12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

Solving for Exact Designs in Optimal Experiment Campaigns under Uncertainty ⋆

Marco Sandrin [*,**] **Kennedy P. Kusumo** [*]
Constantinos C. Pantelides [*,**] **Benoît Chachuat** [*]

[*] *The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, UK*
[**] *Siemens Process Automation Software, London, UK*

Abstract: Applying model-based design of experiments to compute maximally-informative campaigns with multiple parallel runs is challenging. Effort-based methods can overcome some of these challenges through discretizing the experimental space into a finite set of candidate experiments, then applying convex optimization techniques to determine the optimal efforts for each candidate, and finally rounding the efforts to integer numbers of runs for a target experimental campaign size. For small experiment campaigns in particular, the final rounding can result in large suboptimality. This paper presents an approach to solving the exact design problem, where the effort variables being optimized are constrained to taking integer values. We consider model parametric uncertainty and formulate risk-inclined, risk-neutral and risk-averse exact design problems as mixed-integer nonlinear programs (MINLPs) with convex participating functions. We demonstrate the tractability of an outer-approximation algorithm to solve such MINLPs to global optimality on a case study involving the exothermic esterification of propionic anhydride with over 1000 experiment candidates and 100 uncertainty scenarios.

Keywords: Optimal experiment design, Model-based design of experiments, Experiment campaign, Mixed-integer optimization, Uncertainty

1. INTRODUCTION

Model-based design of experiments (MBDoE) for parameter precision can greatly accelerate the development of predictive mechanistic models (Franceschini and Macchietto, 2008). A particular focus in the literature has been on sequential MBDoE, where (possibly dynamic) experiments are designed one-at-a-time using gradient-based techniques to maximize a chosen information criterion. For nonlinear models, the resulting optimization problems are typically nonconvex, and thus prone to converging to local optima. Numerical failure caused by singular information matrices is also commonplace. These computational challenges hinder the design of experimental campaigns comprising multiple parallel runs.

Effort-based methods (Fedorov and Leonov, 2014; Kusumo et al., 2021; Vanaret et al., 2021) can overcome these challenges through discretizing the experimental space into a finite set of candidate experiments. These methods determine the fraction of the total number of experiments (the *effort*) to be associated with each candidate, with a view to selecting the combination of experiments having maximal information content. Treating the efforts as continuous variables leads to convex problems that can be solved efficiently using nonlinear programming techniques. This convexity is even preserved when accounting for uncertain model parameters, either in risk-neutral formulations (e.g., maximizing the average information) or in risk-averse formulations (e.g., maximizing the conditional-value-at-risk (CVaR) of information) (Kusumo et al., 2022b). Current practice entails solving such convex problems in a first step, which typically results in fractional values for the optimized efforts, then applying a rounding procedure to obtain integral efforts after specifying the total number of experiments in the campaign. But such *a posteriori* effort rounding could result in large suboptimality, especially for campaigns with a small number of experiments.

The focus of this article is on tackling the exact design of experiment campaigns, where the efforts are directly treated as integer decision variables in the optimization problem. Such problems can be formulated as mixed-integer nonlinear programs (MINLPs) with convex participating functions. The use of branch-and-bound has long been proposed to tackle these MINLPs (Welch, 1982), however this approach can result in enumerating a large number of candidate designs. Exact designs can also be formulated as mixed-integer quadratically-constrained programs (MIQCPs) (Sagnol and Harman, 2015), but this requires the introduction of a large number of auxiliary variables and can lead to memory storage issues and long computational times in cases involving a large number of candidate experiments. Herein, we investigate the

⋆ This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 955520 (Digitalgaesation). KPK, CCP and BC gratefully acknowledge funding by Eli Lilly & Company through the Pharmaceutical Systems Engineering Lab (PharmaSEL) and by the Engineering and Physical Sciences Research Council (EPSRC) as part of its Prosperity Partnership Programme under grant EP/T518207/1. Correspondence should be directed to Benoît Chachuat: b.chachuat@imperial.ac.uk

tractability of a tailored outer-approximation algorithm to solve these MINLPs to global optimality (Sec. 3). We consider risk-inclined, risk-neutral and risk-averse exact design problems (Sec. 2), and we use the esterification of propionic anhydride as an illustrative case study (Sec. 4).

