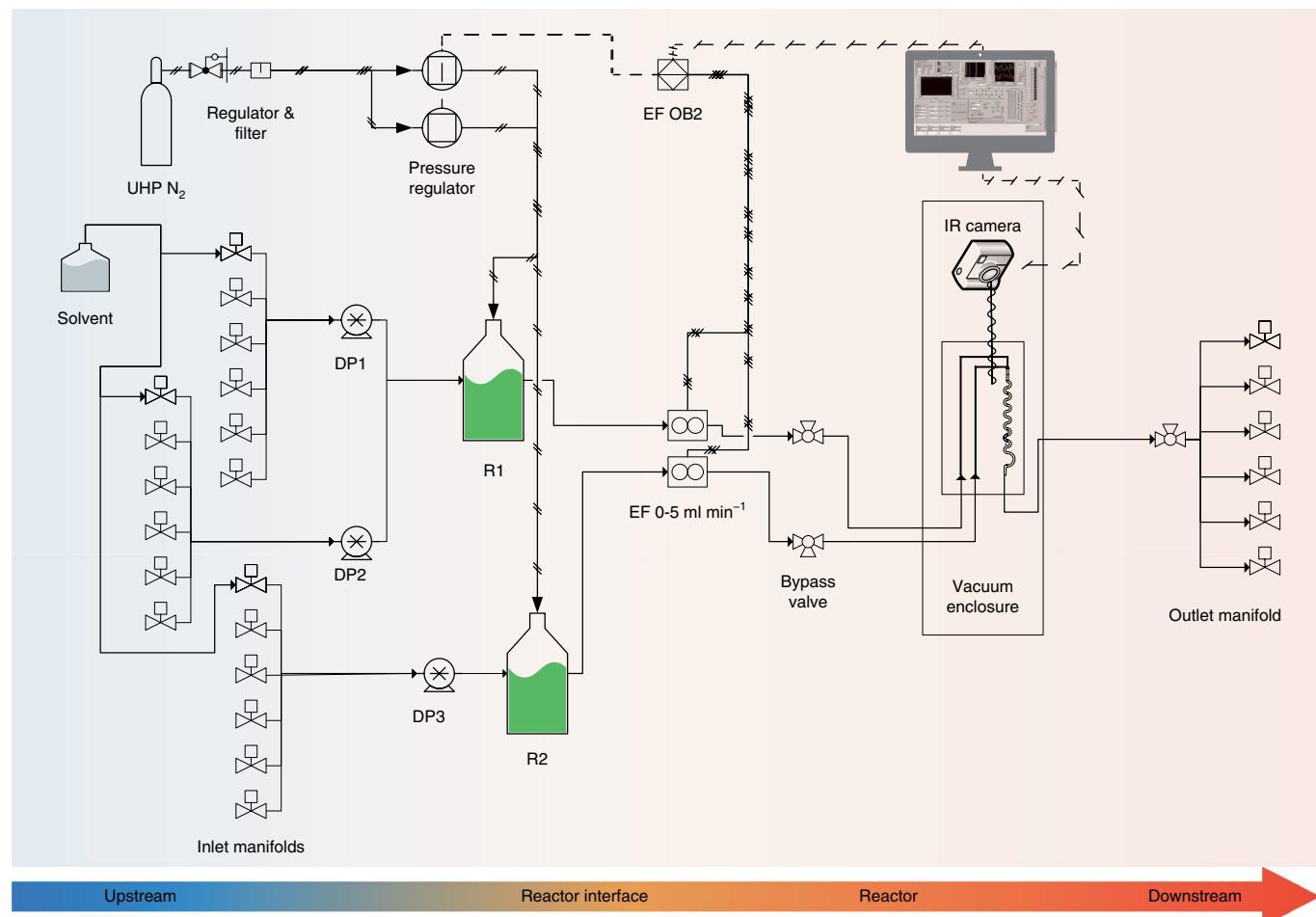


Fig. 3 | Process flow diagram for an automated thermographic microreactor system used in the understanding of metallocene catalysts. Left to right: input manifolds with solvent tank, nitrogen supply, electromagnetic dosing pumps (DP) used for mixing, reagent reservoirs (R1 and R2), pressure controller (EF OB2), mass flow meters (EF 0-5 ml min⁻¹), bypass valves, infrared (IR) camera, vacuum enclosure (used to reduce measurement noise) and outlet manifold. UHP, ultrahigh purity.

using DoE algorithms with ANN fitting and spectroscopic microfluidics to quickly gain process understanding.

The concept presented herein is that a semi-autonomous reaction system can perform as a machine-assisted chemist to help understand the complex reaction space for a homogeneous metallocene catalyst. The first component of this comprises the control and data interpretation systems that perform the experiments, gather the data and generate the fitted ANN models. An overview of the system used is provided in the top panel of Fig. 4. The process is controlled by a combination of MATLAB and LabVIEW, as each offers certain advantages. MATLAB allows for the development of advanced computational algorithms and includes a robust ML toolbox with different training algorithms. LabVIEW offers a real-time control environment with independent loops running to provide control of the system, including interfacing with external devices. Reagent mixing is provided by electromagnetic diaphragm pumps, flow through the reactor is established by the use of a pressure controller and thermal control is established through Peltier elements. A full description of the system is provided in the Supplementary Methods and page 3 of the Supplementary Information.

