Reaction Chemistry & Engineering

Tuning reaction products by constrained optimisation

Journal:	*Reaction Chemistry & Engineering*
Manuscript ID	RE-ART-08-2017-000123.R1
Article Type:	Paper
Date Submitted by the Author:	01-Sep-2017
Complete List of Authors:	Walker, Barnaby; Imperial College London, Department of Chemistry Bannock, James; Imperial College London, Chemistry; NTNU. Faculty of Natural Sciences, Chemistry Nightingale, Adrian; Southampton University, Faculty of Engineering and the Environment de Mello, John; Imperial College London, Department of Chemistry; NTNU. Faculty of Natural Sciences, Chemistry

Tuning reaction products by constrained optimisation

Barnaby Walker[1], James Bannock[1,2], Adrian Nightingale[3] and John deMello[1,2*]

[1]Centre for Plastic Electronics, Department of Chemistry, Imperial College London, Exhibition Road, South Kensington, London SW7 2AY

[2]Centre for Organic Electronic Materials, Department of Chemistry, NTNU N-7491 Trondheim, Norway

[3]Engineering and the Environment, University of Southampton, Building 13, Highfield, Southampton SO17 1BJ

[*]j.demello@imperial.ac.uk

We describe an effective means of defining optimisation criteria for self-optimising reactors, applicable to situations where a compromise is sought between several competing objectives. The problem is framed as a constrained optimisation, in which a lead property is optimised subject to constraints on the values that other properties may assume. Compared to conventional methods (using weighted-sum- and weighted-product-based merit functions), the approach described here is more intuitive, easier to implement, and yields an optimised solution that more faithfully reflects user preferences. The method is applied here to the synthesis of o-xylenyl adducts of Buckminsterfullerene, using a cascadic reaction of the form $X_0 \rightarrow X_1 \rightarrow X_2 \rightarrow \ldots X_N$. Specifically, we selectively target the formation of the (technologically useful) first- and second-order adducts X_1 and X_2, while at the same time suppressing the formation of unwanted higher-order products. More generally, the approach is applicable to any chemical optimisation involving a trade-off between competing criteria. To assist with implementation we provide a self-contained software package for carrying out constrained optimisation, together with detailed tutorial-style instructions.

The goal of finding an efficient route to a target molecule, while at the same time minimising the formation of unwanted side products, lies at the heart of synthetic chemistry. In the ideal case, where the target molecule corresponds to the sole end-point of a reaction, a near-quantitative product yield may be readily achieved by allowing the reaction to progress to completion. More often than not, however, the target is just one of several possible end products, or else it is an intermediate that can only be obtained by quenching the reaction before it has reached completion. In such circumstances, a mixture of reaction products is inevitably obtained, with the yield of the target molecule depending on the (typically unknown) kinetics of the reaction and the specific reaction conditions employed. Manually searching for reaction conditions that deliver an acceptable yield of the target molecule is a laborious undertaking, requiring extensive experimentation and chemical intuition. Even then, there is no guarantee the chosen conditions will correspond to the best attainable solution.

In this paper we set out an easily-implemented and fully automated approach for preferentially synthesising one or more target molecules amongst a larger group of possible products, using a technique that (given sufficient time) will yield a globally optimised solution. Our approach builds on previous work in the area of 'intelligent' or 'self-optimising' reactors[1–5], using an automated reactor with on-

line analysis and algorithmic control to repeatedly update the reaction conditions until a desired goal has been achieved. For each set of reaction conditions tested, the system is allowed to stabilise, a measurement is made using the on-line analysis system, and a scalar *merit value* that quantifies the acceptability of the product is then extracted from the data. In this way the overall physical process may be treated as a mathematical *objective* function in which the inputs are the selected reaction conditions and the output is the merit value. Assuming lower merit values signify superior products, optimisation of the chemical process is formally equivalent to minimisation of the associated objective function, and may accordingly be achieved using numerical techniques.

In two of the earliest reports in the field, self-optimisation was used by Krishnadasan et al.[1] in 2007 to tailor the spectral characteristics of metal chalcogenide quantum dots and by McMullen et al.[6] in 2010 to optimise the Knoevenagel condensation reaction of p-anisaldehyde and malononitrile. Further important contributions in the area have been made by the groups of Bourne[7,8], Jensen[9–11], and Poliakoff.[12–14] Using a variety of in-line/on-line analysis techniques – including infrared absorption spectroscopy[9], visible fluorescence spectroscopy[1], chromatographic separations[6,8,10–15], nuclear magnetic resonance[16], and mass spectrometry[7,17] – self-optimisation has been successfully used to optimise the yield and/or production rate of a variety of target molecules[6–17] and to control the physical properties of materials[1,18]. The above optimisations were carried out using a mixture of local[6,9–18] and global[1,6,8,7] search methods. For unknown chemical systems that may potentially exhibit multiple optima, global routines that fit measured data to approximating surfaces are typically preferred since they do not get trapped in sub-optimal local minima, can cope with measurement noise, and – by avoiding the need for derivative calculations – require relatively few function evaluations to locate optima. In situations where the chemical parameter space is monotonic with a single minimum (or multiple minima exist but the approximate location of the global minimum is known) local search methods may sometimes offer faster convergence.

In many cases a trade-off or compromise must be reached between several competing criteria. Mathematically, this may be achieved through the use of a compound merit function, typically formed from a weighted sum[19,20] or weighted product[20,21] of individual merit functions that separately take into account each property being optimised. Weighted-product-based merit functions were used by Krishnadasan et al. to maximize the intensity of quantum dot emission at a target wavelength[1] and by Jumban et al. to achieve an optimised trade-off between the production rate and yield of methylated ethers[15]; while weighted-sum-based merit functions were used by Moore et al. to achieve an optimised trade-off between the production rate and the conversion efficiency of a Paal–Knorr reaction[9].

The above studies showed that merit-based multi-objective optimization can be a powerful method for chemical optimization. However, its success hinges on the ability of the merit function to reduce multiple property values to a single, meaningful number that can be used to objectively rank the adequacy of

different outcomes. Unfortunately, devising a suitable merit function can be a fraught endeavor[17,19], especially when there are more than two parameters to balance: extensive physical experimentation and mathematical manipulation is often required to find an appropriate form of merit function that sensibly balances the different optimisation criteria and, even then, there is no guarantee that the merit-based ranking will fully accord with user perception.

The lack of a straightforward method for codifying product requirements in the form of chemical merit functions is a major obstacle to the widespread deployment of self-optimising reactors. What is needed is an easily implemented procedure that allows a user to set out all requirements in a simple, intuitive form that can then be directly translated into a usable merit function without significant experimentation or mathematical effort. Here we demonstrate how this may be readily achieved by configuring the problem as a constrained optimisation, in which a lead property is optimised, subject to lower and upper limits being placed on the values that other properties of interest may assume. By way of example, in a typical polymerization reaction, the lead property might be the conversion rate (which we wish to maximize), while typical constrained properties could include the weight-averaged molecular weight (which should fall between certain application-dependent bounds) and the polydispersity index (which should not exceed a maximum level).

The constraints are handled here using an analytical procedure due to Huyer and Neumaier.[22,23] In the discussion below we focus primarily on implementational aspects of the method; a description of its mathematical basis may be found in the SI. We assume that our goal is to minimise the lead property subject to specific constraints on other properties. (If the goal were to maximize the lead property, we would minimize its negative). For each constrained property the user specifies a range of values [F^{Lower}, F^{Upper}] within which that property should *preferably* lie plus a parameter Δ corresponding to the maximum permitted deviation from the preferred range (see Fig. S1a). Property values that lie within the preferred range [F^{Lower}, F^{Upper}] completely satisfy user specifications and hence are said to be "fully feasible". Values that lie outside the preferred range but within the expanded range [F^{Lower}-Δ, F^{Upper} +Δ] partially satisfy user requirements and are said to be "semi feasible" (since they lie within a permitted margin of the preferred range). Values outside the expanded range do not meet user requirements and are said to be "infeasible". It is the goal of the optimisation procedure to identify the set of reaction conditions that minimises the value of the lead property, while at the same time ensuring the values of all constrained properties are feasible or at worst semi-feasible (i.e. they lie within or as close as possible to their preferred windows).

The procedure begins with an initial search of the chemical parameter space to identify at least one data point that completely satisfies all constraints (see Methods). This is typically a straightforward task since any fully feasible point will suffice, irrespective of its lead property value. Two experimental parameters are extracted from these initial measurements: f_0 the value of the lead property at the best feasible point (i.e. the point with the lowest lead property value); and

Δ, the median value of $|f(\mathbf{x}) - f_0|$ where $f(\mathbf{x})$ is the value of the lead property at a data point $\mathbf{x}$.

Using the experimentally determined parameters f_0 and Δ, together with the user defined-parameters F^{Lower}, F^{Upper} and Δ for each constrained property, one may then construct a merit function $f^{\text{soft}}(\mathbf{x})$ that takes into account both the value of the lead property and the specified constraints:

$$f^{\text{soft}}(\mathbf{x}) = f^*(\mathbf{x}) + p(\mathbf{x}) = \frac{f(\mathbf{x}) - f_0}{\Delta + |f(\mathbf{x}) - f_0|} + 2\left(\frac{\sum_i p_i^2(\mathbf{x})}{1 + \sum_i p_i^2(\mathbf{x})}\right) \qquad (1)$$

where the index i runs over all constrained properties and $p_i(\mathbf{x})$ is assigned a value $[F_i^{\text{Lower}} - F_i(\mathbf{x})]/\,\Delta F_i$, zero, or $\left[F_i(\mathbf{x}) - F_i^{\text{U}}\right]/\,\Delta F_i$ according to whether $F_i(\mathbf{x})$ lies below, within, or above the preferred bounds of the window. With the merit function defined in this way, constrained optimisation is straightforwardly carried out by searching for the location $\hat{\mathbf{x}}$ of the global minimum of $f^{\text{soft}}(\mathbf{x})$.