2. PROBLEM DEFINITION

Consider a system with n_x experimental controls $\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^{n_x}$ and n_y measured responses $\boldsymbol{y} \in \mathbb{R}^{n_y}$,

$$\boldsymbol{y} = \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}) + \boldsymbol{\epsilon}, \tag{1}$$

where $\boldsymbol{\theta} \subset \mathbb{R}^{n_\theta}$ are uncertain parameters in the mathematical model $\boldsymbol{\eta}$. For simplicity, the measurements are assumed to be independent, and the measurement error $\boldsymbol{\epsilon} \in \mathbb{R}^{n_y}$ is assumed to have zero mean $\mathbb{E}(\boldsymbol{\epsilon}) = \mathbf{0}$ with uncorrelated and homoscedastic covariance Σ_y. Although written in closed form in Eqn. (1), the model $\boldsymbol{\eta}$ may also be defined implicitly via a set of algebraic and/or differential equations without the loss of generality.

We consider an experimental campaign comprising $N_t > 0$ experimental runs for the purpose of generating data for the estimation of the model parameters $\boldsymbol{\theta}$. Since such campaigns often consist of repeated runs with identical experimental controls (replicates), it is convenient to denote an experimental design ξ as

$$\xi \doteq \left\{ \begin{array}{c} \boldsymbol{x}_1 \ \dots \ \boldsymbol{x}_{N_c} \\ p_1 \ \dots \ p_{N_c} \end{array} \right\}, \tag{2}$$

where $N_c \leq N_t$ is the number of distinct runs, and $p_i \in (0, N_t]$ is the number of replications, or *effort*, of the i-th experimental candidate with controls $\boldsymbol{x}_i$ for each $i = 1, \dots, N_c$, so that the p_i's add up to N_t. The set $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_{N_c}\}$ is called the *support* of the experimental design ξ, denoted by $\text{supp}(\xi)$.

Designs having integral efforts, $p_i \in \mathbb{Z}_+$, are called *exact* designs. In contrast, a *continuous* design allows p_i to vary continuously in the standard simplex $\sum_i p_i = N_t$ with $p_i > 0, \forall i$. The distinction between continuous and exact designs generally becomes less important in practice as N_t grows larger.

An optimal design ξ^* is one that maximizes some scalar information criterion ϕ,

$$\xi^* \in \arg\max_{\xi} \phi(\xi), \tag{3}$$

applicable to both continuous and exact designs. Determining such an optimal design entails searching over all possible number of supports $N_c \leq N_t$, the experimental controls $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{N_c} \in \mathbb{R}^{n_x}$, and the corresponding efforts $p_1, \dots, p_{N_c} \in (0, N_t]$ simultaneously.

Herein, our focus is on computing experimental campaigns which are optimal for model calibration. The information criterion ϕ is expressed as a function of the Fisher information matrix (FIM), $\mathbf{M} \in \mathbb{R}^{n_\theta \times n_\theta}$ given by (Atkinson et al., 2007)

$$\mathbf{M}(\xi, \boldsymbol{\theta}) \doteq \sum_{i}^{N_c} p_i \mathbf{A}(\boldsymbol{x}_i, \boldsymbol{\theta}). \tag{4}$$

Under the assumption of uncorrelated homoscedastic error in (1), the atomic matrix $\mathbf{A}$ of a given experimental support $\boldsymbol{x}_i$ is computed as

$$\mathbf{A}(\boldsymbol{x}_i, \boldsymbol{\theta}) \doteq \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}}(\boldsymbol{x}_i, \boldsymbol{\theta})^\intercal \Sigma_y^{-1} \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}}(\boldsymbol{x}_i, \boldsymbol{\theta}). \tag{5}$$

In the classical D-optimal sense, the information criterion in (3) for a maximal local design (LD) is expressed as

$$\phi^{\text{LD}}(\xi) \doteq \log\det\left(\mathbf{M}(\xi, \boldsymbol{\theta}_0)\right), \tag{6}$$

where $\boldsymbol{\theta}_0$ are the nominal parameter values. This formulation is risk-inclined in the sense that it ignores the uncertainty in the model parameters.