Data from the infrared camera were collected using the native camera software at a maximum speed of 24 frames per second and recorded across the reaction zone into a database. Reaction exotherm data were used to interpret the catalytic productivity,

mainly the grams of polymer produced per (mole catalyst × mole monomer × hour). Catalytic productivity is an important metric for polymerization catalyst design and is used when discovering new catalysts, understanding existing catalysts and designing industrial plants. Due to the environmental concerns associated with catalyst production and reagent recycling, including GHG emissions and chemical waste, it is important to adjust catalytic performance to meet optimistic goals and government regulations.

By interpolating the data from a minimal number of experiments using a quick and efficient fitting algorithm, it is possible to visualize the full range of the experimental topology, which could not be achieved using traditional trials.

Automated experimentation

Experimentation was conducted in such a way as to establish an understanding of the reaction space without the need for extensive trials. The reagents were manually prepared in an inert environment glovebox and connected to the experimental manifolds. An automated routine was used to establish control, clean the reactor and perform experiments (see flowchart presented in the bottom panel of Fig. 4). The purpose of the algorithm was to efficiently perform experiments in an automated fashion. The next experimental point was selected from a specified monomer, catalyst, activator concentration bank based on the recommendation of a Latin hypercube DoE algorithm. This allows for a nearly random distribution to be

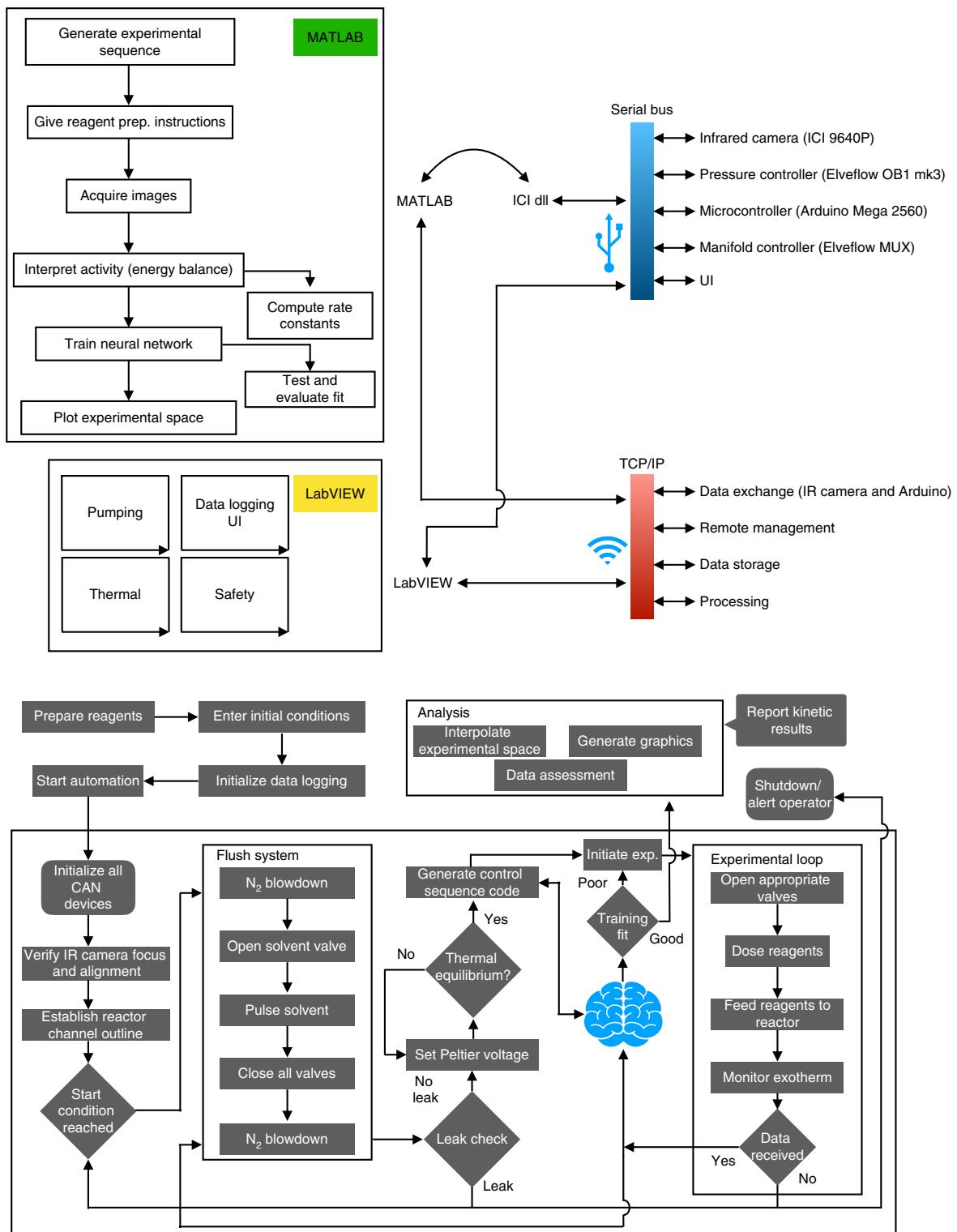


Fig. 4 | Flowcharts for experimental control and data handling. Top: diagram of the system architecture of the high-throughput experimentation system with the functions of MATLAB and LabVIEW broken down by control loops and communication protocols. Bottom: flowchart of how experiments are performed, from reagent preparation to a report of the kinetic results. Analysis was performed after the experimental sequence was complete. CAN, controller area network; UI, user interface.

sampled across the experimental space, enhancing the robustness of the resultant model. The experimental system consisted of the reagent mixing and storage equipment, thermal control, the reactor and associated control systems (see experimental flowchart in Fig. 3). The reactor and thermal camera were contained within a

vacuum enclosure to reduce the effect of atmospheric interference, as air and water vapour contained in it would introduce noise into the data. The reactor used in this study was fabricated by photo-polymerization stereolithography and bonded to an infrared-transparent fluoropolymer film (for an overview of the reactor assembly

and dimensions see Fig. 2a; for full information on fabrication see Supplementary Information).