We note here some pertinent properties of Eq. (1). The first term $f^*(\mathbf{x})$ is simply a rescaled variant of the original unconstrained merit function $f(\mathbf{x})$: $f^*(\mathbf{x})$ varies monotonically with $f(\mathbf{x})$, so the optima of $f^*(\mathbf{x})$ occur at the same locations as the optima of $f(\mathbf{x})$. The second term $p(\mathbf{x})$ is a penalty term that has the effect of increasing the value of $f^{\text{soft}}(\mathbf{x})$ whenever a property value lies outside its preferred window (otherwise if all constraints are satisfied its value is zero). From the construction of Eq. (1), the merit values are bounded to lie between -1 and 3. Fully feasible points have merit values less than one (with $\mathbf{x}_0$ having a merit value of zero), infeasible points have merit values greater than one, while semi-feasible points can span the full range of merit values from -1 to 3. It follows from these properties that the optimisation procedure will never prefer an infeasible point over a feasible point (since the infeasible point will always have a higher merit value). Moreover, it will only prefer a semi-feasible point over the best identified feasible point if the value of the lead property at the semi-feasible point is substantially lower than that at the best identified feasible point (resulting in a lower overall merit value). The complete step-by-step procedure for carrying out the constrained optimisation is summarised in the flow diagram of Fig. S1b for a problem involving two constrained variables. The diagram makes clear the simplicity of the procedure, which in practice is scarcely more difficult to implement than a standard unconstrained optimisation.

To exemplify the application of the procedure to chemical optimisation, we show how it may be used to tune the products of a cascadic reaction of the type $X_0 \rightarrow X_1 \rightarrow X_2 \rightarrow \ldots X_N$. It is a characteristic feature of such reactions that a mixture of products is present at all intermediate times, with the instantaneous distribution of products depending (in an often complicated way) on the underlying kinetics of the reaction. For the purposes of exposition, we specifically focus on the synthesis of *o*-xylenyl adducts of Buckminsterfullerene by the reaction of C_{60} with cyclic esters of a hydroxy sulfinic acid (sultines)[25], see Scheme 1. While these molecules have important applications as light-harvesting agents and electron conductors[25–27], the details of their use need not concern us here. Suffice to say, it is frequently the first- and second-order adducts that are used in

practice, with the presence of significant quantities of higher-order adducts having a detrimental impact on optoelectronic behaviour[28]. Hence, there is a need to identify reaction conditions that maximise the yields of singly- and doubly-functionalised molecules, while minimising the fractions of higher-order adducts.

The high stabilities of the reagents used in the synthesis of *o*-xylenyl-functionalised C_{60} make it a viable candidate for self-optimisation, with the one-pot nature of the reaction lending itself to straightforward automation. Here we used a simple single-phase, capillary-based flow reactor*, incorporating: two syringe pumps, separately loaded with C_{60} and sultine solutions; a passive y-shaped mixer for bringing the two solutions into contact; and a cylindrical solid-state heater for thermally activating the reaction (see Fig. 1 and Methods). Control over the time, temperature and chemical composition of the reaction was achieved by making independent adjustments to the infusion rates of the two reagent streams and the temperature of the heater.

High performance liquid chromatography (HPLC) was selected for on-line analysis, being a moderately fast, flow-compatible method for analysing multicomponent solutions. HPLC has been successfully applied to self-optimising reactors by the groups of Jensen and Bourne[6,8,10,11]. For the current work, discrimination of the *o*-xylenyl adducts was achieved using pyrenylpropyl-functionalised silica as the stationary phase and a mixture of toluene and hexane as the mobile phase[27]. (The affinity of the C_{60} adducts to the mobile phase rises substantially with increasing order number, resulting in progressively shorter and well separated elution times). Following each change of reaction conditions, the system was allowed to stabilize for a time period equal to twice the current calculated residence time. A sample of the product mixture was then taken by diverting the out-flow of the reactor to a sample coil, from where it was injected into an HPLC column using a high-pressure switching valve. Detection was carried out optically by absorption spectroscopy. For ease of comparison all chromatograms reported here have been normalised to the total area under the measured peaks. The relative concentrations of the adducts were determined from the areas under the chromatographic peaks, using a calibration curve.

Initial testing of the reactor was carried out by varying in turn the reaction temperature, reaction time and molar ratio of sultine to C_{60}, while holding the other two parameters constant. The effects of varying these parameters on the measured chromatograms are shown in Fig. 2a(i, ii, iii), while the effects on the mole fractions of the adducts are shown in the stacked area plots of Fig. 2b(i, ii, iii). Up to four distinct and well separated chromatographic peaks were observed in each case at elution times of approximately 4.4, 5.4, 7.5, and 11.4 min, corresponding to triply- (X_3), doubly- (X_2), singly- (X_1) and un- (X_0) functionalised C_{60}, respectively (see Methods). Similar trends are evident in each plot, with there being a reduction in the C_{60} peak accompanied by an increase in the other three peaks as the variable parameter was increased. Hence, for the

* A flow synthesis for indene C_{60} adducts has previously been reported by Seyler *et al*[29].

conditions tested, increases in the temperature, reaction time and sultine:C_{60} ratio all resulted in increased conversion of C_{60} into higher-order adducts in broad accordance with expectation.

The plots in Fig. 2 show smooth trends in the concentrations of the adducts as the reaction parameters were varied, indicating a well controlled reaction environment and low noise in the measurement system – necessary characteristics for developing a reliable self-optimising reactor. However, collectively, the plots represent a very limited data set since only one reaction parameter was varied in each case, with the other two being held fixed. There is no guarantee similar trends would be obtained for different values of the fixed parameters. (Indeed, given the cascadic nature of the reaction, the mole fractions of all species must eventually decrease as the reaction proceeds and they are converted into higher-order adducts in contrast to the behaviour seen in Fig. 2.)

For a more complete understanding of how the distribution of reaction products depends on the reaction conditions, the three-dimensional parameter space should be mapped out by varying all reaction parameters in parallel. The result of doing this at a coarse level – using a $6 \times 6 \times 6$ set of evenly spaced grid points for the temperature, time and sultine:C_{60} ratio – is shown in Fig. 3. The measurements were carried out in a randomized order, with several sets of reaction parameters being repeated multiple times. Chromatograms for the replicate measurements showed only slight differences (see Fig. S2), indicating negligible system drift over the timescale of the measurement run, with only small sample-to-sample variations due to minor reactor instability and/or measurement errors.

A number of general observations may be made about the data in Fig. 3: as before, in all cases a mixture of reaction products was obtained; increasing the reaction time, temperature and/or sultine:C_{60} ratio resulted in a progressive reduction in the mole fraction of unreacted C_{60} and a progressive increase in the mole fraction of the (typically unwanted) third-order adduct; at lower temperatures the (typically preferred) first- and second-order adducts were the dominant products, while at higher temperatures and sultine:C_{60} ratios the third-order adduct dominated; an increase in higher-order adducts was evident at higher temperatures and sultine concentrations, under which conditions C_{60} was fully depleted during the reaction, consistent with the cascadic reaction mechanism.

The systematic, reproducible nature of the data in Figs. 3 and S2 suggest the complete system – i.e. the reagents, reactor and measurement system taken as a whole – is a good candidate for self-optimisation (reproducibility being a pre-requisite for successful optimisation). As noted above, first- and second-order adducts of C_{60} are typically preferred for electronic applications, with the presence of higher-order adducts often having a detrimental impact on optoelectronic behaviour. Hence, as an initial test, a simple unconstrained optimisation was carried out, in which we sought to minimise the formation of the third-order adduct by setting the merit function – i.e. the quantity to be

minimised – equal to the mole fraction of the third-order adduct, [X_3] (see Table 1).

For all optimisation runs reported here we used the global optimisation code Stable Noisy Optimisation by Branch and Fit (SNOBFIT)[22] – a noise-tolerant routine that first divides the search space into separate boxes that each contain one sampled data point, and then forms quadratic models around each point; local searching is handled by selecting the model minima as new evaluation points, and global searching is handled by making measurements in large boxes (which correspond to large regions of unexplored territory). In each iteration, a batch of new points is selected for testing, some for local optimisation and others for global searching. In all cases: the temperature was varied between 100 and 150 °C; the reaction time was varied between 3 and 31 min; the flow-rate ratio was varied between 2:1 and 1:2[†]; the routine was started 'cold', i.e. with no pre-supplied measurement data; and approximately one hundred trial measurements were carried out during each search, of which 30 % were selected for global searching and 70 % for local refinement (see Methods).