To mitigate the risk of uninformative experiments, one can adopt the Bayesian paradigm and describe the model parameter uncertainty with a probability distribution $\pi(\boldsymbol{\theta})$. The average design (AD) approach maximizes the expected value of information content over the model uncertainty,

$$\phi^{\text{AD}}(\xi) \doteq \int_{\boldsymbol{\theta}} \pi(\boldsymbol{\theta}) \log\det\left(\mathbf{M}(\xi, \boldsymbol{\theta})\right) d\boldsymbol{\theta}. \tag{7}$$

The above takes a neutral attitude towards risk. By contrast, a risk-averse attitude can look at those model uncertainty scenarios corresponding to a given lower percentile of the information content. For instance, information criteria for a maximal risk-averse design both in terms of the value-at-risk (VaR) and conditional-value-at-risk (CVaR) can be expressed as (Kusumo et al., 2022b)

$$\phi_\beta^{\text{VaR}}(\xi) \doteq \max\left\{ V : \int_{\boldsymbol{\theta}:\Phi(\mathbf{M}(\xi,\boldsymbol{\theta}))\leq V} \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} \leq (1-\beta) \right\} \tag{8}$$

$$\phi_\beta^{\text{CVaR}}(\xi) \doteq \int_{\boldsymbol{\theta}:\Phi(\mathbf{M}(\xi,\boldsymbol{\theta}))\leq \phi_\beta^{\text{VaR}}(\xi)} \pi(\boldsymbol{\theta}) \Phi\left(\mathbf{M}(\xi, \boldsymbol{\theta})\right) d\boldsymbol{\theta}, \tag{9}$$

where a higher confidence value $\beta \in [0, 1)$ corresponds to higher risk avoidance.

3. COMPUTATIONAL METHODOLOGY

3.1 Optimization Formulations

The proposed computational framework builds on the continuous-effort design methodology (Fedorov and Leonov, 2014; Kusumo et al., 2021; Vanaret et al., 2021), where the experimental design space is discretized into a finite collection of experiment candidates, denoted by $\mathbb{X}_s \doteq \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_{N_s}\} \subset \mathcal{X}$ with $N_s \gg N_t$. In the case of a simple experimental design space $\mathcal{X}$, a sample can be obtained via gridding or the application of low-discrepancy sequences such as Sobol' sampling (Sobol', 1967). In the presence of operational constraints, adaptive sampling techniques may instead be used to draw samples from a restricted experimental space (Kusumo et al., 2022a).

Likewise, the distribution π of the uncertain model parameters is discretized as a set of N_π parameter scenarios, collected in the set Θ_π. For simplicity and without loss of generality, such samples are assumed to be equally weighted. In practice, the samples could be obtained as the posterior distribution from a Bayesian inference problem with an appropriate likelihood function and prior.

These discretizations recast the search over ξ in (3) into a more tractable, finite-dimensional search over the experimental efforts p_i associated with each experiment candidate $\boldsymbol{x}_i \in \mathbb{X}_s$,

$$(p_1^*, \ldots, p_{N_s}^*) \in \arg\max_{\mathbf{p}} \phi(\mathbf{p}). \tag{10}$$

