Electronic Supplementary Material for Reaction Chemistry & Engineering
April 3, 2018

Supporting Information for

"Optimum catalyst selection over continuous and discrete process variables with a single droplet microfluidic reaction platform"

Lorenz M. Baumgartner,[a‡] Connor W. Coley,[a‡] Brandon J. Reizman,[a] Kevin W. Gao,[a,b] and Klavs F. Jensen*[a]

[a] Department of Chemical Engineering
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139, USA.

[b] Department of Chemical and Biomolecular Engineering
University of California Berkeley
Berkeley, CA 94702, USA.

1 Methods

1.1 Optimization algorithm

1.1.1 Derivation of the model

Following Reizman *et al.*, we derive a simple quadratic model that can be used to approximate the kinetic behavior of a general bimolecular reaction $\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{R}$ catalyzed by a transition metal complex.[1] Assuming a single rate determining step, an expression for the production rate Eq. 1 is derived from a kinetic power law. The product concentration depends on the temperature, residence time, catalyst concentration, and catalyst identity.

$$\frac{C_\mathrm{R}}{t_\mathrm{res}} \propto k_\mathrm{R}\left(\mathrm{Catalyst}, C_\mathrm{cat}, T\right) C_{\mathrm{A}_0}^{p} C_{\mathrm{B}_0}^{q} \tag{1}$$

The kinetic constant is assumed to follow an Arrhenius relationship Eq. 2 and is additionally dependent on the catalyst concentration C_cat with some reaction order r. To consider the effect of different catalysts, the pre-exponential factor is divided into a catalyst-specific prefactor A_i and an inherent prefactor A_R. Similarly, the activation energy is separated into the catalyst specific E_{A_i} and the inherent E_{A_R}.

$$k_\mathrm{R}\left(\mathrm{Catalyst}, C_\mathrm{cat}, T\right) \propto \left(C_\mathrm{cat}^{r} A_i e^{-E_{A_i}/RT}\right)\left(A_\mathrm{R} e^{-E_{A_\mathrm{R}}/RT}\right) \tag{2}$$

Substituting the kinetic constant in Eq. 1 with Eq. 2 gives an expression for the scaling of the product concentration Eq. 3. This relationship is linearized by taking the natural logarithm Eq. 4.

$$C_\mathrm{R} \propto A_i A_\mathrm{R} e^{-\left(E_{A_i}+E_{A_\mathrm{R}}\right)/RT} C_{\mathrm{A}_0}^{p} C_{\mathrm{B}_0}^{q} C_\mathrm{cat}^{r} t_\mathrm{res} \tag{3}$$

$$\begin{aligned}\log\left(C_\mathrm{R}\right) \propto \log(A_i) + \log(A_\mathrm{R}) - \frac{E_{A_i}+E_{A_\mathrm{R}}}{R}\frac{1}{T} + p\log(C_{\mathrm{A}_0}) \\ + q\log(C_{\mathrm{B}_0}) + r\log(C_\mathrm{cat}) + \log(t_\mathrm{res})\end{aligned} \tag{4}$$

To further simplify the scaling of product concentration, the initial reactant concentrations were held constant during the optimization. Rearranging Eq. 4 results in expressions for the scalings of the TON (5) and yield (6).

$$\log\left(TON\right) = \log\left(\frac{C_\mathrm{R}}{C_\mathrm{cat}}\right) \propto \log(A_i) - \frac{E_{A_i}}{R}\frac{1}{T} - \frac{E_{A_\mathrm{R}}}{R}\frac{1}{T} + (r-1)\log(C_\mathrm{cat}) + \log(t_\mathrm{res}) \tag{5}$$

$$\log\left(Y\right) = \log\left(\frac{C_\mathrm{R}}{C_{\mathrm{A}_0}}\right) \propto \log(A_i) - \frac{E_{A_i}}{R}\frac{1}{T} - \frac{E_{A_\mathrm{R}}}{R}\frac{1}{T} + r\log(C_\mathrm{cat}) + \log(t_\mathrm{res}) \tag{6}$$

We can summarize the continuous variables in the vector of unscaled continuous variables $\tilde{\mathbf{x}}$; these are scaled to the interval $[-1,+1]$ as $\mathbf{x}$. The three continuous variables are $\tilde{x}_1 = T^{-1}$, $\tilde{x}_2 = \log(C_\mathrm{cat})$, and $\tilde{x}_3 = \log(t_\mathrm{res})$. Including interaction and quadratic effects among all continuous variables and an additional coefficient to weigh the influence of the residence time, we are lead to the final semi-empirical model Eq. 7. These additional empirical model parameters can capture nonlinear effects, such as the influence of side reactions, inaccuracies in scaling relationships, more complex kinetics, and the change in reaction rate as the reagents are consumed.

$$\hat{b} = \mathbf{r}\boldsymbol{\theta} = \underbrace{\sum_{i=1}^{N_{\mathrm{dv}}} y_i \left(c_i + a_i x_1\right)}_{\text{Cat.: Const. + Lin. } T} + \underbrace{\sum_{j=2}^{N_{\mathrm{cv}}} a'_j x_j}_{\text{Lin. effects}} + \underbrace{\sum_{j=1}^{N_{\mathrm{cv}}} \sum_{k=j}^{N_{\mathrm{cv}}} a_{jk} x_j x_k}_{\text{Quad. + Inter. effects}} \tag{7}$$

Eq. 7 describes how the model response $\hat{b}$, representing the logarithmic yield or the objective function value, depends on vector of experimental settings $\mathbf{r}$ and the vector of fitting parameters $\boldsymbol{\theta}$. The first sum on the right side describes the catalyst specific offsets c_i and the catalyst specific temperature interaction a_i for each discrete variable y_i. The second sum contains the linear effects of all continuous variables a'_j except the temperature. The double sum includes the interaction and quadratic effects of all continuous variables a_{jk}. Multiple experiments can be summarized in vector notation (8).

$$\hat{\mathbf{b}} = \mathbf{X}\boldsymbol{\theta} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_k \\ \vdots \\ \mathbf{r}_{N\mathrm{exp}} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{N\mathrm{param}} \end{pmatrix} \tag{8}$$

The model response vector $\hat{\mathbf{b}}$ depends on the experimental design matrix $\mathbf{X}$ comprised of the line vectors $\mathbf{r}_k$ and the vector of model fitting parameters $\boldsymbol{\theta}$.
The model derived with this approach incorporates all discrete catalyst candidates i into a single global response surface. This allows the sharing of continuous variable parameters that are assumed to be independent of the discrete variables. Compared to an approach that performs a separate optimization for each discrete variable, using a shared response surface reduces the number of experiments required to fit the model. For example, if the linear effect of time is independent of the catalyst, then experiments run with one catalyst can provide data that inform experiments for another catalyst.