The left side of Fig. 4 shows a scatter plot of the sampled data from Run I, in which the marker locations indicate the reaction conditions and the colours denote the merit values: lower merit values are denoted by darker colours; and, for ease of distinction, merit values higher than the median value of 0.011 are coloured red, while those lower than the median value are coloured blue. (Pale, near-white markers denote points with merit values equal or close to the median value). The wire-frame 'cage' denotes the bounds defined by the flow rate and temperature constraints. The algorithm has evidently sampled certain regions of the parameter space preferentially – in particular the region next to the lower half of the right face of the cage, corresponding to lower temperatures and lower sultine concentrations. The data markers in this region are all coloured blue, signifying low merit values, i.e. low mole fractions of the third-order adduct.

The mole fractions [X_0], [X_1], [X_2] and [X_3] of the four adducts are plotted against measurement number in Fig. 5a(i). The mole fraction distribution can be seen to fluctuate substantially between successive measurements due to local searching at the beginning of each batch of points, and global searching at the end of each batch: in the local phase, the parameter space is sampled preferentially in

[†] When the sultine flow rate (F_S), the C_{60} flow rate (F_{C60}) and the temperature (T) are plotted along the x, y and z axes, respectively, the constraints define a *right prism* shaped parameter space with vertical walls and a trapezoidal base (see Fig. 4) – the non-parallel sides of the base being due to the constraints imposed on the flow-rate ratio. SNOBFIT by contrast accepts only box-bounded constraints, corresponding to a *rectangular prism* shaped parameter space. The trapezoidal flow constraints were handled by a two-stage transformation of the external variables F_S and F_{C60} to box-bounded internal variables. The first stage involved a rotation of each coordinate [F_S, F_{C60}] by 45°, the angle between the axis of symmetry of the trapezium and the y-axis; while the second stage involved a mapping of each rotated coordinate to the internal rectangular space used by SNOBFIT, using the shadow-map algorithm of Tobias and Tiow-Seng[30].

regions where the existing merit values are low, yielding new merit values that are typically low also; in the global phase, unexplored regions of the parameter space are sampled where the merit values tend to be large (but where superior, as yet undiscovered, minima might potentially exist). Insight into the behaviour of the algorithm can be drawn from Figs. 6a(i) and 6a(ii) which show successive merit values and a histogram of the merit values, respectively. The plots indicate that the algorithm preferentially explored regions of the parameter space that yielded low merit values, with more than 80 % of sampled data points having merit values of 0.050 or less (compared to a maximum measured value of 0.389). The best point with the lowest merit value of 0.001 had mole fractions of 0.318, 0.649, 0.033 and 0.001 for the zero-, first-, second- and third-order adducts, respectively, confirming effective suppression of the third-order adduct.

Fig. 5b(i) shows mole fractions for the best result obtained so far versus measurement number, *N*. Changes in the best result occurred whenever the most recently tested reaction conditions gave rise to a product distribution with a lower merit value than the previous best result, i.e. a lower mole fraction for the third-order adduct. Improvements occurred at N = 5, 15, 18 and 53, with the mole fraction of the third-order adduct falling from an initial value of 0.384 at N = 1 to a value of 0.001 at N = 53. At the same time, the concentration of unreacted C_{60} increased from an initial value of 0.000 at N = 1 to a value of 0.318 at N = 53. Hence, it is evident that the reduction in $[X_3]$ was primarily achieved at the expense of an undesirable drop in C_{60} conversion.

To improve the conversion rate, a second optimisation (Run II) was carried out in which the mole fraction of the third-order adduct was again minimized, but constraints were applied to the combined mole fraction of the (desired) first- and second-order adducts: soft and hard lower limits of 90 and 60 %, respectively, were imposed on the combined mole fraction $[X_{1,2}] = [X_1] + [X_2]$, using a constraint window for $[X_{1,2}]$ of [0.9, 1.0] and a σ-value of 0.3. The right side of Fig. 4 shows a scatter plot of the measurements made during Run II, where merit values above the median value (0.020) are again coloured red and those below the median value are coloured blue. As before, the algorithm preferentially sampled the low temperature, low sultine region close to the right face of the cage. However, this time the majority of sub-median data points occurred close to the foremost 'spine' of the cage, corresponding to lower flow rates, i.e. longer reaction times. Hence, in common with the unconstrained case of Run I, the algorithm ensured a low mole fraction of the third-order adduct by selecting low temperatures and low sultine concentrations, but this time it selected longer reaction times that resulted in higher C_{60} conversion. The best point with the lowest merit value of -0.166 had mole fractions of 0.101, 0.784, 0.110 and 0.004 for $[X_0]$, $[X_1]$, $[X_2]$ and $[X_3]$.

The mole fraction distribution is plotted against measurement number in Fig. 5a(ii). While the observed behaviour is similar to that seen in Fig. 5a(i) for Run I – with the mole fractions again fluctuating substantially between successive measurements as the optimisation routine switched between local and global searching – the average height of the dark purple bars that denote unreacted C_{60} is significantly lower. Hence, compared to the first optimisation

run, the algorithm preferentially selected reaction conditions that resulted in substantial conversion of C_{60}. Fig. 5b(ii) shows mole fractions for the best result to date versus N, where the best result corresponds to the outcome with the lowest (soft) merit value. Improvements occurred at N = 3, 13, 14 and 33, with $[X_3]$ assuming respective values of 0.013, 0.006, 0.005 and 0.004 and $[X_0]$ assuming respective values of 0.067, 0.155, 0.134 and 0.101. Hence, although the initial reduction in $[X_3]$ was again achieved at the cost of an increase in $[X_0]$, the C_{60} concentration subsequently dropped without the concentration of the third-order adduct increasing. The algorithm therefore succeeded in its aim of minimizing the amount of the third-order adduct, while also achieving a close to ninety percent yield of first- and second-order adducts.

The contrasting behaviour of the constrained and unconstrained optimisation runs can be seen more easily by ranking the data from Figs. 5a(i) and 5a(ii) in order of decreasing third-order adduct. While the two ranked plots in Figs. 5c(i) and 5c(ii) are qualitatively similar in appearance – with the concentrations of the second- and third-order adducts *decreasing* and the concentrations of the zero- and first-order adducts *increasing* from left to right – clear differences are evident: the average mole fraction of the third-order adduct is slightly higher in the case of the constrained optimisation (Run II), while the average mole fraction of unreacted C_{60} is markedly lower. Hence it is evident that, during Run II, the algorithm preferentially probed regions of the parameter space that resulted in high C_{60} conversion.

The behaviour of (the constrained optimisation) Run II can be understood by examining the plots in Fig. 6b, which show successive values and the corresponding histograms for the total objective function $f^{\text{soft}}(\mathbf{x})$, the transformed merit function $f^*(\mathbf{x})$ and the constraint function $p(\mathbf{x})$. It can be seen that the algorithm preferentially explored regions of the chemical parameter space that yielded low merit values, with these low values being achieved through a combination of low $f^*(\mathbf{x})$ values and low $p(\mathbf{x})$ values. The $f^*(\mathbf{x})$ values, while running from –0.227 to 0.887, were strongly biased towards negative values, indicating that the algorithm identified many points that yielded a lower third-order adduct concentration than the first identified feasible point $\mathbf{x}_0$ (identified at N = 1). In addition, the $p(\mathbf{x})$ values were preferentially skewed towards zero, signifying a strong bias towards points that satisfied or nearly satisfied the constraint: of the 100 measurements, 46 resulted in a $p(\mathbf{x})$ value of zero, corresponding to fully feasible points that completely satisfied the applied constraint (i.e. yielded a combined mole fraction for X_1 and X_2 of > 90 %).

Fig. 7a(i) shows the chromatogram corresponding to the lowest merit value (0.001) obtained during Run I, while Fig. 7a(ii) shows chromatograms for the lowest merit value (-0.166) and the lowest fully feasible merit value (-0.123) obtained during Run II. Comparing these chromatograms, it can be seen that $[X_{1,2}]$ was substantially higher for Run II (~0.9) than Run I (0.68), consistent with the successful application of the constraint in the former case. The fully feasible and semi-feasible chromatograms of Fig. 7a(ii) are very similar to one another, implying similar amounts of the four adducts in both cases. The best feasible point had a third-order adduct mole fraction of 0.007 and a combined mole

fraction [$X_{1,2}$] of 0.905 for the first- and second-order adducts, compared to values of 0.004 and 0.894 for the best semi-feasible point. Hence, comparing the best point and the best feasible point, it is clear that – in accordance with the discussion above – an improvement (reduction) in the primary parameter [X_3] was achieved through a slight violation of the constraint on [$X_{1,2}$]. The violation in this particular case was rather small since the best feasible point had an [X_3] value that was already close to the lowest possible value of zero, so straying far outside the feasible zone would have caused a substantial increase in $p(\mathbf{x})$ without significantly reducing $f^*(\mathbf{x})$.

In Run II, a constraint was applied to [$X_{1,2}$], the combined mole fraction of the first- and second-order adducts, but the relative amounts of the two mole fractions were allowed to vary freely. While the best point and best feasible points corresponded to product mixtures that contained substantially more of the first-order adduct than the second-order adduct, this was not specifically encoded within the merit function. To *ensure* such an outcome, it is necessary to impose a constraint on the ratio $\mathfrak{R}_{12} = [X_1]/[X_2]$, alongside the existing constraint on [$X_{1,2}$]. To demonstrate the feasibility of doing this, in the next optimisation run (Run III), we additionally sought to enforce a lower limit of 4:1 on the ratio $\mathfrak{R}_{12} = [X_1]/[X_2]$, using a constraint window of [4,∞] and a sigma value of 0.3‡ (while maintaining the existing constraint on [X_{12}]).