A noteworthy difference with the experimental campaign (2) is that the optimal efforts $(p_1^*, \ldots, p_{N_s}^*)$ are allowed to take a value of zero. The support of the optimal design ξ^* is then *recovered* after solving (10) as the collection of experimental samples $x_j \in \mathbb{X}_s$ that have an effort $p_j > 0$,

$$\text{supp}(\xi^*) \doteq \{x_j \in \mathbb{X}_s : p_j > 0\}. \tag{11}$$

A maximal local exact design is computed by solving the following pure integer nonlinear program (INLP),

$$\max_{\mathbf{p}} \ \log\det\left(\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_0)\right) \tag{12}$$

$$\text{s.t.} \ \sum_{i=1}^{N_s} p_i = N_t, \ p_i \in \mathbb{Z}_+, \ \forall i, \tag{13}$$

and for a maximal average exact design, the following INLP,

$$\max_{\mathbf{p}} \ \frac{1}{N_\pi} \sum_{j=1}^{N_\pi} \log\det\left(\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_j)\right) \tag{14}$$

$$\text{s.t.} \ \sum_{i=1}^{N_s} p_i = N_t, \ p_i \in \mathbb{Z}_+, \ \forall i, \tag{15}$$

where the objective function is an estimator of the average design criterion (7) using a sampled average approximation (SAA) based on the discretized model uncertainty set Θ_π.

A maximal CVaR exact design is computed by solving the following mixed-integer nonlinear program (MINLP),

$$\max_{\mathbf{p}, v, \boldsymbol{\delta}} \ v - \frac{1}{(1-\beta)N_\pi} \sum_{j=1}^{N_\pi} \delta_j \tag{16}$$

$$\text{s.t.} \quad \delta_j \geq v - \log\det\left(\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_j)\right), \ \delta_j \geq 0, \ \forall j \tag{17}$$

$$\sum_{i=1}^{N_s} p_i = N_t, \ p_i \in \mathbb{Z}_+, \ \forall i \tag{18}$$

where the auxiliary (continuous) variables v and δ_j model the VaR and the potential shortfall of information below the VaR value in each uncertainty scenario $j = 1, \ldots, N_\pi$. This formulation also uses an SAA and follows the seminal work by Rockafellar and Uryasev (2000); see Kusumo et al. (2022b) for a step-by-step derivation.

Recall that the continuous design counterparts to the optimization problems (12)–(13), (14)–(15) and (16)–(18) are readily obtained upon relaxing the integrality restrictions $p_i \in \mathbb{Z}_+$ as $p_i \geq 0$. Since the continuous efforts will not be exact repetitions of the experimental candidates in general, a rounding procedure needs to be applied. While typically suboptimal, such rounding procedures may still provide a good guess for the exact design. Where the number of continuous supports N_c is no larger than the target number of experiments N_t in the campaign, a popular rounding procedure, called *efficient rounding* (Pukelsheim and Rieder, 1992), constructs rounded efforts $p_i^\dagger$ as follows:

- Calculate $p_i^\dagger \doteq \lceil (1 - \frac{N_c}{2N_t}) p_i \rceil$ for $i = 1, \ldots, N_c$.
- While $N_t \neq \sum_{i=1}^{N_c} p_i^\dagger$, repeat:
  If $N_t > \sum_{i=1}^{N_c} p_i^\dagger$, increment $p_j^\dagger \leftarrow p_j^\dagger + 1$ for $j \in \arg\min\{p_i^\dagger / p_i : 1 \leq i \leq N_c\}$;
  Else if $N_t < \sum_{i=1}^{N_c} p_i^\dagger$, decrement $p_j^\dagger \leftarrow p_j^\dagger - 1$ for $j \in \arg\max\{p_i^\dagger / p_i : 1 \leq i \leq N_c\}$.

Alternatively, if $N_c > N_t$, the *greatest effort rounding* procedure simply assigns an equal effort of 1 to the N_t number of supports with the largest efforts and zero effort to the rest.

3.2 Numerical Solution Algorithm

Except for the integrality restrictions $p_i \in \mathbb{Z}_+$, the optimization problems (12)–(13), (14)–(15) and (16)–(18) are otherwise convex. This follows from the property that the term $\log\det\left(\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_j)\right)$ is concave in the variable p_i, while the other terms are all linear in p_i, v or δ_i.