Overall the process and system were designed to handle several challenges, including maintaining an inert environment, ensuring safety, performing experiments automatically and producing scalable data. The chosen catalyst and activator molecules are highly sensitive to moisture and oxygen, and must therefore be handled in such a way that they never come into contact with the atmosphere. Additionally, the solvent and monomer used present safety concerns as they are both toxic and highly flammable, producing vapours that may be explosive. The next challenge in the system design was the integration of robust process automation. The experiment needed to integrate a thermal camera, analytics and fluid handling seamlessly and autonomously. This was accomplished through the combined use of MATLAB and LabVIEW with the integration of an open-source Arduino microcontroller for manifold control. The final challenge in system design was the microreactor itself. The reactor was designed from the ‘bottom up’ to ensure optimal fluidic and thermal performance supported with both computational and experimental evidence. Overall, the system was designed in such a way as to allow for automated experimentation using sensitive chemistries with *in situ* analytics integrated with the microreactor platform.

Data collection and analysis

On verifying reactor performance, a Latin hypercube sampling algorithm was used to test the effects of the concentrations of monomer, activator and catalyst, with 29 initial experiments and eight additional trials performed at the random temperature setpoints of 16, 11, 8, 6, 42, 46, 61 and 79 °C. An overview of the experimental concentration used for network training is shown in Fig. 5a with all thermal trials using a low catalyst loading and a high concentration of activator and monomer. It was determined that decoupling anisothermal trials from the anisotropic ones would yield the best quality of data because the experimental effects are not convoluted. If both sets of changing conditions were incorporated into the original Latin hypercube design, it would be difficult to separate which effects are caused by changing concentration and which by changing temperature, resulting in questionable statistical interpretations of the data. In total, these experiments generated (with reactor rinses) less than 30 g of chemical waste, a two to three order of magnitude reduction from traditional experiments. The footprint of the experiment and all associated support equipment was contained to a nine-square-foot area in a fume hood.

The reactor performed as designed, generating a series of exotherms with both dimensional and temporal resolution within the reactor. These exotherms were then used in conjunction with an energy balance for reacting laminar flow in a microchannel to glean information both on the catalytic productivity (a technoeconomic metric) and reaction kinetics (necessary for reactor design and scale-up). Figure 5b shows how the catalytic productivity changes with changing reaction conditions. Finally, a series of ANNs were generated based on the experimental results in an effort to model the catalytic productivity throughout the experimental space. The results of this training in the form of error percentage over the test dataset (15% of the experimental data) are provided in Supplementary Fig. 3, ranging over both training method and the number of hidden-layer neurons. Cross-validation was performed while taking care to keep the training and test data distinct. To test the fit quality of the chosen network architecture, training was repeated five times using random indices for training versus test data; detailed results are provided in Supplementary Table 2. A Bayesian regularization backpropagation training algorithm with a feed-forward network consisting of an input layer, two hidden layers with nine neurons each and an output layer resulted in the best quality of fit over the experimental space and was retrained after

hyperparameter tuning using MATLAB’s automated routines for selecting the training rate parameters and manual selection of hidden layer/neuron numbers. The final network used for the study had a mean square error (MSE) losses of 1.5×10^{-2} and 6.0×10^{-5} over the testing and training datasets, respectively. The Bayesian regularization backpropagation algorithm was predicted to have the best representation because it leverages the benefits of both the Levenberg–Marquardt optimization algorithm with robust regularization to enhance prediction stability. Neural networks were used as opposed to polynomial fits due to the ease of adaptability to new experiments and input/output parameters. In the case presented here, the inputs consisted of concentration parameters and temperature, but the architecture of the network can be easily adapted to include other inputs including flow parameters, chemical metrics such as Lewis acidity and adaptations for different reactants. The network can also be updated using reinforcement learning to adapt to different α -olefin polymerization reactions. The network was designed to output the catalytic activity, a technoeconomic parameter used to determine how much polymer is made per quanta of catalyst per quanta of time. Bayesian regularization and normalization were used to provide a stable and scalable fit across a range of conditions. Raw data from the thermal camera were processed through ICI’s proprietary libraries, which provided non-uniformity correction and autocalibration. For additional details on network training see the Supplementary Information.

Kinetic and ANN analysis

The first step in kinetic analysis was determining the catalytic productivity and kinetic rate constants for the various experiments. This was performed using automated MATLAB code, which extracted the exotherm from the experimental database and converted it into meaningful data through a heat balance. The raw experimental results are shown in Fig. 5b, where the experimental exotherms are compared with the neural network prediction over all trials. The average percent error of the fit is under 0.5%, with notable deviation being observed in the low-temperature and first high-temperature trials. It is hypothesized that this deviation was caused by adsorption of gelled polymer product from the low-temperature experiment to the reactor walls, which then became desorbed in the first high-temperature trial. Following this deviation, the AI predictions once again lie in near-perfect agreement with the experimental results. Finally Fig. 5c,d shows the kinetic rate constants of chain initiation and propagation, respectively. From this data it can be seen how the rate constants change with varying concentrations and temperature (plots 1–29 and 30–37, respectively).