Scatter plots for Run III, plots of mole fractions and merit values versus measurement number, and histograms of $f^{\mathrm{soft}}(\mathbf{x})$, $f^*(\mathbf{x})$ and $p(\mathbf{x})$ are provided in Fig. S3. The chromatograms for the best point and the best feasible point, obtained at measurement numbers 26 and 99 respectively, are shown in Fig. 7a(iii). The best feasible point had mole fractions of 0.085, 0.778, 0.128 and 0.009, implying an $\mathfrak{R}_{12}$ value of ~6.1:1 and an [X_{12}] value of 0.906 – both values being consistent with the specified limits. The best point by contrast had mole fractions of 0.129, 0.776, 0.091 and 0.004, implying an $\mathfrak{R}_{12}$ value of ~8.5:1 and an [X_{12}] value of 0.867 - the slight violation of the [X_{12}] constraint having led to a beneficial reduction in [X_3] from 0.009 to 0.004.

Obtaining a product mixture that is rich in the first-order adduct is not especially difficult, and indeed occurred *by chance* in Run II, even without imposing a constraint on $\mathfrak{R}_{12}$. For the fourth optimisation run (Run IV), we sought to obtain a product mixture that contained more of the second-order adduct than the first-order adduct. Owing to the cascadic nature of the reaction, this is a substantially harder challenge since the second-order product lies adjacent in the reaction sequence to the unwanted third-order adduct, meaning conditions that favour the formation of the second-order product are liable to promote (to a lesser extent) the unwanted formation of the third-order product.

‡ In this work, we were primarily interested in achieving fully feasible solutions, for which the exact σ-value chosen is of secondary importance. Hence a common value of 0.3 was used for all constraints.

To assess the feasibility of attaining an end-product rich in the second-order adduct, we placed what we hoped would be a physically achievable upper limit of 1:2 on the ratio $\Re_{12} = [X_1]/[X_2]$, using a constraint window of [0, 0.5] and a sigma value of 0.3, while again maintaining the constraint on $[X_{12}]$. The resultant scatter plots, plots of mole fraction and merit value versus measurement number, and histograms of $f^{\text{soft}}(\mathbf{x})$, $f^*(\mathbf{x})$ and $p(\mathbf{x})$, are provided in Fig. S4. In contrast to the previous results, no feasible point was identified during the course of the run, with the histogram for $p(\mathbf{x})$ spanning the range 0.280 to 1.999 due to partial ($0 < p(\mathbf{x}) \leq 1$) or complete ($p(\mathbf{x}) > 1$) violation of at least one of the constraints in all cases. There were 50 cases of partial violations, of which 36 were due to partial violation of the yield constraint only, and 14 were due to partial violation of both constraints. There were 50 cases of full violations, of which 2 were due to full violation of the yield constraint, 26 were due to full violation of the ratio constraint, and 22 were due to full violation of both constraints. The constraint violations are consistent with the difficulty noted above of suppressing the formation of the third-order adduct, while at the same time trying to achieve a high mole fraction of the second-order adduct.

The chromatogram for the best point in Run IV, i.e. the point with the lowest soft merit value (-0.212), is shown in Fig. 7a(iv). From the areas under the chromatographic peaks, $[X_0]$, $[X_1]$, $[X_2]$ and $[X_3]$ were determined to be 0.001, 0.288, 0.510 and 0.201, respectively. Hence, even at the best point, a substantial amount of the third-order adduct was present and both constraints were partially violated, with $[X_{12}]$ having a value of 0.798 (i.e. less than 0.9) and $\Re_{12}$ having a value of 0.564 (i.e. greater than 0.5). As expected from the above discussion, in an effort to find conditions that came close to satisfying the constraint on $\Re_{12}$, the algorithm identified conditions that resulted in a high concentration of the third-order adduct and virtually no C_{60}.

The difficulty of simultaneously achieving a high ratio of the second- to first-order adducts, while at the same time suppressing formation of the third-order adduct is evident from Fig. 7b, which shows chromatograms (acquired during Run IV) at several illustrative values of $[X_3]$. From the chromatograms, it is evident that the ratio of the first-order adduct to the second-order adduct decreases steadily as $[X_3]$ increases. In Fig. 7c, the ratio $\Re_{12}$ is plotted against the mole fraction $[X_3]$ for each measurement in Run IV. The data points lie on a trade-off curve, with desired reductions in $\Re_{12}$ leading to an unwanted increase in $[X_3]$. From the trade-off curve, it is evident that $[X_3]$ can only be kept below the 10 % level (which we consider to be an acceptably low value) if $\Re_{12}$ is greater than approximately 1.5. Armed with this information, a fifth optimisation run (Run V) was carried out using an expanded constraint window for $\Re_{12}$ of [0, 1.5] and the same sigma value of 0.3. Scatter plots for Run V, plots of the mole fraction distributions and merit values versus measurement number, and histograms of $f^{\text{soft}}(\mathbf{x})$, $f^*(\mathbf{x})$ and $p(\mathbf{x})$ are provided in Fig. S5.

Using the expanded constraint window for $\Re_{12}$, an initial feasible point $\mathbf{x_0}$ was found at N = 21. The same point turned out to be both the best feasible point and the point with the lowest overall merit value (see Fig. 7a(v) for chromatogram). The mole fractions for $[X_0]$, $[X_1]$, $[X_2]$ and $[X_3]$ were 0.007, 0.535, 0.373 and 0.085,

respectively, corresponding to values of 0.909 and 1.43 for $[X_{12}]$ and $\Re_{12}$ in close agreement with the trade-off curve of Fig. 7c. Hence, given the successful discovery of a fully feasible solution in Run V, it is evident that data generated in an initial optimisation run based on unsatisfiable constraints (i.e. Run IV) may nonetheless still be used to identify more appropriate constraints for future runs. In this way, it is easy to learn from experience and progressively modify constraints over a number of repeat runs until the constraints are appropriately matched to the underlying chemical system.

While the merit functions proposed above are sensible choices for achieving the intended outcomes, they are not the only options. The same (or at least a similar) result should be obtained for any sensibly constructed merit function that has been designed to achieve a particular goal. Based on the trade-off curve of Fig. 7c, as an alternative to the merit function used for optimisation Run V, the ratio $\Re_{12}$ could be used as the principal property to be minimised, subject to $[X_{12}]$ lying in the constraint window [0.9, 1.0] (using here the same sigma value of 0.3). The results of framing the merit function in this way were investigated in Run VI. The chromatograms for the best feasible point and the best point are shown in Fig. 7a(vi) and can be seen to closely match those of the previous run, confirming the approximate equivalence of the optimisation criteria. The best feasible point had mole fractions of 0.007, 0.527, 0.378 and 0.087 for $[X_0]$, $[X_1]$, $[X_2]$ and $[X_3]$, corresponding to values of 0.906 and 1.393 for $[X_{12}]$ and $\Re_{12}$in reasonably close agreement with the best feasible point of Run V. The best point by contrast had mole fractions of 0.004, 0.382, 0.458 and 0.157, corresponding to a favourable reduction in $\Re_{12}$ (to 0.833) at the expense of an unfavourable reduction in $[X_{12}]$ (to 0.840) .

Discussion and Conclusion

The results presented above illustrate the use of self-optimisation in two distinct forms: the first form, *blind discovery*, relates to the optimisation of an unknown system, for which little information is available at the outset; while the second, *rediscovery*, relates to a repeat optimisation of a (partially) known system. In the case of blind discovery, it is not known in advance what can be achieved by the system. Physically plausible constraints must therefore be proposed on the basis of physicochemical intuition in the hope that an acceptable solution will be attained. The acceptability or otherwise of the solution is determined by the appropriateness of the constraints chosen. In cases where the solution is not acceptable to the user, blind optimisation may be straightforwardly followed by one or more refinement stages, in which the constraint windows are iteratively modified to achieve a superior outcome.

Rediscovery relates to repeat optimisations of a well understood system for which a near optimal outcome is known in advance, but the detailed reaction conditions needed to achieve that outcome remain to be discovered. This might be the case, for instance, when resuming a previously optimised reaction after changing reagent batches or otherwise servicing/modifying the reactor, or on transferring the reaction to a similar, but untested, reactor. In such cases, it is

reasonable to expect broadly equivalent behavior across the reaction runs and reactors, but the detailed mapping of reaction conditions onto the final product may differ due to slight differences in reagent compositions or the mechanical configuration of the reactor(s).

Most of the optimisation runs reported above were carried out in the manner of blind runs, where we postulated appropriate constraints without using information gained in previous runs to guide our choice. Rather loose constraints were applied that had a significant chance of being satisfied immediately, recognising they could if necessary be tightened in subsequent runs to achieve a superior outcome. Run V is an example of rediscovery in the sense that we used the trade-off information acquired during Run IV to identify an achievable solution with an acceptable mole fraction distribution. We then modified the upper limit on $\Re_{12}$ accordingly to deliver that solution in Run V. (Run VI may be considered an example of re-optimisation for similar reasons.)

We stress again that the merit functions used here are constructed entirely on the basis of easily acquired physical information and consequently, once the appropriate constraints are established, they may be written down directly with no further work or mathematical manipulation being required on the part of the user. For the benefit of readers wishing to implement the procedure described in this paper, we have provided an easy-to-use self-contained software package (see https://github.com/jdmgroup/SNOBFit_for_chemical_optimisations) that takes care of the construction and subsequent optimisation of the merit function, together with detailed tutorial-style instructions. We hope the provision of this software will substantially simplify the implementation of self-optimising reactors, and so encourage their wider adoption by the general chemistry community.