1.1.2 Execution of the optimization algorithm

For the selection of the D-optimal design, we create an 11-level full factorial design with all $N_{\mathrm{dv}} \cdot 11^{N_{\mathrm{cv}}}$ possible combinations with eleven evenly spaced factor settings for every continuous variable. The experimental design matrix of the full factorial design is created using our quadratic model with N_{param} parameters. The Matlab row-exchange algorithm `candexch()` selects an experimental design matrix $\mathbf{X}$ with $N_{\mathrm{param}} + N_{\mathrm{extra}}$ diversified settings to maximize the determinant of the Fisher information matrix $|\mathbf{X}^{\mathrm{T}}\mathbf{X}|$.

The model parameters of the logarithmic product yield $\boldsymbol{\theta}_Y$ and the objective function value $\boldsymbol{\theta}_\varphi$ are calculated with a weighted least squares regression (9).

$$\boldsymbol{\theta} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{b} \tag{9}$$

The experimental data is contained in the vector of observed values $\mathbf{b}$ whose kth element corresponds to the kth experiment $b_k = \log\left(TON_k\right)$ or $b_k = \log\left(Y_k\right)$. The normalized weighting matrix $\mathbf{W}$ is calculated with Eq. 10. It reduces the influence of experiments with relatively low yield.

$$\mathbf{W} = \frac{N_{\text{exp}}}{\sum_{i=k}^{N_{\text{exp}}} Y_k} \begin{bmatrix} Y_1 & 0 & \dots & 0 \\ 0 & Y_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & Y_{N_{\text{exp}}} \end{bmatrix} \tag{10}$$

The Matlab solver `fmincon()` is used to determine the yield and constrained TON optima of the models for each discrete candidate under consideration. We determine the maximum yield Y^* and the constrained global objective function optimum φ^*. For the hypothesis test, the discrete variable TON optima are compared against the lower bound on the 99% confidence interval of the global optimum φ^*_- estimated by (11).

$$\varphi^*_- = \varphi^* - \sigma_{\varphi^*} \cdot t_{1-\alpha,\nu} \tag{11}$$

The interval is determined using a two-sided t-test with $\alpha = 1\%$ and the degrees of freedom $\nu = N_{\text{exp}} - N_{\text{param}}$. The uncertainty σ_{φ^*} of the global optimum is calculated by taking the square root of the corresponding variance calculated with Eq. 12 following Goos and Jones[3], p. 87.

$$\sigma^2_{\varphi^*} = \sigma^2_{\varepsilon,\varphi} \mathbf{r}_{\varphi^*} \left(\mathbf{X}^{\mathrm{T}} \mathbf{W} \mathbf{X}\right)^{-1} \mathbf{r}^{\mathrm{T}}_{\varphi^*} \tag{12}$$

The vector $\mathbf{r}_{\varphi^*}$ is a potential matrix row of $\mathbf{X}$ that corresponds to the variable settings of the global TON optimum. Eq. 13 provides the weighted mean squared error of the models $\sigma^2_{\varepsilon,\varphi}$ and $\sigma^2_{\varepsilon,Y}$.

$$\sigma^2_{\varepsilon} = \frac{(\mathbf{b} - \mathbf{X}\boldsymbol{\theta})^{\mathrm{T}} \mathbf{W} (\mathbf{b} - \mathbf{X}\boldsymbol{\theta})}{N_{\text{exp}} - N_{\text{param}}} \tag{13}$$

New G-optimal experiments are determined by minimizing Eq. 14 with the Matlab solver `fmincon()`. Note that the proposed next conditions are not necessarily the experimental setting of the predicted yield or constrained TON optima.

$$G_{\text{opt},i} = \sqrt{\sigma^2_{\varepsilon,\varphi} \mathbf{r}_{\varphi^*} \left(\mathbf{X}^{\mathrm{T}}_{\text{aug}} \mathbf{W}_{\text{aug}} \mathbf{X}_{\text{aug}}\right) \mathbf{r}^{\mathrm{T}}_{\varphi^*}} + \sqrt{\sigma^2_{\varepsilon,Y} \mathbf{r}_{Y^*} \left(\mathbf{X}^{\mathrm{T}}_{\text{aug}} \mathbf{W}_{\text{aug}} \mathbf{X}_{\text{aug}}\right) \mathbf{r}^{\mathrm{T}}_{Y^*}} \tag{14}$$

Eq. 13 provides the mean squared errors of the TON and yield models $\sigma^2_{\varepsilon,\varphi}$ and $\sigma^2_{\varepsilon,Y}$. The potential experimental design matrix rows $\mathbf{r}_{\varphi^*}$ and $\mathbf{r}_{Y^*}$ correspond to the predicted optimal variable settings. The experimental design matrix $\mathbf{X}$ is augmented by a row describing the candidate conditions for the next experiment resulting in the augmented matrix $\mathbf{X}_{\text{aug}}$. To calculate the augmented weighting matrix $\mathbf{W}_{\text{aug}}$, the unknown yield of the candidate experiment is estimated with the response surface model.

1.1.3 Primary differences between MINLP 1 and MINLP 2

The previous algorithm MINLP 1 uses a fractional factorial design and second refined fractional factorial design with 16 experiments each for the initialization phase instead of the D-optimal design employed by MINLP 2. The previous algorithm uses the data of *all* catalysts (remaining and fathomed) to calculate the optima and the uncertainty of the global optimum for the hypothesis test. In contrast to the new algorithm, the uncertainty of the global optimum σ_{φ^*} is calculated with a jackknife resampling strategy and the two-sided t-test was conducted with $\alpha = 5\%$. MINLP 1 performs the hypothesis test for all discrete variables allowing candidates that have been fathomed during previous iterations to be re-added to the pool of variables under consideration. Instead of having a fathoming queue, the algorithm can skip the fathoming of a discrete variable if the number

of experiments associated with the remaining candidates would be lower than $N_{\text{param}}+1$. MINLP 1 uses the data of the remaining catalysts to calculate the optima and uncertainties for the next G-optimal experiments and the global optimum. The convergence criterium was an improvement of 2% for the *predicted* global optimum φ^* *and* the lower bound φ^*_-.[1]

1.2 Experimental validation with a Suzuki-Miyaura coupling

1.2.1 Experimental optimization with MINLP 1

The experimental optimization with MINLP 1 was performed with reagent preparation methods and the reaction system described by Reizman *et al.*[1].