As the final component of the investigation, a neural network was trained and used to compute the catalytic productivity over a broad range of plausible points from the literature within the reaction space. This methodology presents a way to efficiently visualize the reaction space topology for complex catalytic cycles with a minimum number of experiments. The results of the ANN-based computation are shown in Fig. 6, where the first row shows the reaction space over a range of activator concentrations, the second row over monomer concentrations and the third row over catalyst concentrations. It is observed that the catalytic productivity decreases with increasing catalyst and monomer concentrations, consistent with the mathematical interpretation of the term. This is also consistent with the reaction mechanism, because as the concentration of monomer goes up, the prevalence of vinylene and vinylidene termination mechanisms may increase, and at high catalyst concentrations the activity per gram of catalyst is inherently lower due to the inverse relationship. There is also some nonlinear behaviour at mid-range concentrations of monomer with high activator concentrations, perhaps due to competing branching and termination steps. It is also observed that the productivity tends to increase with increasing activator concentration, consistent with its purpose in

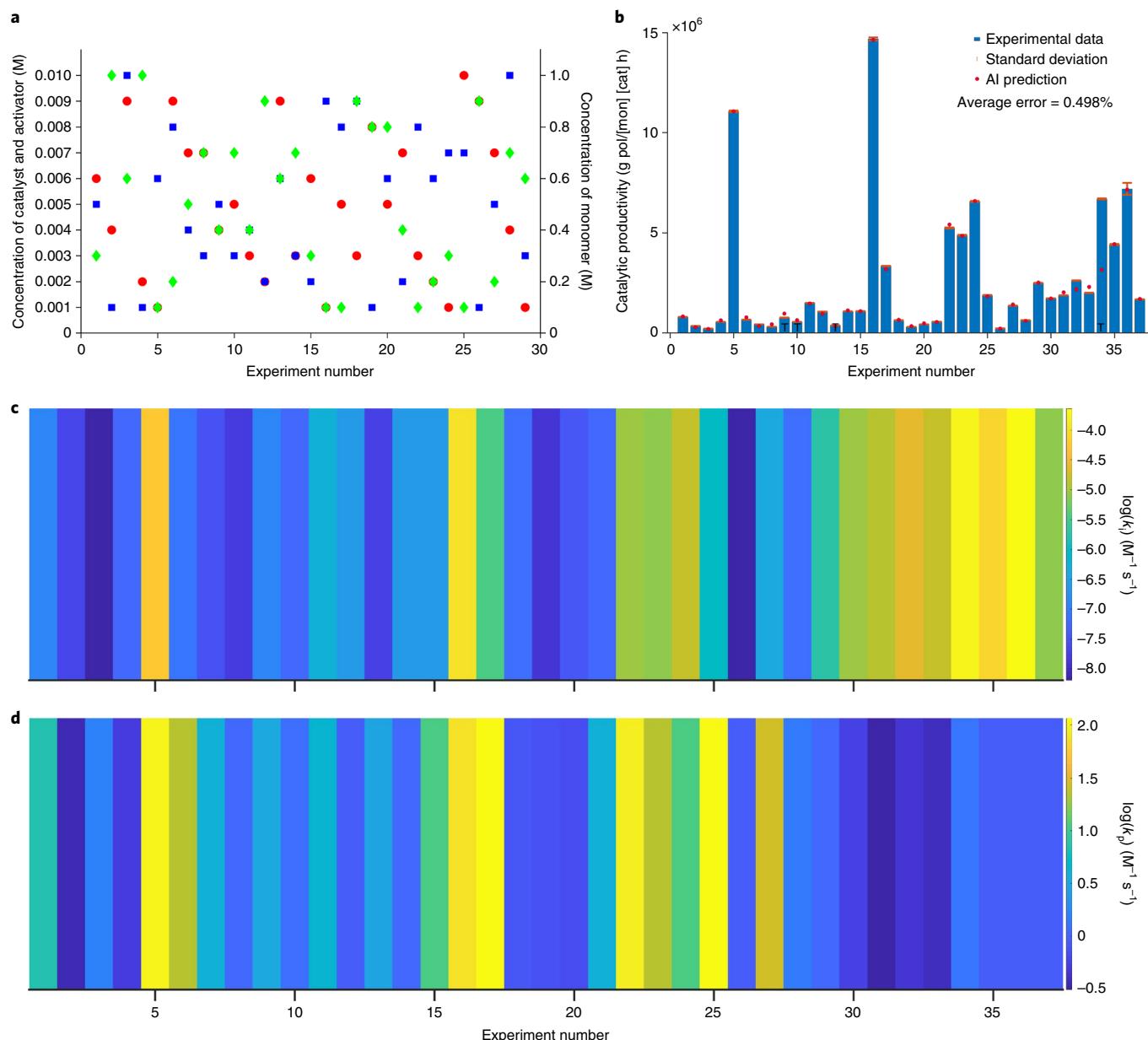


Fig. 5 | Experimental results for catalytic productivity and kinetic rate constants. **a**, Experimental concentrations of the various reagents over the 29 training trials. Red circles represent catalyst concentration, blue squares represent activator concentration and green diamonds represent monomer concentration. **b**, Catalytic productivity ($\text{g polymer}/(\text{mol}(\text{Zr}) \times \text{mol (1-hexene)} \times \text{h})$) of the $\text{SBI}(\text{Zr})\text{Me}_2$ catalyst for the various experiments performed. Blue bars represent catalytic activity and red circles represent the AI prediction. The error bars represent the differences in analysed exotherm between frames in flow over the course of each experiment (standard deviation over the course of the experiment). Trials 9, 10, 13 and 34 were used for cross-validation (indicated by the 'T' markings in the figure) and different random combinations were used for additional testing, as shown in Supplementary Table 2. **c**, Natural logarithm of the kinetic rate constant for chain initiation as computed through a nonlinear fit over the various experiments. **d**, The same as in **c**, but for the chain propagation rate constant.

the reaction mixture. This presents an interesting technoeconomic problem, which is investigated in the bottom half of Fig. 5, showing the catalytic productivity as a function of monomer concentration and temperature. These data represent a snapshot of 1,000 random concentration combinations over a range of temperatures, giving a randomized decision matrix from which an optimal value can be inferred given desired parameters. Over the experimental space presented here, this value was determined to be a catalyst concentration of 54 mM, a nearly equimolar activator concentration of 56 mM, a monomer concentration of 1.04 M and an operating temperature between 40 and 78 °C. Overall, the ANN analysis allows

for the generation of a very large dataset, which can then be used to make decisions about operating points as part of the catalyst discovery process.