Beyond the tuning of product distributions, the procedure used here is also applicable to reactions where product yield must be balanced against practical factors such as production rates and/or materials and energy costs. The approach has further applications in materials optimisation, where a compromise must frequently be reached between several physicochemical properties. For instance, using a conventional weighted product based multi-objective merit function, Krishnadasan et al.[1] reported a self-optimising reactor that optimised the emission intensity of quantum dots at a chosen emission wavelength. Owing to the difficulty of identifying weights that accurately encapsulated the intended outcome, small (nm-level) deviations from the target wavelength were heavily penalized even when they resulted in a substantial improvement in emission intensity. Framing the same problem as a constrained optimisation – in which the intensity is maximized subject to the emission wavelength lying in a prescribed range – would allow the trade-off to be precisely encoded within the merit function in a way that more closely reflects typical user preferences.

In summary, we have described a simple procedure for constructing multi-objective merit functions for self-optimising reactors. Framing the problem as a constrained optimisation, in which a principal property is optimised subject to

soft and hard constraints on the other parameters, allows the optimisation criteria to be set out in a manner that is intuitive even for the non-specialist. The specific method for constructing chemical utility functions used here offers substantial advantages over conventional approaches based on weighted sums and products, both in terms of their ease of construction and their mathematical behaviour. In particular, the merit function may be written down directly from the specified constraints without the need to tune weighing coefficients or penalty parameters, and given sufficient time (if satisfiable constraints are selected) the solution is guaranteed to minimise the lead property, while ensuring all other properties lie within the prescribed boundaries. The generic approach is not tied to any specific optimisation algorithm and consequently can be expected to simplify the implementation of self-optimising reactors in many situations, while at the same time yielding improved reaction products that more closely match user requirements.

Methods

Preparation of precursors for o-xylenyl C_{60} adducts

C_{60} was obtained from Solenne BV, while all other chemicals were obtained from Sigma-Aldrich. Sultine (1,4-dihydro-2,3-benzoxothiin 3-oxide) was synthesised using the protocol described in Kim *et al*[25]. A stock solution of sultine was prepared by dissolving under argon the unpurified product in o-dichlorobenzene (o-DCB) at a concentration of 1.4 mg/mL. A stock solution of C_{60} was prepared by dissolving under argon the as-received C_{60} in o-DCB at a concentration of 2 mg/mL. The stock solutions were stored under argon for up to a week before use.

Reactor setup

The reagent solutions were transferred to separate 500 mL flasks, where they were stored under argon and delivered to the reactor by a dual-channel continuous-flow syringe pump (Syrris Asia), using PTFE tubing (1 mm I.D., 2 mm O.D., Polyflon Technology Ltd.). The two solutions were merged using a static *y*-shaped mixer (PEEK, tP-512, Upchurch Scientific). Using the same diameter PTFE tubing, the mixed reagents were passed helically around a heater formed from a solid 88-mm-diameter cylindrical block of aluminium containing three symmetrically disposed inset cartridge heaters (3/8" × 2", 150 W, RS Components Ltd.). The temperature of the heater surface was monitored using a K-type flag-style thermocouple (25x13 mm, FL-K-2M, LABFACILITY) connected to a microcontroller (Arduino Uno) via a thermocouple amplifier (MAX6675, Adafruit). The total length of tubing in contact with the heater was 119 cm, equating to a heated reactor volume of 0.931 mL. The heater was enclosed in a plastic box, containing two fans inset into the walls for air cooling. The heater and fans were controlled by the microcontroller, using the Arduino PID library. A back-pressure regulator (20 psi, PEEK, Upchurch Scientific) was placed at the outflow of the reactor. The destination of the product stream was controlled by an injection valve as described below. The total length of tubing in the reactor was 202 cm, equating to a total reactor volume V of 1.587 mL.

On-line HPLC analysis

Eluent solvent (Hexane/Toluene, 1:3) was passed through a pyrenylpropyl-functionalised silica column (BuckyPrep, Cosmosil) using an HPLC pump (Model 12-6, SSI). Sample loading and injection were controlled by an injection valve (MXP-7900, Rheodyne) connected to the outflow of the reactor via a sample coil (Stainless Steel, Upchurch Scientific, V_{coil} = 5 μL). The sample transmittance after the column was monitored using a 390 nm light-emitting diode and an amplified silicon photodiode (OPT101, Texas Instruments), which were placed either side of transparent perfluoroalkoxy tubing (I.D. 0.50 mm, O.D 1/16 in, Upchurch Scientific). A second amplified silicon photodiode arranged at 90° to the LED allowed the signal from the first photodiode to be corrected for fluctuations in LED light intensity. The signals from the photodiodes were acquired using a data acquisition card (NI-6211, National Instruments) controlled by a MATLAB script.

Manual Operation

The heater, injection valve and syringe pump were controlled by a PC across the Universal Serial Bus (USB). A custom-written MATLAB script was used to step through a pre-determined sequence of reaction conditions. The set-point temperature of the heater was updated at the start of each step. Following a delay of at least one minute as described below, the two syringe pumps containing the C_{60} and sultine solutions were set to the specified volumetric flow-rates (F_{C60} and F_{Sultine}), with the injection valve oriented in the inject position (i.e. with the sample loop between the eluent stream and the column), so that the product flowed directly to the collection flask. The flow was allowed to stabilize for a duration Δt equal to twice the mean residence time ($\Delta t = 2V/[F_{C_{60}} + F_{\mathrm{Sultine}}]$). The injection valve was then switched to the load position (i.e. the sample loop was inserted between the product stream and the collection flask) and held there for a duration sufficient for 50 µL (= 10 V_{coil}) of fluid to pass through. With the product loaded in the sample coil, the injection valve was switched back to the inject position, causing the product to be carried by the eluent stream into the column. Chromatograms were obtained by recording the signals from the two photodiodes for thirteen minutes. The next step was started while the chromatogram from the previous step was still being recorded by first updating the target temperature and waiting for it to stabilize and then, at a time Δt before the end of the current HPLC measurement, setting the syringe pumps to infuse at the new flow rates.

Automated Operation

For ease of use a class-based MATLAB wrapper was written for the SNOBFit functions provided by Huyer and Neumaier.[§] At the start of each optimisation run, using a Latin Hypercube design, SNOBFit selected a batch of n_{point} randomised data points inside the region bounded by the parameter constraints. In subsequent iterations, SNOBFit selected new data points for measurement in batches of size n_{req}. New batches were generated until the total number n of function evaluations exceeded a pre-set limit n_{call}. (The final batch was carried out to completion). To initialise the soft optimisations, a feasible point $\mathbf{x}_0$ satisfying the condition $p(\mathbf{x_0}) = 0$ was first identified by running an unconstrained optimisation, using the penalty function $p(\mathbf{x})$ as the objective function to be minimised. Once a feasible point had been found, soft merit values for all data points so far measured were calculated using Eq. (1). The optimisation run was then restarted using the soft merit function $f^*(\mathbf{x})$ as the objective function. In cases where no feasible point had been found after fifty function evaluations, the soft optimisation was instead started using $2f_{\mathrm{max}} - f_{\mathrm{min}}$ as a proxy for a feasible point (see main text).

[§] http://www.mat.univie.ac.at/~neum/software/snobfit/

The following SNOBFit parameters (see Ref. 2) were used:

Parameter	Value	Description
N	3	Number of reaction parameters
Δf	0.02	Uncertainty in f, used for fitting
n_{point}	N+4 = 7	Number of points in initial batch
n_{req}	N+4 = 7	Number of points in subsequent batches
n_{call}	100	Maximum number of function calls
p	0.3	Probability of generating a point away from a minimum: used to control the balance of local and global searching.

Author contributions

B. W and J. dM designed the experiments. A. N. and B. W contributed to software development. J. B and B. W. developed the hardware. B. W. carried out the experiments. All authors contributed to the interpretation and analysis of the results and writing of the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Data statement

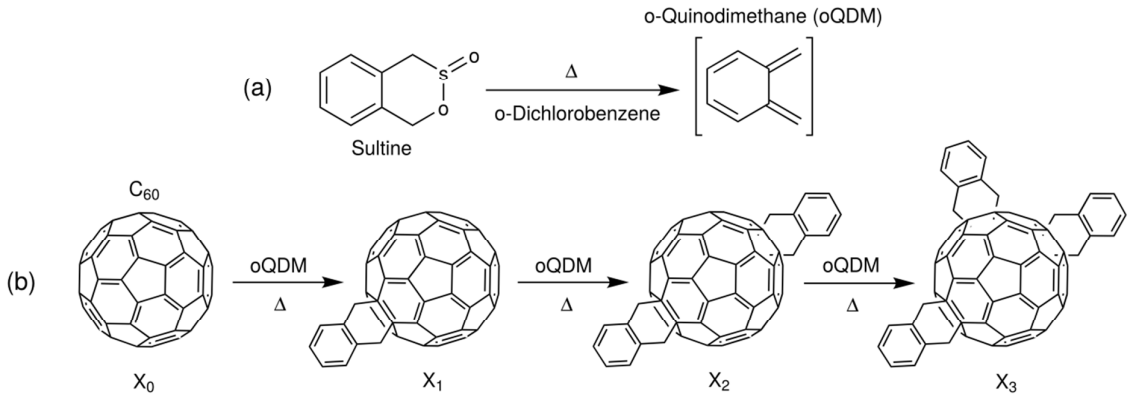
The datasets generated during the current study are available in the Imperial College Box repository at:
https://imperialcollegelondon.box.com/v/tuning-reaction-products