1.2.2 Reagent preparation for MINLP 2 experiments

The commercially obtained chemicals are listed in Table 1. The precatalyst powders **P1L1** and **P2L1** were obtained from Sigma-Aldrich. All other precatalysts were prepared using the procedure of Bruno *et al.*[2].

Table 1 List of commercially obtained chemicals

Name of chemical	Vendor	Abbreviation
Acetone, ACS Reagent ≥99.5%	Sigma-Aldrich	-
Argon	Airgas	-
1,8-Diazabicyclo[5.4.0]undec-7-ene	Sigma-Aldrich	DBU
2-Fluoropyridine-3-boronic acid pinacol ester	AK Scientific	-
3-Chloropyridine, Reagent plus, ≥99%	Sigma-Aldrich	Reagent A
Acetonitrile solution, 0.1% (v/v) HPLC Grade	Sigma-Aldrich	ACN + 0.1% FA
Naphthalene, 99%	Sigma-Aldrich	ISTD
Tetrahydrofuran, anhydrous, 250 ppm Inhibitor, ≥99.9%	Sigma-Aldrich	THF
Water, HPLC Grade	Sigma-Aldrich	-
XPhos Pd G2	Sigma-Aldrich	-
XPhos Pd G3	Sigma-Aldrich	-

For the preparation of all solvents, nitrogen was bubbled through the liquid in 50 mL vials with rubber stoppers while sonicating under vacuum for 10 min to remove oxygen. As a general procedure, chemicals were weighed into volumetric flasks (Sartorius MC21S Micro Balance, 0.001 mg). The flasks were sealed with rubber stoppers and the gas space exchanged three times by applying vacuum and backfilling with nitrogen. Deoxygenated THF was added up to the mark using a syringe and the solution sonicated. The base solution was withdrawn with a glass syringe (Hamilton Gas Tight Series). For the reagent solutions, glass vials were sealed with silicone septum caps and the gas space exchanged. The solutions from the volumetric flasks were transferred to the sealed glass vials with a syringe.

For the reagents 2-fluoropyridine-3-boronic acid pincacol ester (0.996 M) and 3-chloropyridine (1.434 M), 5 mL stock solutions were prepared in THF and transferred to flat bottom glass vials. The 3-chloropyridine stock solution was prepared with the internal standard naphthalene at a concentration of 81.15 g L^{-1}. The pre-synthesized precatalyst powders were used to prepare stock solutions with a concentration of 0.017 M, which were transferred into tappered 2 mL glass vials.

All stock solutions and solutions of the deoxygenated solvents water and THF were placed into the vial rack of the liquid handler. Argon was feed into a manifold with a flow rate of 10 mL min^{-1} and distributed to every vial using a 0.03 in inner diameter PFA tubing with syringe needles (BD PrecisionGlide, 20 G x 1 1/2, 0.9 mm x 40 mm). A vent needle was added to the each vial to equilibrate it with atmospheric pressure (Fig. 1). Catalyst stock solutions were replaced after 14 h of operation to prevent changes in reactivity due to evaporation. A DBU

base solution of 1.645 M in THF was prepared in a 10 mL volumetric flask and transferred to a 2.5 mL glass syringe with LUER lock adapter. The quench solution was prepared mixing equal volumes of deoxygenated acetone and water in a 50 mL vial and withdrawing them with a 10 mL glass syringe.

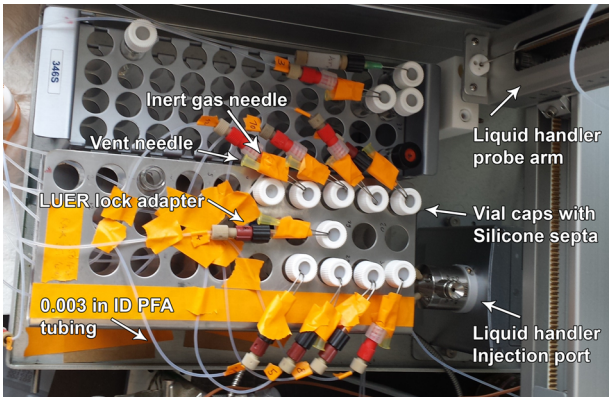

Fig. 1 Detailed view of the vials in the liquid handler attached to inert gas manifold.

1.2.3 Microfluidic Reaction Platform for MINLP 2 experiments

An overview of the microfluidic reaction platform is depicted in Fig. 2. A flow diagram is given in Fig. 3.

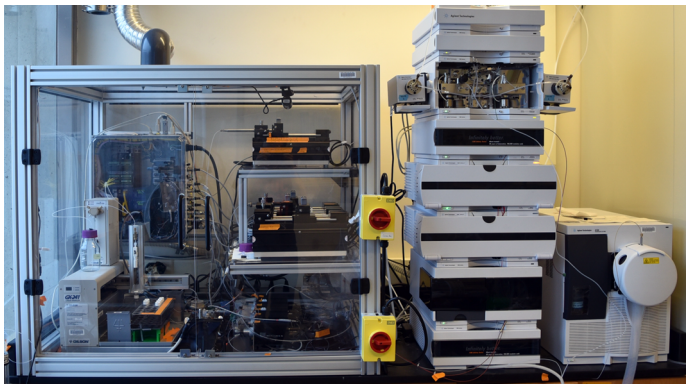

Fig. 2 Detailed view of the open oscillatory flow reactor (aluminum cover removed).