Conclusions

Semi-autonomous microfluidic platforms integrated with *in situ* thermography may be used in conjunction with fitted regularized and normalized neural networks to better understand catalytic cycles for homogeneous polymerization reactions. It has been shown that a metallocene-catalysed polymerization can be performed in a machine-assisted microreactor, and the results can be

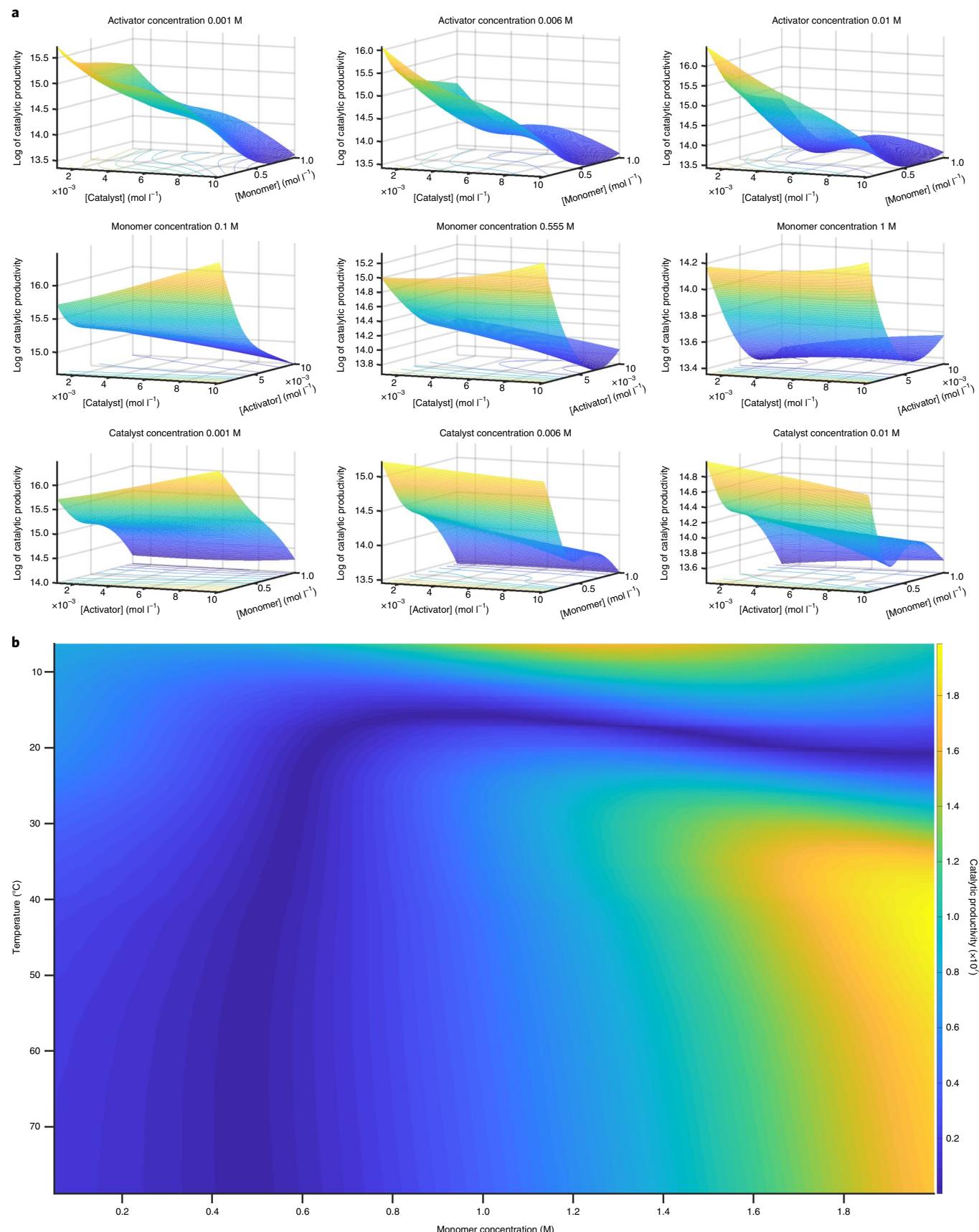


Fig. 6 | ANN results for catalytic productivity over various concentrations and temperatures. a, Reaction space topology simulated from the ANN, trained based on experimental data. The top row shows how the reaction space changes with changing activator concentration, the middle row the change with monomer concentration and the bottom row the change with catalyst concentration. **b,** Catalytic productivity plotted as a function of monomer concentration and temperature.

quantified using an infrared camera, interpreted using an energy balance and used to train an ANN. The ANN may then, in turn, be used to visualize the reaction space with a higher resolution than would otherwise be possible using traditional experimentation. The behaviour of the rate constants of initiation and propagation was also investigated, along with the behaviour of the catalyst at varying temperatures. The reaction space for an industrially and academically relevant homogeneous metallocene polymerization system was investigated and the behaviour towards varying concentrations of reactants was visualized. The combination of these data provides a useful tool for both scientific and technoeconomic analysis of catalyst systems. It was also determined that the Bayesian regularization backpropagation algorithm with nine hidden-layer neurons, used in conjunction with normalization, provides the best quality of fit. It is hypothesized that this result is due to the fact that in Bayesian regularization the training algorithm determines a combination of squared errors and weights that generalizes well. This enables construction of a network that is resilient across a wide range of conditions.