The optimisations were carried out using an easy-to-use, custom-written, class-based wrapper for SNOBFit which, together with detailed tutorial-style instructions, may be obtained from :
https://github.com/jdmgroup/SNOBFit_for_chemical_optimisations

References

1. Krishnadasan, S., Brown, R. J. C., deMello, a J. & deMello, J. C. Intelligent routes to the controlled synthesis of nanoparticles. *Lab Chip* **7,** 1434–41 (2007).
2. Fabry, D. C. *et al.* Online monitoring and analysis for autonomous continuous flow self-optimizing reactor systems. *React. Chem. Eng.* **1,** 129–133 (2016).
3. Reizman, B. J. & Jensen, K. F. Feedback in Flow for Accelerated Reaction Development. *Acc. Chem. Res.* **49,** 1786–1796 (2016).
4. Tibbetts, K. M., Feng, X.-J. & Rabitz, H. Exploring experimental fitness landscapes for chemical synthesis and property optimization. *Phys. Chem. Chem. Phys.* (2017). doi:10.1039/C6CP06187G
5. Holmes, N. & Bourne, R. A. in *Chemical Processes for a Sustainable Future* (eds. Letcher, T., Scott, J. & Patterson, D.) 28–43 (The Royal Society of Chemistry, 2015).
6. McMullen, J. P. & Jensen, K. F. An Automated Microfluidic System for Online Optimization in Chemical Synthesis. *Org. Process Res. Dev.* **14,** 1169–1176 (2010).
7. Holmes, N. *et al.* Online quantitative mass spectrometry for the rapid adaptive optimisation of automated flow reactors. *React. Chem. Eng.* **1,** 96–100 (2016).
8. Holmes, N. *et al.* Self-optimisation of the final stage in the synthesis of EGFR kinase inhibitor AZD9291 using an automated flow reactor. *React. Chem. Eng.* **1,** 366–371 (2016).
9. Moore, J. S. & Jensen, K. F. Automated multitrajectory method for reaction optimization in a microfluidic system using online IR analysis. *Org. Process Res. Dev.* **16,** 1409–1415 (2012).
10. Reizman, B. J. & Jensen, K. F. Simultaneous solvent screening and reaction optimization in microliter slugs. *Chem. Commun.* **51,** 2–5 (2015).
11. McMullen, J. P., Stone, M. T., Buchwald, S. L. & Jensen, K. F. An integrated microreactor system for self-optimization of a heck reaction: From micro- to mesoscale flow systems. *Angew. Chemie - Int. Ed.* **49,** 7076–7080 (2010).
12. Bourne, R. A., Skilton, R. A., Parrott, A. J., Irvine, D. J. & Poliakoff, M. Adaptive process optimization for continuous methylation of alcohols in supercritical carbon dioxide. *Org. Process Res. Dev.* **15,** 932–938 (2011).
13. Amara, Z. *et al.* Automated Serendipity with Self-Optimizing Continuous-Flow Reactors. *European J. Org. Chem.* **2015,** 6141–6145 (2015).
14. Parrott, A. J., Bourne, R. A., Akien, G. R., Irvine, D. J. & Poliakoff, M. Self-optimizing continuous reactions in supercritical carbon dioxide. *Angew. Chemie - Int. Ed.* **50,** 3788–3792 (2011).
15. Jumbam, D. N., Skilton, R. A., Parrott, A. J., Bourne, R. A. & Poliakoff, M. The Effect of Self-Optimisation Targets on the Methylation of Alcohols Using Dimethyl Carbonate in Supercritical CO2. *J. Flow Chem.* **2,** 24–27 (2012).
16. Sans, V., Porwol, L., Dragone, V. & Cronin, L. A self optimizing synthetic organic reactor system using real-time in-line NMR spectroscopy. *Chem. Sci.* **6,** 1258–1264 (2015).
17. Fitzpatrick, D. E., Battilocchio, C. & Ley, S. V. A Novel Internet-Based Reaction Monitoring, Control and Autonomous Self-Optimization Platform

for Chemical Synthesis. *Org. Process Res. Dev.* **20,** 386–394 (2016).
18. Krishnadasan, S., Yashina, A., DeMello, A. J. & DeMello, J. C. in *Advances in Chemical Engineering* (ed. Schouten, J. C.) **38,** 195–231 (Elsevier, 2010).
19. Marler, R. T. & Arora, J. S. The weighted sum method for multi-objective optimization: New insights. *Struct. Multidiscip. Optim.* **41,** 853–862 (2010).
20. Triantaphyllou, E. & Shu, B. Multi-criteria decision making: an operations research approach. *Encycl. Electr. Electron. Eng.* **15,** 175–186 (1998).
21. Mateo, J. R. S. C. in *Multi Criteria Analysis in the Renewable Energy Industry* 19–22 (Springer London, 2012). doi:10.1007/978-1-4471-2346-0_4
22. Huyer, W. & Neumaier, A. SNOBFIT -- Stable Noisy Optimization by Branch and Fit. *ACM Trans. Math. Softw.* **35,** 1–25 (2008).
23. Dallwig, S., Neumaier, A. & Schichl, H. in *Developments in Global Optimization* (eds. Bomze, I. M., Csendes, T., Horst, R. & Pardalos, P. M.) 19–36 (Springer US, 1997). doi:10.1007/978-1-4757-2600-8_2
24. Yeniyay, Ö. Mathematical and Computational Applications ,. *Math. Comput. Appl.* **10,** 45–56 (2005).
25. Kim, K. *et al.* Facile Synthesis of o -Xylenyl Fullerene Multiadducts for High Open Circuit Voltage and Efficient Polymer Solar Cells. *Chem. Mater.* **23,** 5090–5095 (2011).
26. Kang, D. J. *et al.* Efficient light trapping in inverted polymer solar cells by a randomly nanostructured electrode using monodispersed polymer nanoparticles. *Nanoscale* **5,** 1858–63 (2013).
27. Meng, X. *et al.* Dihydronaphthyl-based [60]fullerene bisadducts for efficient and stable polymer solar cells. *Chem. Commun. (Camb).* **48,** 425–7 (2012).
28. Kang, H. *et al.* Effect of fullerene tris-adducts on the photovoltaic performance of P3HT:fullerene ternary blends. *ACS Appl. Mater. Interfaces* **5,** 4401–8 (2013).
29. Seyler, H., Wong, W. W. H., Jones, D. J. & Holmes, A. B. Continuous flow synthesis of fullerene derivatives. *J. Org. Chem.* **76,** 3551–6 (2011).
30. Martin, T. & Tan, T.-S. Anti-aliasing and Continuity with Trapezoidal Shadow Maps. in *Proceedings of the Fifteenth Eurographics Conference on Rendering Techniques* 153–160 (Eurographics Association, 2004). doi:10.2312/EGWR/EGSR04/153-160



Scheme 1: Synthesis of o-xylenyl C_{60} adducts of varying order via in-situ conversion of sultine (1,4-dihydro-2,3-benzoxathiin 3-oxide) to o-quinodimethane (oQDM) **(a)** followed by successive attachments of oQDM to C_{60} (X_0) by Diels-Alder cycloadditions **(b)**

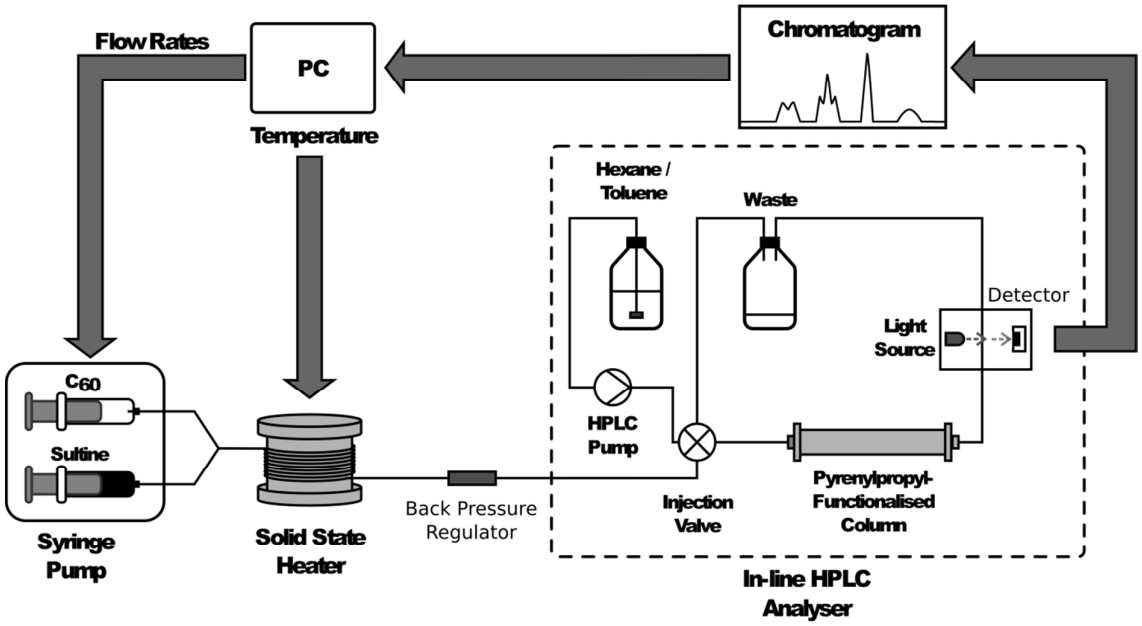


Figure 1: Schematic showing experimental set-up for manual and automated synthesis of o-xylenyl C_{60} adducts. Sultine and C_{60} solutions are separately injected into the two inlets of a y-shaped mixer, and the resulting mixture is then passed helically around a cylindrical solid-state heater to initiate the reaction prior to collection in a flask. For on-line analysis, a small amount of the product mixture is diverted to a sample coil, from where it is injected into an HPLC column using a high-pressure switching valve. Detection is carried out optically by absorption spectroscopy. The transient signal from the detector is passed to a personal computer (PC) for display and analysis. In manual mode, the flow rates and temperatures used in successive measurements are read in sequence from a pre-written file; in automated mode they are determined at runtime by an optimisation routine on the basis of previously acquired data.