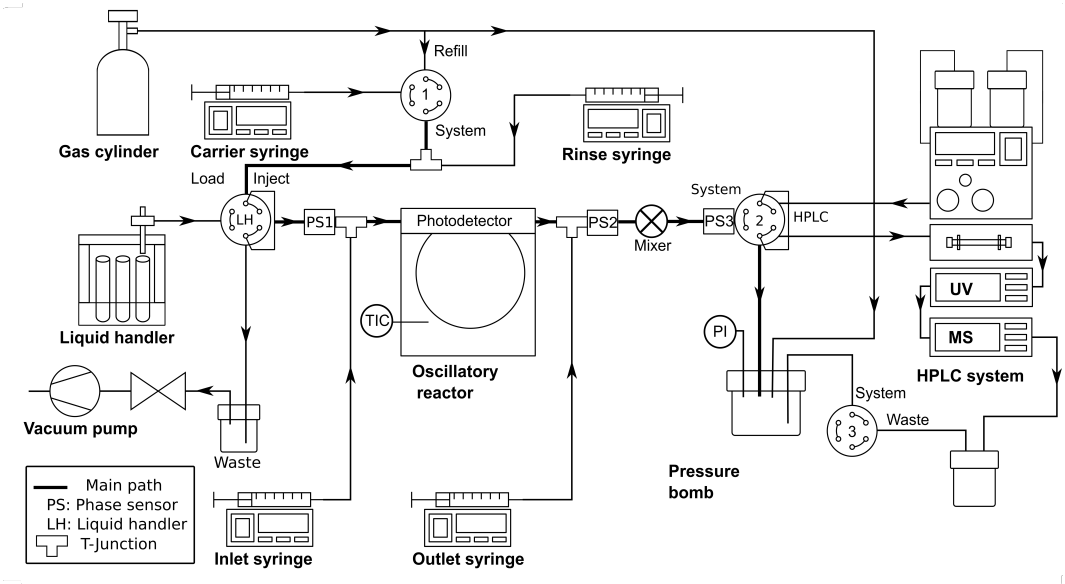

Fig. 3 Flow diagram of microfluidic segmented flow reaction platform with oscillatory flow reactor.

Configuration of components

All connections were made with super flangeless PEEK fittings. The carrier syringe pump (8 mL stainless steel, Harvard Apparatus) is connected to the 2-way 4-port PEEK valve 1 (C2-2344D EUHA, VICI Valco) using 0.03 in inner diameter (ID) PFA tubing. This configuration allows the syringe to be refilled from the gas cylinder position and to infuse into the system position. The 0.03 in ID line in system position is connected to a PEEK T-junction with 0.02 in ID attached to the rinse syringe and leads to the 6-way stainless steel valve of the liquid handler (GX 241, Gilson). All other system components are connected with 0.02 in ID PFA tubing and super flangeless PEEK fittings.

The liquid handler needle is mounted to the probe arm (Fig. 4) used to aspirate sample volumes from stock solutions and solvents from glass vials. We modified the liquid handler (LH) by adding a 5 cm long piece of 1/16 in ID PFA tubing above the needle. This section with a larger diameter helps to eliminate gas bubbles between volume fractions and improves mixing. The 0.02 in PFA tube attached with a PEEK union (P-702, IDEX

Health & Science) on top contains the transfer fluid THF. It is connected to the liquid handler pump using a 100 µL glass syringe.

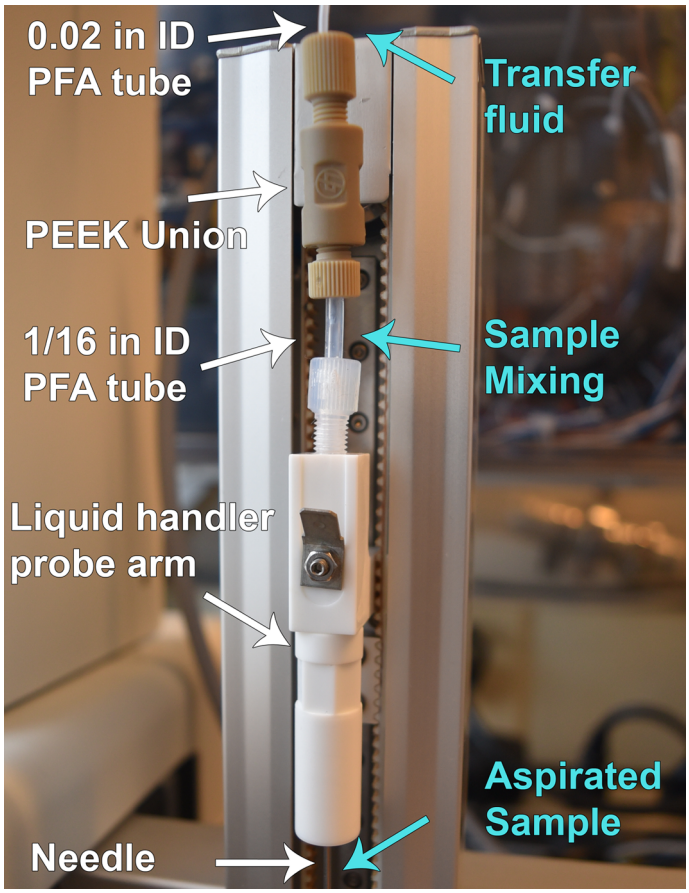

Fig. 4 Detailed view of sample mixer on liquid handler z-axis arm.

The GX Direct Injection Module is connected to a waste bottle with vacuum pump used to drain the injection port (Fig. 3). The inlet and outlet syringe pumps are attached to the main path with additional T-junctions before and after the reactor. They are equipped with the phase sensors PS1 & PS2 (OPB350L062Z, TT Electronics) to time online injections. The reactor (Fig. 5) consists of 1/16 in inner diameter PFA tubing in an aluminum case equipped with two cartridge heaters and a thermocouple.
The main flowpath (Fig. 3) leads from the reactor outlet to the mixer comprised of larger 1/16 in ID PFA tubing connected with two PEEK unions . It eliminates small gas bubbles and improves mixing of the quench solution and reaction droplets. The 2-way 6-port PEEK valve 2 (C2-1346D EUHA, VICI Valco) and phase sensor PS3 are used to transfer samples into the flow path of the HPLC. The main flowpath ends at the pressure bomb with provides the system backpressure and is used to collect waste. The 2-way 4-port PEEK valve 3 (C2-2344D EUHA, VICI Valco) is set to system position during operation. It is switched to the waste position to empty the pressure bomb and vent the reaction system and the end of a reaction campaign.