Overall, this study has demonstrated that high-throughput microfluidics can be aided by ML algorithms for the investigation of complex chemical reactions. This opens doors to new types of research, primarily the 'robotic chemist', increasing throughput, data fidelity and the efficiency of experimental campaigns. Future work could incorporate statistically based Monte Carlo design algorithms to aid in understanding the relationships between parameters. Depending on the use of the experimental results and knowledge of the underlying model, other numerical fitting techniques may also be used. Performing reactions at the microscale with automation reduces the amount of energy input and chemical waste generated, while also increasing safety because any failure in the reaction system is small and contained. Finally, the study aimed to contribute knowledge to online learning for complex systems, as the training methodology could be applied to other chemical, mechanical and electric/electrical systems.

In the future, the methodology presented here may also be used to investigate other catalyst systems where a thermal signature may be expected, including other polymerizations, exothermic biochemical transformations and catalytic breakdown of harmful emissions including SO_x and NO_x compounds. This experimentation lays the groundwork for the ability to predict optimal catalytic performance conditions autonomously. Optimization of industrially relevant catalysts or catalyst/activator pairs may be accomplished by combining the neural network analysis with optimization techniques such as technoeconomic model minimization or genetic algorithms. Additionally, the high-resolution ANNs can be combined with various topology/structural analysis techniques to design new experiments. This AI methodology and the design of machine-assisted microreactors reduces the amount of chemical waste and energy input while enhancing the time to actionable data and the resolution of reaction space topology.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The datasets generated and/or analysed during the current study are available in the Zenodo repository, <http://doi.org/110.5281/zendodo.3706730>[50].

Code availability

The code used for data analysis in the current study is publicly available in the Zenodo repository, <http://doi.org/110.5281/zendodo.3706730>[50]. The code used for experimental control is available in the Zenodo repository, <http://doi.org/zendodo.3706734>[51].

This code is not publicly available due to inclusion of third party code but is available upon reasonable request from the corresponding author.

Received: 17 September 2019; Accepted: 11 March 2020;
Published online: 6 April 2020

References

- Bräntzinger, H. H., Fischer, D., Mülhaupt, R., Rieger, B. & Waymouth, R. M. Stereospecific olefin polymerization with chiral metallocene catalysts. *Angew. Chem. Int. Ed.* **34**, 1143–1170 (1995).
- Shamiri, A. et al. The influence of Ziegler–Natta and metallocene catalysts on polyolefin structure, properties and processing ability. *Materials* **7**, 5069–5108 (2014).
- Kaminsky, W. Highly active metallocene catalysts for olefin polymerization. *J. Chem. Soc. Dalton Trans.* **1998**, 1413–1418 (1998).
- Sinclair, K. B. Future trends in polyolefin materials. *Macromol. Symp.* **173**, 237–261 (2001).
- Plastics and Polymers Global Market Briefing 2018* (The Business Research Company, 2018).
- Sumerin, V. & Thorman, J. Ziegler–Natta catalyst and preparation thereof. US patent 10,118,977 (2018).
- Kesti, M. R., Coates, G. W. & Waymouth, R. M. Homogeneous Ziegler–Natta polymerization of functionalized monomers catalyzed by cationic group IV metallocenes. *J. Am. Chem. Soc.* **114**, 9679–9680 (1992).
- Society of the Plastics Industry, US Department of Energy *Improving Energy Efficiency at US Plastics Manufacturing Plants* 40 (US DOE, 2005).
- Worrell, E., Bernstein, L., Roy, J., Price, L. & Harnisch, J. Industrial energy efficiency and climate change mitigation. *Energy Efficiency* **2**, 109–123 (2009).
- Khrapko, D., Schlüter, B. A., Rommel, B., Rosano, M. & Hesselbach, J. Energy demand and efficiency measures in polymer processing: comparison between temperate and Mediterranean operating plants. *Int. J. Energy Environ. Eng.* **7**, 225–233 (2016).
- Sinn, H. & Kaminsky, W. Ziegler–Natta catalysis. *Adv. Organomet. Chem.* **18**, 99–149 (1980).
- Chien, J. C. W. & Wang, B.-P. Metallocene–methylaluminoxane catalysts for olefin polymerizations. IV. Active site determinations and limitation of the ¹⁴CO radiolabeling technique. *J. Polym. Sci. A* **27**, 1539–1557 (1989).
- Kaminsky, W. (ed.) *Metalorganic Catalysts for Synthesis and Polymerization* (Springer, 2011).
- Rieger, B., Jany, G., Steimann, M. & Fawzi, R. Synthesis of ethylene bridged biscyclopentadiene ligand precursor compounds and some of their ansa-zirconocene derivatives chiral epoxides: a synthetic strategy of high variability. *Z. Naturforsch. B Chem. Sci.* **49**, 451–458 (1994).
- Kolthammer, B. W. S., Mangold, D. J. & Gifford, D. R. Polymerization kinetics of octene-1 catalyzed by metallocene methylaluminoxane investigated with attenuated total reflectance Fourier transform infrared (ATR-FT-IR) spectroscopy. *J. Polym. Sci. A* **30**, 1017–1026 (1992).
- Charpentier, P. A., Zhu, S., Hamielec, A. E. & Brook, M. A. Continuous solution polymerization of ethylene using metallocene catalyst system, zirconocene dichloride/methylaluminoxane/trimethylaluminum. *Ind. Eng. Chem. Res.* **36**, 5074–5082 (1997).
- D'Agnillo, L., Soares, J. B. P. & Penlidis, A. Effect of operating conditions on the molecular weight distribution of polyethylene synthesized by soluble metallocene/methylaluminoxane catalysts. *Macromol. Chem. Phys.* **199**, 955–962 (1998).
- Martínez, S., Cruz, V. L., Ramos, J. & Martínez-Salazar, J. Polymerization activity prediction of zirconocene single-site catalysts using 3D quantitative structure–activity relationship modeling. *Organometallics* **31**, 1673–1679 (2012).
- Moscato, B. M., Zhu, B. & Landis, C. R. GPC and ESI-MS analysis of labeled poly(1-hexene): rapid determination of initiated site counts during catalytic alkene polymerization reactions. *J. Am. Chem. Soc.* **132**, 14352–14354 (2010).
- Santos, L. S. & Metzger, J. O. Study of homogeneously catalyzed Ziegler–Natta polymerization of ethene by ESI-MS. *Angew. Chem. Int. Ed.* **45**, 977–981 (2006).
- Silveira, F., De Sá, D. S., Da Rocha, Z. N. & Dos Santos, J. H. Z. Metallocene combinations in ethylene polymerization: a cyclic and differential pulse voltammetry study. *Macromol. React. Eng.* **2**, 253–264 (2008).
- Moscato, B. M., Zhu, B. & Landis, C. R. Mechanistic investigations into the behavior of a labeled zirconocene polymerization catalyst. *Organometallics* **31**, 2097–2107 (2012).
- Gonzalez-Ruiz, R. A., Quevedo-Sánchez, B., Laurence, R. L., Henson, M. A. & Bryan Coughlin, E. Kinetic modeling of slurry propylene polymerization using *rac*-ET(Ind)₂ZrCl₂/MAO. *AIChE J.* **52**, 1824–1835 (2006).
- Christianson, M. D., Tan, E. H. P. & Landis, C. R. Stopped-flow NMR: determining the kinetics of [*rac*-(C₂H₄(1-indenyl)₂)ZrMe][MeB(C₆F₅)₃]–catalyzed dopolymerization of 1-hexene by direct observation. *J. Am. Chem. Soc.* **132**, 11461–11463 (2010).