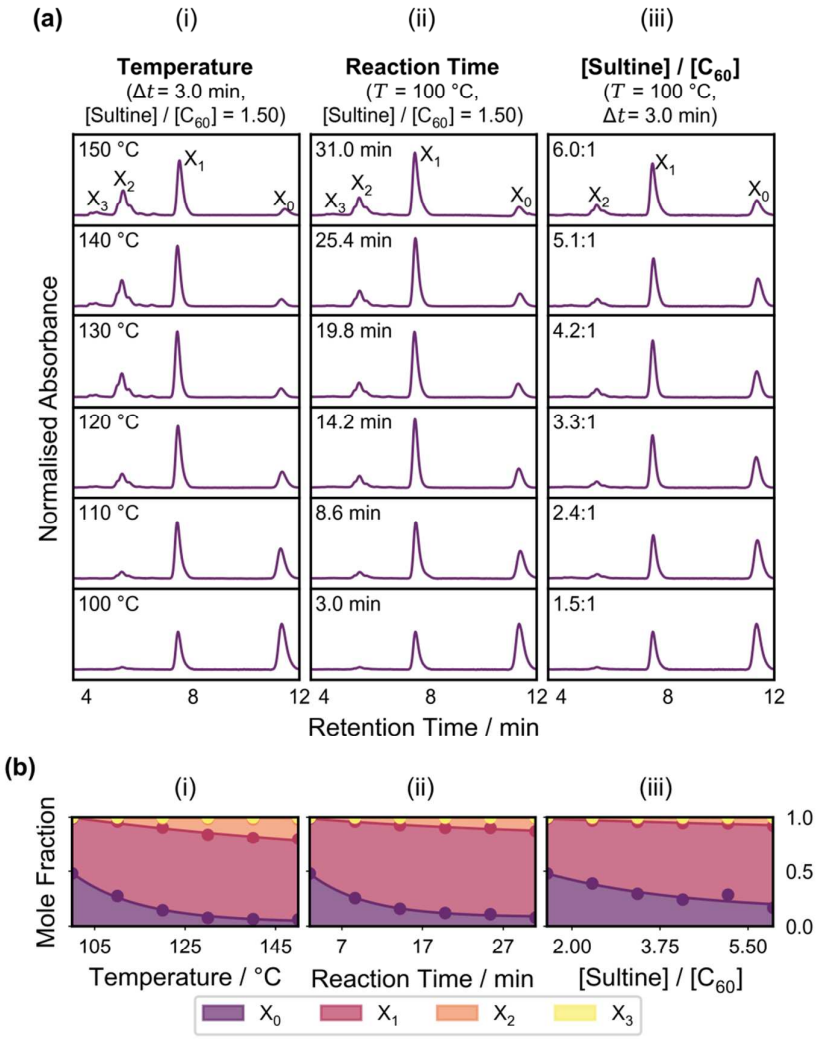


Figure 2: (a) Graphs showing chromatograms for flow-synthesised o-xylenyl fullerene adducts as a function of temperature (i), reaction time (ii), and sultine to C_{60} ratio (iii), holding in each case the other two reaction parameters fixed. The peaks corresponding to each adduct are labelled in the uppermost plots. **(b)** Stacked plots showing mole fraction distributions for the fullerene adducts versus temperature (i), reaction time (ii), and sultine to C_{60} ratio (iii), using mole fraction values extracted from the chromatograms in (a). For the reaction conditions chosen, increases in temperature, reaction time and sultine to C_{60} ratio all lead to increased conversion of C_{60} into higher order adducts, with the concentration of C_{60} decreasing and the concentration of higher order adducts increasing.

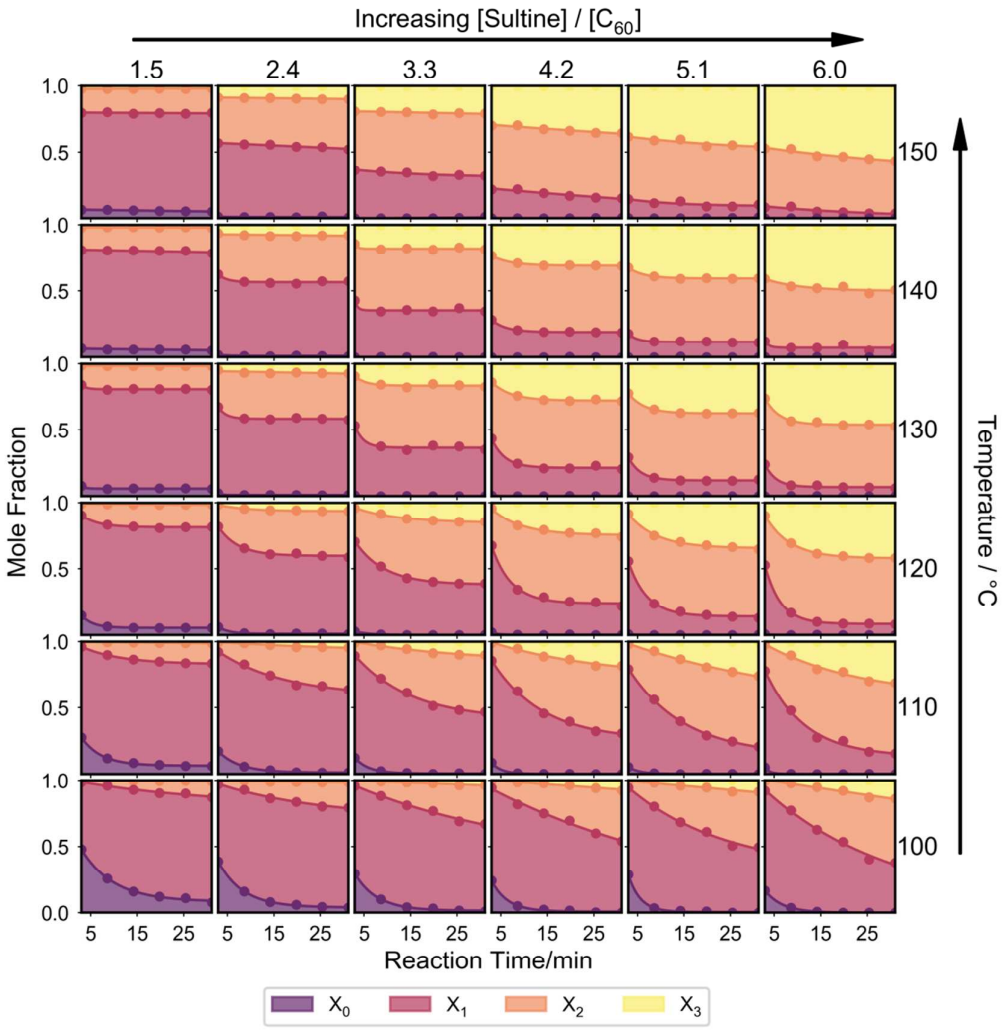


Figure 3: Stacked plots showing the effect on the mole fraction distribution of concurrently varying the temperature, reaction time, and sultine to C_{60} ratio. Each individual plot shows the variation of mole fraction distribution with reaction time at a given temperature and sultine to C_{60} ratio. Plots in the same row correspond to reactions undertaken at a common temperature, while plots in the same column correspond to reactions undertaken at a common sultine to C_{60} ratio. The measurements were carried out in a randomized order to eliminate possible bias due to system drift. Increases in the reaction time, temperature and/or sultine:C_{60} ratio cause a progressive reduction in the mole fraction of unreacted C_{60} and a progressive increase in the mole fraction of the third-order adduct; at lower temperatures the first- and second- order adducts are the dominant products, while at higher temperatures and sultine:C_{60} ratios the third-order adduct dominates.