System startup

The system is turned on and the hardware communication with the virtual instrument in Labview is started. The HPLC is set up to trigger the start of the method remotely by a relay controlled with Labview. Valve 1 is switched to the refill position. The liquid handler valve is set to the inject position. The HPLC valve 2 and waste valve 3 are set to system position. After the user pressurized the system with 100 psig nitrogen, valve 1

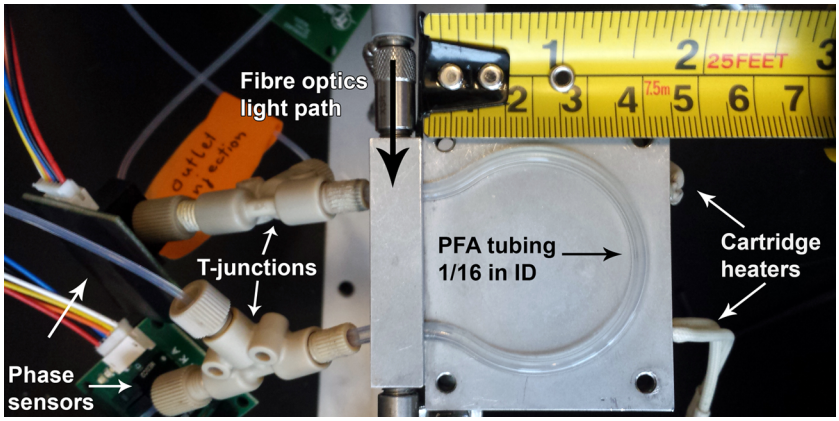

Fig. 5 Detailed view of the open oscillatory flow reactor (aluminum cover removed).

is switched to system position. In the next step, the user primes the rinse, inlet and outlet syringe. The liquid handler primes the tubing with the transfer liquid THF and rinses the injection port completing the startup procedure.

Droplet preparation

Before the preparation of each reaction droplet, four THF rinse slugs with a volume of 15 µL each are injected to clean the tubing and reduce the carry-over between experiments. The liquid handler inserts the needle into a vial with inert gas and withdraws a gas buffer of 20 µL. Next, it aspirates the components necessary to mix a 40 µL reaction droplet into the needle. The probe tip is dipped into a vial with THF between sampling from vials to avoid cross-contamination. In order to mix all components withdrawn into the needle, the probe tip is inserted into the inert gas vial. The syringe pump of the liquid handler withdraws the sample into the mixer on top and infuses it back into the needle four times to eliminate gas bubbles and mix the components to a homogeneous solution. After switching the LH valve into load position, the liquid handler inserts the needle into the injection port and loads the prepared sample into the the 15 µL sample loop.

If the all rinse slugs have left the system, the reactor has reached its target temperature and the previous HPLC run is going to be completed on time for the next injection, the LH valve is switched into inject position transferring the reaction droplet into the system. During the reaction the liquid handler rinses the injection port and cleans the needle. The vacuum pump is used to remove the rinse liquid from the injection port into the waste bottle (Fig. 3).

Base injection, Reaction and Quench

The reaction slug is moved to the inlet T-junction (P-715, IDEX Health & Science) by the carrier gas at a flow rate of 60 µL min^{-1}. As soon as the slug has passed PS1, the inlet syringe injects 3.5 µL of the base solution into the reaction slug to establish a DBU concentration of 0.333 M (2 equiv.) leading to the formation of the active catalyst complex in mixture.

The slug gets detected by the photodetector while entering the reactor. The carrier gas syringe sets an oscillatory flow rate of about 200 µL min^{-1} to 360 µL min^{-1}. The flow direction is reversed as soon as the slug is detected at the reactor outlet keeping the slug inside the reactor. The flow rate and the number of oscillations depend on the temperature and the target residence time.

After the final oscillation, the gas flow rate is reduced to 80 µL min^{-1} to make the timing of the quench injection reliable. As soon as the slug leaves the reactor and is detected at PS2, a quench solution volume equal to the

reaction slug volume (prepared slug volume + base injection volume) is injected into the reaction slug. The carrier gas flow rate is increased to 300 µL min^{-1} to speed up the transfer to the mixer. The temperature set point of the reactor is adjusted to the next reaction.

Analysis

Phase sensor PS3 is used to detect the quenched reaction slug at the inlet of sample loop. As soon as the droplet has passed PS3, the valve is switched to HPLC position for a sampling time of 100 ms. This corresponds to a sampled volume of 0.2 µL at a HPLC method standby flow rate of 0.1 mL min^{-1}. After switching the valve back to the system position, the HPLC method is started remotely. Only a part of the 1.7 µL sample loop is used for the HPLC analysis because a lower sample amount is advantageous for the separation. The system starts with the preparation of the next reaction slug.

When the HPLC analysis is finished, Chemstation automatically integrates the resulting chromatogram and saves a report Excel sheet under a user specified path. The Labview codes checks this directory for the newest file and uses a Matlab function to import the peak area and calculate the reaction yield and objective function. In case no product peak was detected, a value of 0.000001 was assigned to yield and objective function value. This avoids undefined values due to the logarithmic scaling in the case of a 0% reaction yield.

1.2.4 MINLP 2 HPLC analysis

The reaction droplets were analyzed with an Agilent LC/MS system (Table 2). The Suzuki-Miyaura cross-coupling product **11** was synthesized and purified to calibrate the UV detector.

Table 2 Components of Agilent LC/MS system

Product	Description
G4225A 1260 HiP Degasser	Inline Degasser
G1312B 1260 BinPump	Binary pump
G1315D 1260 DAD VL	UV Detector
G7116 B 1290 MCT	Column thermostat
Eclipse Plus C8 RRHT 1.8 μm	Column (4.6 mm) x 50 mm
G1329B 1260 ALS	Autosampler

Product purification

The product 2-fluoro-3,3'-bipyridine **11** was synthesized in batch following the procedure of Reizman *et al.* [1]. A magnetic stir bar, SPhos Pd G2 (72 mg, 0.10 mmol, 0.05 equiv.), THF (8 mL) and water (2 mL) was added to a dry and nitrogen-filled 20-mL septum vial under nitrogen atmosphere. Using syringes, 3-chloropyridine **9** (190 µL, 2.0 mmol, 1.0 equiv.) and DBU (598 µL, 4.0 mmol, 2.0 equiv.) were added sequentially. The reaction mixture was then heated to 65 °C, followed by addition of the THF (1 mL) solution of 2-fluoropyridine-3-boronic acid pincol ester **10** (669 mg, 3.0 mmol, 1.5 equiv.). The reaction mixture was allowed to stir overnight. The next day, the reaction was diluted with ethyl acetate, washed with brine and dried over Na_2SO_4, filtered and concentrated under reduced pressure. The resulting residue was then purifed by flash column chromatography (ethyl acetate/heptane = 1:1) to afford the desired product **11** (330 mg, 95% yield) as a white solid. The purity was confirmed with LC/MS (m/z = 174.06, Fig. 6) and NMR (Fig. 7 and Fig. 8). ^{1}**H NMR** (400 MHz, $CDCl_3$) δ 8.86-8.84 (m, 1H), 8.70 (dt, J = 4.8, 1.2 Hz, 1H), 8.32-8.30 (m, 1H), 7.99-7.94 (m, 1H), 7.93-7.91 (m, 1H), 7.50-7.46 (m, 1H), 7.37 (ddt, J = 7.3, 4.9, 1.3 Hz, 1H).