25. Rubens, M., Vrijen, J. H., Laun, J. & Junkers, T. Precise polymer synthesis by autonomous self-optimizing flow reactors. *Angew. Chem. Int. Ed.* **58**, 3183–3187 (2019).

26. Knox, S. T. & Warren, N. J. Enabling technologies in polymer synthesis: accessing a new design space for advanced polymer materials. *React. Chem. Eng.* **5**, 405–423 (2020).

27. Kaminsky, W. Zirconocene catalysts for olefin polymerization. *Catal. Today* **20**, 257–271 (1994).

28. Jensen, K. Flow chemistry—microreaction technology comes of age. *AIChE J.* **63**, 858–869 (2017).

29. Heider, P. L. et al. Development of a multi-step synthesis and workup sequence for an integrated, continuous manufacturing process of a pharmaceutical. *Org. Process Res. Dev.* **18**, 402–409 (2014).

30. Hartman, R. L., Naber, J. R., Buchwald, S. L. & Jensen, K. F. Multistep microchemical synthesis enabled by microfluidic distillation. *Angew. Chem. Int. Ed.* **49**, 899–903 (2010).

31. Kim, J. O. et al. A monolithic and flexible fluoropolymer film microreactor for organic synthesis applications. *Lab Chip* **14**, 4270–4276 (2014).

32. Hu, C., Morris, J. E. & Hartman, R. L. Microfluidic investigation of the deposition of asphaltenes in porous media. *Lab Chip* **14**, 2014–2022 (2014).

33. Rizkin, B. A., Popovic, F. G. & Hartman, R. L. Spectroscopic microreactors for heterogeneous catalysis. *J. Vac. Sci. Technol. A* **37**, 050801 (2019).

34. Gromski, P. S., Granda, J. M. & Cronin, L. Universal chemical synthesis and discovery with ‘The Chemputer’. *Trends Chem.* **1**–9 (2019); <https://doi.org/10.1016/j.trechm.2019.07.004>

35. Coley, C. W. et al. A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* **365**, eaax1566 (2019).

36. Theurkauff, G., Bondon, A., Dorcet, V., Carpentier, J. F. & Kirillov, E. Heterobi- and -trimetallic ion pairs of zirconocene-based isoselective olefin polymerization catalysts with AlMe_3 . *Angew. Chem. Int. Ed.* **54**, 6343–6346 (2015).

37. Song, F., Cannon, R. D. & Bochmann, M. Zirconocene-catalyzed propene polymerization: a quenched-flow kinetic study. *J. Am. Chem. Soc.* **125**, 7641–7653 (2003).

38. Christopher, J. N., Diamond, G. M., Jordan, R. F. & Petersen, J. L. Synthesis, structure and reactivity of *rac*- $\text{Me}_2\text{Si}(\text{indenyl})_2\text{Zr}(\text{NMe}_2)_2$. *Organometallics* **15**, 4038–4044 (1996).