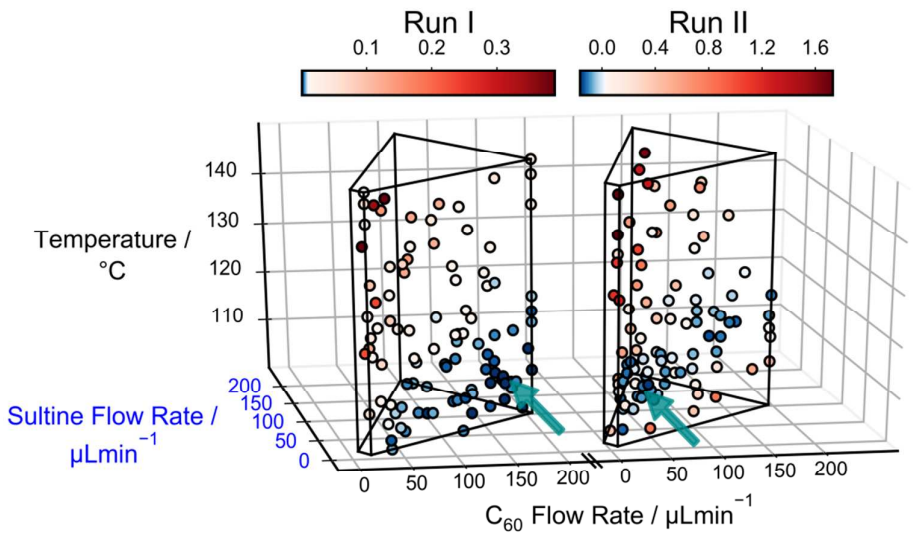


Figure 4: Scatter plots for Runs I and II, showing the influence of the sultine flow rate, the C_{60} flow rate and the temperature on the merit values f_I and f_{II}. The location of each data point indicates the reaction conditions used, while the colour denotes the corresponding merit value. For ease of interpretation, points with merit values above the median merit value have been coloured red, while those with merit values below the median value have been coloured blue; points with merit values close to the median value appear as white. The black cage defines the flow rate and temperature constraints. The arrows indicate the locations of the best point (i.e. the point with the lowest merit value) for the two runs. In Run I, the algorithm preferentially sampled the region next to the lower half of the right face of the cage, corresponding to lower temperatures and lower sultine concentrations. In Run II it preferentially sampled the same broad region, but placed a stronger emphasis on data points close to the foremost 'spine' of the cage, corresponding to longer reaction times.

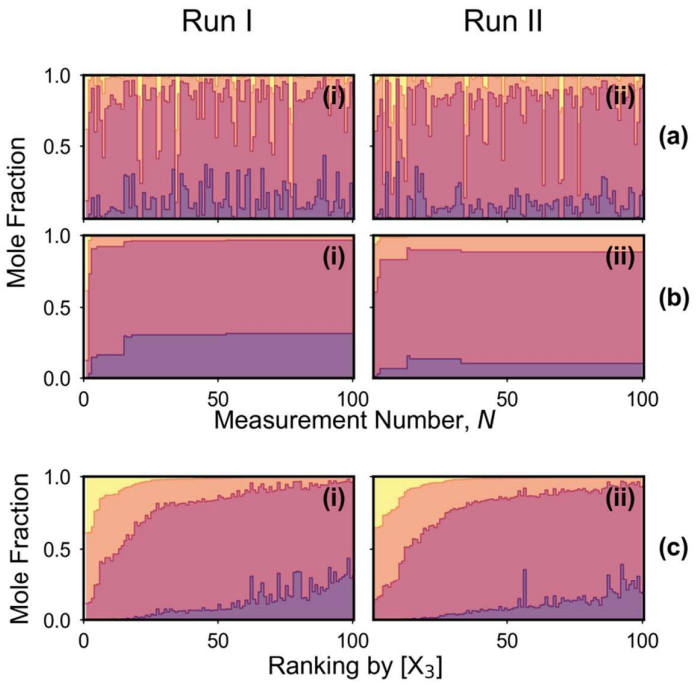


Figure 5: (a) Time series for Runs I and II, showing the mole fraction distribution of the o-xylenyl adducts versus measurement number N. **(b)** Time series showing the mole fraction distribution for the best result to date (i.e. the data point yielding the lowest merit value) versus N. In the case of Run I, the reduction in $[X_3]$ was achieved at the cost of an unwanted increase in $[X_0]$, whereas in the case of Run II $[X_0]$ was maintained at a low value as intended. **(c)** Mole fraction distributions from **(a)** arranged in order of decreasing $[X_3]$.

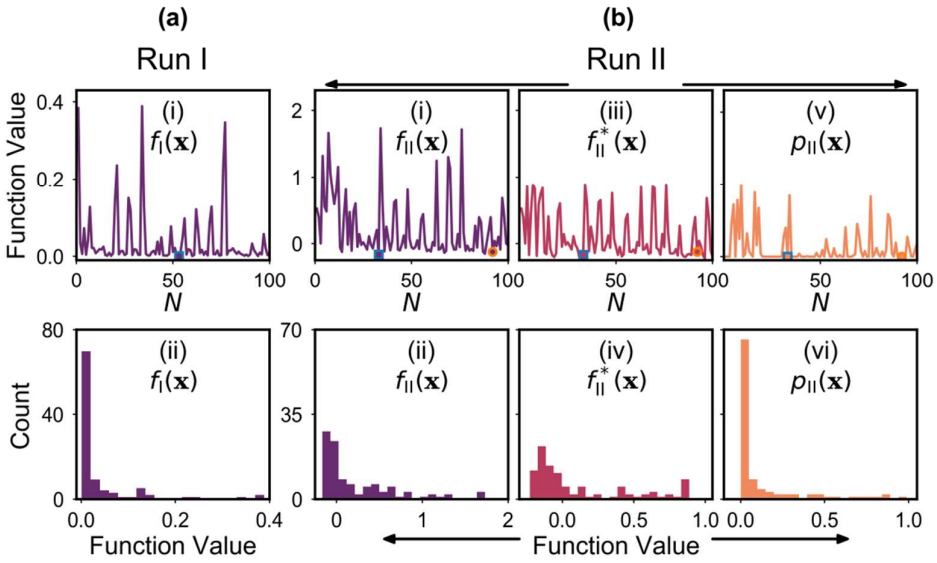


Figure 6: (a) Merit values for optimisation Run I expressed as a time series (i) and a histogram (ii). The square-shaped marker in the time-series plot corresponds to the best point. **(b)** Merit values for optimisation Run II expressed as a time series (i) and a histogram (ii). Also shown for Run II are the time series and histograms for $f_{II}^*(\mathbf{x})$ (iii, iv) and $p_{II}(\mathbf{x})$ (v, vi). The square- and circle-shaped markers in the time-series plots correspond to the best point and the best feasible point, respectively. As expected, during both runs the optimiser preferentially explored regions of the chemical parameter space that yielded low merit values. In the case of Run II, these low values were achieved through a combination of low $f^*(\mathbf{x})$ values (corresponding to points with low $[X_3]$ values) and low $p(\mathbf{x})$ values (corresponding to points that satisfied or nearly satisfied the constraint).

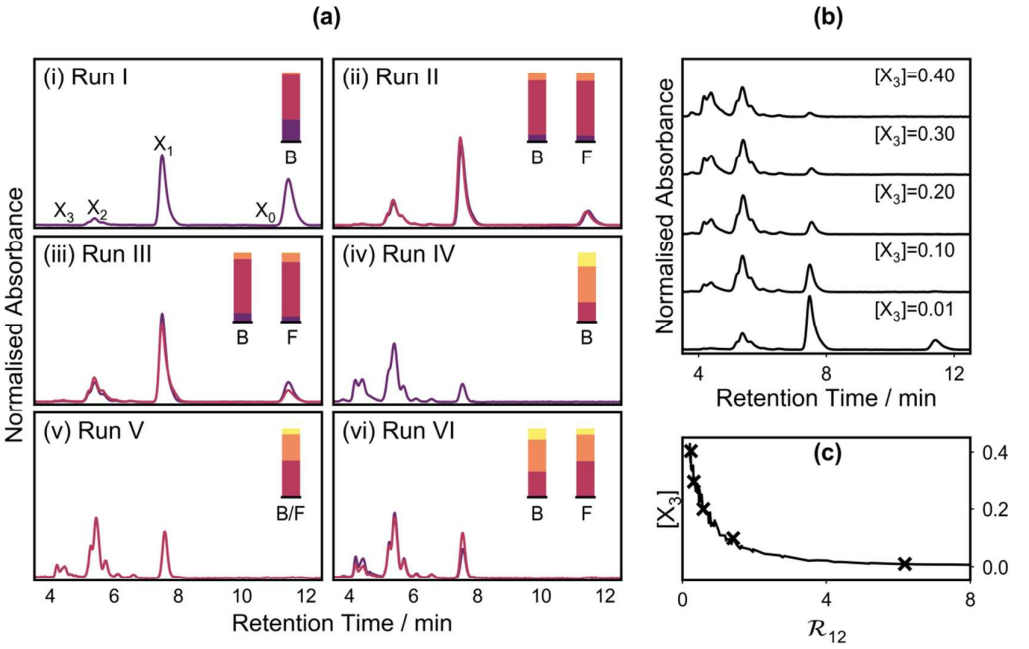


Figure 7: (a) Chromatograms for the best points (dark purple curves) and the best feasible points (red curves) obtained during the six optimisation runs. (Note: for varying reasons only a single chromatogram is shown for Runs I, IV and V: Run I was an unconstrained optimisation, meaning all points were feasible; no feasible point was found for Run IV; and the best feasible point for Run V was also the best overall point for that run). The mole fraction distributions extracted from the chromatograms are shown inset, with "B" denoting the best point and "F" denoting the best feasible point. **(b)** Illustrative chromatograms from Run IV, obtained at different values of $[X_3]$, ranging from 0.01 to 0.4. **(c)** Trade-off curve for Run IV, showing an unwanted increase in $[X_3]$ as R_{12} decreases.