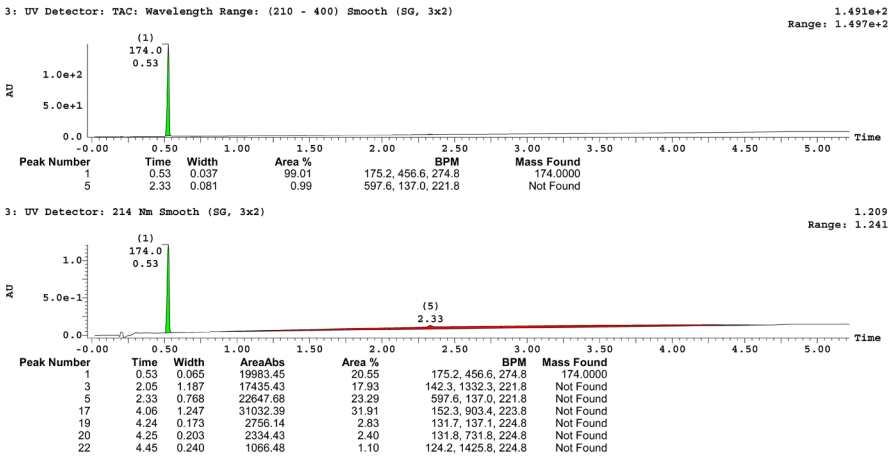

Fig. 6 LC/MS purification report shows pure 2-fluoro-3,3'-bipyridine product **11** (m/z = 174.06).

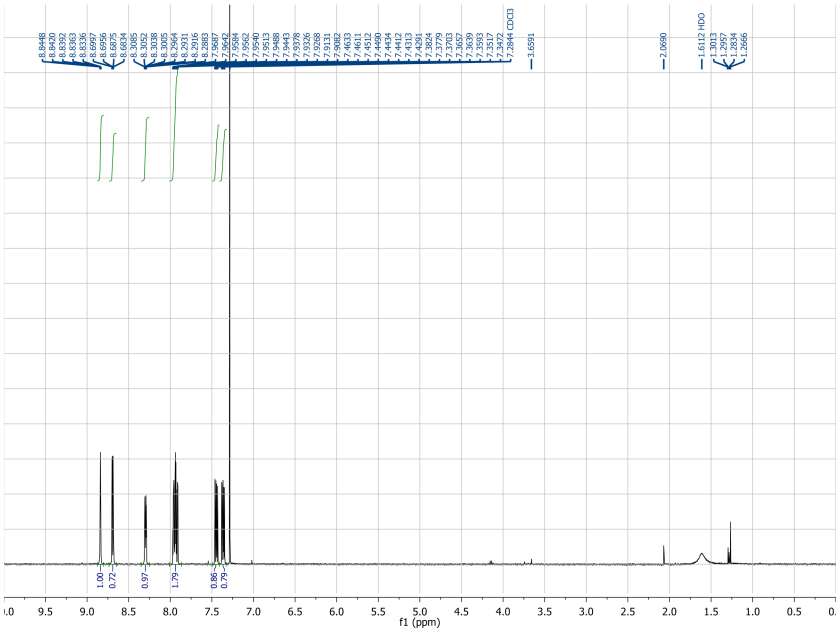

Fig. 7 NMR spectrum of 2-fluoro-3,3'-bipyridine product **11**.

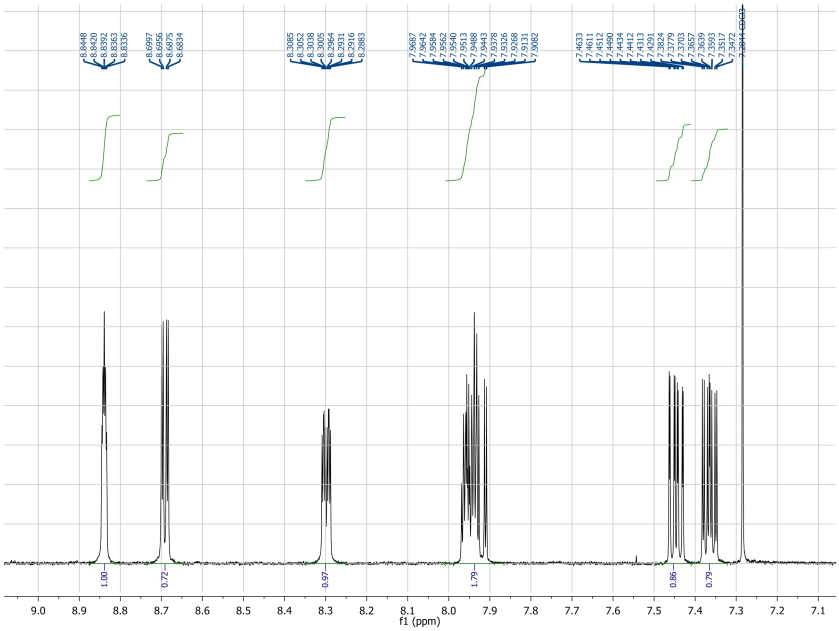

Fig. 8 Detailed NMR peaks of 2-fluoro-3,3'-bipyridine product **11** spectrum.

UV detector calibration

Standards were prepared with volumetric flasks and a micro balance (Sartorius MC21S Micro Balance, 0.001 mg). A make-up solution of THF with naphthelene as an internal standard was prepared in a 5 mL flask. It was diluted 1:1 with a 1:1 THF:water (by volume) solution resulting in an internal standard concentration of 4.676 g L^{-1}. Flasks of 2 mL or 1 mL volume were charged with the purified product and filled to the mark with make-up solution.

The reversed phase HPLC method was carried out at a column temperature of 50 °C. The separation used a gradient of water with 0.1% formic acid (solvent A) and acetonitrile with 0.1% formic acid (solvent B) requiring a total run time of 6 min. The timetable is given in Table 3.