39. Lenton, T. N. et al. Formation of trivalent zirconocene complexes from ansa-zirconocene-based olefin-polymerization precatalysts: an EPR- and NMR-spectroscopic study. *J. Am. Chem. Soc.* **135**, 10710–10719 (2013).

40. Ning, Y., Cooney, M. J. & Chen, E. Y. X. Polymerization of MMA by oscillating zirconocene catalysts, diastereomeric zirconocene mixtures, and diastereospecific metallocene pairs. *J. Organomet. Chem.* **690**, 6263–6270 (2005).

41. Bochmann, M., Cannon, R. D. & Song, F. Kinetic and mechanism of alkene polymerization. *Kinet. Catal.* **47**, 160–169 (2006).

42. Song, F., Hannant, M. D., Cannon, R. D. & Bochmann, M. Zirconocene-catalysed propene polymerisation: kinetics, mechanism and the role of the anion. *Macromol. Symp.* **213**, 173–185 (2004).

43. Su, Y., Song, Y. & Xiang, L. Continuous-flow microreactors for polymer synthesis: engineering principles and applications. *Top. Curr. Chem.* **376**, 44 (2018).

44. Zhang, J. S., Zhang, C. Y., Liu, G. T. & Luo, G. S. Measuring enthalpy of fast exothermal reaction with infrared thermography in a microreactor. *Chem. Eng. J.* **295**, 384–390 (2016).

45. Hany, C., Lebrun, H., Pradere, C., Toutain, J. & Batsale, J. C. Thermal analysis of chemical reaction with a continuous microfluidic calorimeter. *Chem. Eng. J.* **160**, 814–822 (2010).

46. Pradere, C., Joanicot, M., Batsale, J.-C., Toutain, J. & Gourdon, C. Processing of temperature field in chemical microreactors with infrared thermography. *Quant. Infrared Thermogr. J.* **3**, 117–135 (2007).

47. Terms, F. Heat of polymerization. *Polym. Rev.* **3**, 339–356 (1969).

48. Rizkin, B. A. & Hartman, R. L. Supervised machine learning for prediction of zirconocene-catalyzed α -olefin polymerization. *Chem. Eng. Sci.* **210**, 115224 (2019).

49. Iooss, B. & Lemaître, P. A review on global sensitivity analysis methods. *Uncertain. Manag. Simulation-Optimization Complex Syst. Algorithms Appl.* **59**, 101–122 (2015).

50. Rizkin, B. A., Shkolnik, A. S., Ferraro N. J. & Hartman R. L. Combining automated microfluidic experimentation with machine learning for efficient polymerization design. *Zenodo* <https://doi.org/10.5281/zenodo.3706730> (2020).

51. Rizkin, B. A., Shkolnik, A. S., Ferraro N. J. & Hartman R. L. Combining automated microfluidic experimentation with machine learning for efficient polymerization design, control code. *Zenodo* <https://doi.org/10.5281/zenodo.3706734> (2020).

Acknowledgements

This material is based on work supported by the National Science Foundation under grant no. CBET-1701393. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Author contributions

Conceptualization was provided by R.L.H., data curation by B.A.R., formal analysis by B.A.R. and A.S.S., funding acquisition by R.L.H., investigation by B.A.R. and A.S.S., methodology by B.A.R. and N.J.E., project administration by B.A.R. and R.L.H., resources by R.L.H., software by B.A.R., A.S.S. and N.J.E., supervision by R.L.H., validation by B.A.R. and A.S.S., visualization by B.A.R., A.S.S. and R.L.H., writing of the original draft by B.A.R. and A.S.S. and review and editing by R.L.H.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s42256-020-0166-5>.

Correspondence and requests for materials should be addressed to R.L.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Code and Software Submission Checklist

Prior to submitting your work to Nature Research, we strongly recommend that you ask at least one colleague who is unfamiliar with your software to install the tool(s), follow the instructions, and provide feedback. This process will help ensure that reviewers will also be able to run your software.

You must submit all required content as a single zip file prior to peer review or provide a link where editors and reviewers can access all required content.

► Required content

- Compiled standalone software and/or source code
- A small (simulated or real) dataset to demo the software/code

A README file that includes:

1. System requirements
 - All software dependencies and operating systems (including version numbers)
 - Versions the software has been tested on
 - Any required non-standard hardware
2. Installation guide
 - Instructions
 - Typical install time on a "normal" desktop computer
3. Demo
 - Instructions to run on data
 - Expected output
 - Expected run time for demo on a "normal" desktop computer
4. Instructions for use
 - How to run the software on your data
 - (OPTIONAL) Reproduction instructions

We encourage you to include instructions for reproducing all the quantitative results in the manuscript.

► Additional information

Describe your software's license for use. We strongly recommend using a [license](#) approved by the [Open Source Initiative](#).

GPL v3

Provide a link to the code in an open source repository (when available).

<http://doi.org/10.5281/zenodo.3647597>

Your manuscript should include a complete, detailed description of the code's functionality (i.e. pseudocode).

Please indicate where this is found:

- Main text
- Methods section
- Elsewhere (specify):

► Examples of well-structured software packages

1. <https://github.com/neurodata-papers/MGC>
2. <https://github.com/neurodata-papers/LOL>
3. <https://www.nature.com/nbt/journal/v34/n6/abs/nbt.3569.html#supplementary-information>
4. <https://www.nature.com/nature/journal/v548/n7669/full/nature23463.html#extended-data>
<https://github.com/yasharhezaveh/EnsaI>
5. <https://www.nature.com/nbt/journal/v34/n11/full/nbt.3685.html#supplementary-information>
<https://github.com/IFIproteomics/LFQbench>