Table 3 Timetable of HPLC method - MINLP 2

Time in min	Solvent A in vol%	Solvent B in vol%	Flow in ml min^{-1}
Standby	95.0	5.0	0.1
0.10	95.0	5.0	1.3
1.80	95.0	5.0	1.3
3.30	35.0	65.0	1.3
3.55	35.0	65.0	1.3
4.30	0.0	100.0	1.3
4.80	0.0	100.0	1.3
4.81	95.0	5.0	1.4
6.00	95.0	5.0	1.4

The calibration curve was determined with triple injections of 0.1 µL standard per concentration level (Fig. 9). More detailed information about the standard preparation and resulting calibration curve is available in the file "Baumgartner2018_calibration.xlsx".

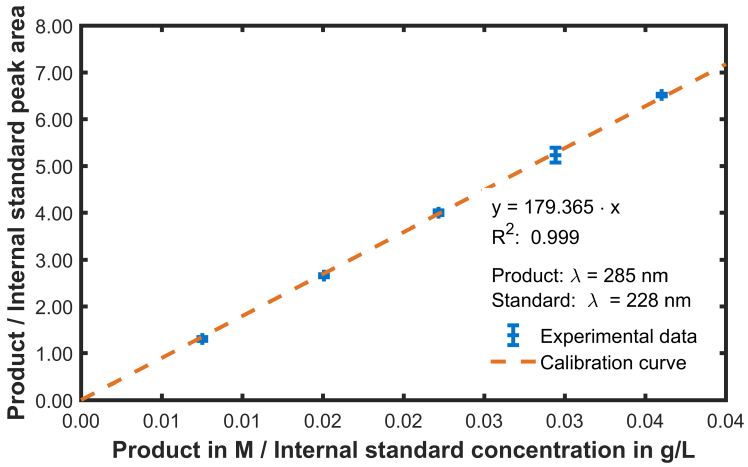

Fig. 9 MINLP 2: 2-fluoro-3,3'-bipyridine product **11** calibration curve for UV detector.

1.2.5 MINLP 1 HPLC analysis

The reversed phase HPLC method was carried out at a column temperature of 40 °C. The separation used a gradient of water with 0.1% formic acid (solvent A) and acetonitrile with 0.1% formic acid (solvent B) requiring a total run time of 9 min. The timetable is given in Table 4.

The experimental optimization with MINLP 1 was performed without calibrating the UV detector using the ratio

Table 4 Timetable of HPLC method - MINLP 1

Time in min	Solvent A in vol%	Solvent B in vol%	Flow in ml min $^{-1}$
Standby	95.0	5.0	0.5
0.10	95.0	5.0	3.5
2.00	95.0	5.0	3.5
3.00	35.0	65.0	3.5
3.75	35.0	65.0	3.5
4.50	0.0	100.0	3.5
6.00	0.0	100.0	3.5
7.00	95.0	5.0	3.5
8.50	95.0	5.0	3.5
9.00	95.0	5.0	0.5

between product peak and internal standard peak in place of the actual product concentration. After the optimization, we determined a calibration curve (Fig. 10) by matching experiments with equivalent experimental settings. It is used to calculate the unknown yield and TON with the known peak ratio for MINLP 1 and the product yields of the MINLP 2 experiments. More details are provided in the file "Baumgartner2018_calibration.xlsx". The uncertainty of the calculated yield and TON (single standard deviation) was determined with Eq. 12. The uncertainty maximum was ±6% for the product yield and ±8 for the catalyst TON. The uncertainties of individual results are listed in the file "Baumgartner2018_experimental_data.xlsx" in the sheet for MINLP 1.

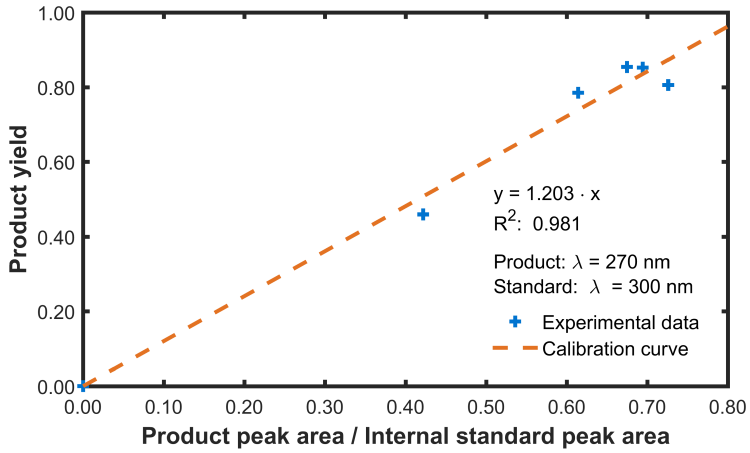

Fig. 10 MINLP 1: 2-fluoro-3,3'-bipyridine product **11** calibration of UV detector peak ratios with known yields for matching experiments of MINLP 2.

2 Results

The raw data of the simulations and experiments are available in the files "Baumgartner2018_simulation_data.xlsx" and "Baumgartner2018_experimental_data.xlsx". The experimental optimization conditions and response values for MINLP 1 and MINLP 2 are visualized in Fig. 11.

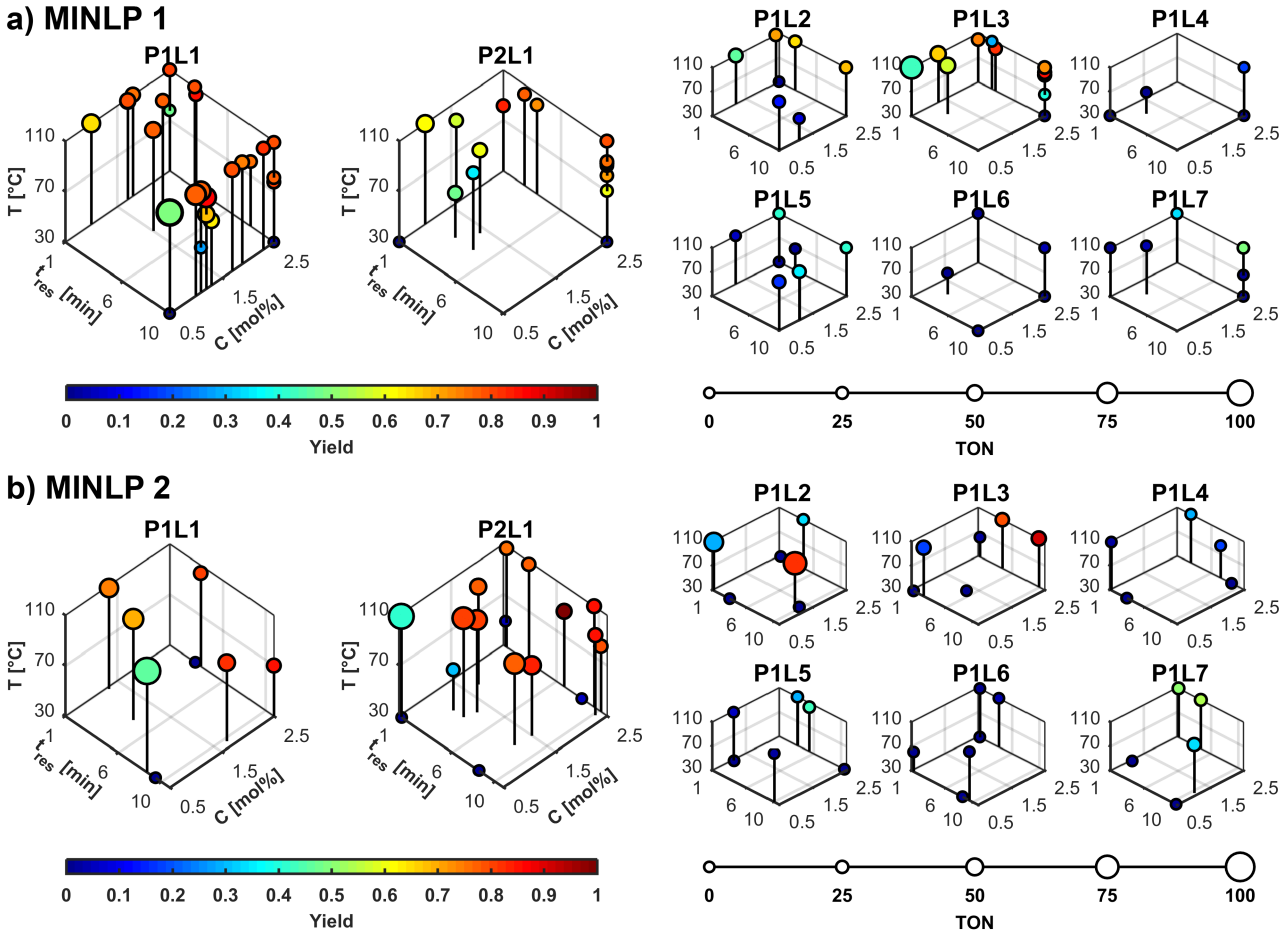

Fig. 11 Comparison of experimental conditions between previous algorithm MINLP 1 (a) and new algorithm MINLP 2 (b).

Nomenclature

Symbol	Definition
$A_\mathrm{R}, A_{\mathrm{S}_1}, A_{\mathrm{S}_2}$	Pre-exponential factor for product R, S_1 or S_2
a_i	Catalyst specific temperature effect
a'_j	Linear effects
a_{jk}	Quadratic and interaction effects
A_i	Catalyst specific pre-exponential factor
$1-\alpha$	Confidence level
$\hat{\mathbf{b}}$	Model response vector
$\mathbf{b}$	Observed response vector
C_cat	(Pre-) Catalyst concentration
$C_\mathrm{R}, C_{\mathrm{S}_1}, C_{\mathrm{S}_2}$	Product R, S_1 or S_2 concentration
$C_{\mathrm{A}_0}, C_\mathrm{A}$	(Initial) Reagent A concentration
$C_{\mathrm{B}_0}, C_\mathrm{B}$	(Initial) Reagent B concentration
c_i	Catalyst specific offset
γ	Yield criterion
$E_{A_\mathrm{R}}, E_{A_{\mathrm{S}_1}}, E_{A_{\mathrm{S}_2}}$	Activation energy for product R, S_1 or S_2
E_{A_i}	Catalyst specific activation energy
$G_{\mathrm{opt},i}$	G-optimality criterion for discrete candidate i
$k_\mathrm{R}, k_{\mathrm{S}_1}, k_{\mathrm{S}_2}$	Kinetic konstant for R, S_1 or S_2 formation
N_exp	Number of experiments
N_cat	Number of catalyst candidates
N_cv	Number of continuous variables
N_dv	Number of discrete variables
N_extra	Number of additional experiments
N_para	Number of model parameters
ν	Degrees of freedom
p, q, r	Power law exponents
R	Ideal gas constant
$\mathbf{r}$	Row describing variable settings in experimental design matrix $\mathbf{X}$
$\mathbf{r}_{\varphi^*}$, $\mathbf{r}_{Y^*}$	$\mathbf{X}$ row for predicted optimal variable settings
σ^2_ε	Mean squared error of model
σ_{φ^*}	Uncertainty of global φ optimum
T	Reaction temperature
t	t-value of student's t distribution
t_res	Residence time in reactor
$\boldsymbol{\theta}$	Model parameter vector
TON	Catalyst turnover number
$\mathbf{x}$, x_j	Vector of scaled continuous variables x_j
Y	Reaction product yield R
$\mathbf{y}$, y_j	Vector of continuous variables y_j
$\mathbf{W}$, $\mathbf{W}_\mathrm{aug}$	Normalized weighting matrix, Augmented $\mathbf{W}$
$\mathbf{X}$, $\mathbf{X}_\mathrm{aug}$	experimental design matrix , Augmented $\mathbf{X}$
$\mathbf{X}_\mathrm{check}$	Reduced $\mathbf{X}$ to check fathoming of candidate
$\tilde{\mathbf{x}}$, $\tilde{x_j}$	Vector of unscaled continuous variables
φ	Objective function value
φ^*_-	Lower boundary of confidence interval on φ^*

Bibliography

1 B. J. Reizman, Y.-M. Wang, S. L. Buchwald and K. F. Jensen, *Reaction Chemistry & Engineering*, 2016, **1**, 658–666.

2 N. C. Bruno, M. T. Tudge and S. L. Buchwald, *Chemical science*, 2013, **4**, 916–920.

3 P. Goos and B. Jones, *Optimal design of experiments: a case study approach*, John Wiley & Sons, 2011.