Of the alternatives to solve MINLPs with convex participating functions, we consider an outer-approximation (OA) decomposition algorithm (Duran and Grossmann, 1986; Fletcher and Leyffer, 1994). Although the MINLPs may comprise hundreds or even thousands of integer variables for practical problems, our hypothesis is that convergence to an exact design optimum may require only a small number of iterations when initialized from the rounded optimum of its continuous design counterpart, thereby avoiding having to consider a large number of suboptimal designs. We put this hypothesis to the test later in the case study.

The OA algorithm is therefore initialized by solving the continuous design counterpart, which provides an upper bound on the exact design optimum; a lower bound is then obtained by rounding the optimal efforts in the continuous design. The nonlinear programming (NLP) solver for the continuous design problem is itself initialized by setting all the efforts p_i at positive values, for instance all equal to N_t/N_s. The rationale behind this initialization is that, if the FIM were to be singular at a point with all $p_i > 0$, then it would also be singular for any convex combination $\sum_{i=1}^{N_s} p_i = N_t$ with $p_i \geq 0$,[1] in which case the discretized (exact or continuous) design problem would fail to have any solutions. Such singularity could indicate that the collection $\mathbb{X}_s$ of experiment candidates needs to be augmented, or that the full experiment design problem (3) itself does not have any solutions for the prescribed measured responses $\mathbf{y}$.

After initialization, the OA algorithm iterates between the numerical solution of a mixed-integer linear program (MILP) master subproblem (which provides an updated

[1] Suppose that $\sum_{i=1}^{N_s} \frac{N_t}{N_s} \mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_j)$ is a singular matrix and let the vector $\mathbf{v} \in \mathbb{R}^{n_\theta}$ be in the null space. Since the atomic matrices $\mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_j)$ are positive semi-definite, it follows that $\mathbf{v}^\intercal \mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_j)\mathbf{v} = 0$ for each $i = 1, \ldots, N_s$. For any non-negative combination of efforts $p_i \geq 0$, we have:

$$\mathbf{v}^\intercal \left(\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_j)\right) \mathbf{v} = \sum_{i=1}^{N_s} p_i \mathbf{v}^\intercal \mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_j)\mathbf{v} = 0.$$

Therefore, $\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_j)$ is also a singular matrix. □

upper bound on the information content and a new exact design candidate) and the solution of a continuous primal subproblem obtained by fixing the integer variables in the original MINLP at the master MILP solution (which may result in obtaining an improved feasible solution). The optimization problems (12)–(13) and (14)–(15) being INLPs, their primal subproblems reduce to a simple evaluation of the (local or average) information criterion. By contrast, the risk-averse problem (16)–(18) is an MINLP and the primal subproblems there search over the subspace of the auxiliary variables v and $\boldsymbol{\delta}$.

New linear cuts are appended to the master MILP subproblem at each iteration through linearizing the convex nonlinear objective and/or constraint functions of the MINLP at the solution of the previous primal subproblem. Recall that the only nonlinearities in the optimization problems (12)–(13), (14)–(15) and (16)–(18) are due to the terms $\log\det\left(\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i,\boldsymbol{\theta}_j)\right)$. Expressions for the derivatives of these terms with respect to the efforts p_i directly follow from Jacobi's formula, [2]

$$\frac{\partial}{\partial p_i}\left[\log\det\left(\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i,\boldsymbol{\theta}_j)\right)\right] = \operatorname{tr}\left[\left(\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i,\boldsymbol{\theta}_j)\right)^{-1} \mathbf{A}(\mathbf{x}_i,\boldsymbol{\theta}_j)\right] \tag{19}$$

Note that such linear outer-approximation cuts may only be appended when the FIMs $\sum_{i=1}^{N_s} \hat{p}_i \mathbf{A}(\mathbf{x}_i,\boldsymbol{\theta}_j)$ are non-singular at the solution point $\hat{p}_i$ of the previous master MILP subproblem, for all model parameter scenarios $\boldsymbol{\theta}_j \in \Theta_\pi$. Otherwise, the integer-cut $\|\mathbf{p}-\hat{\mathbf{p}}\|_1 \geq 1$ is added to the next master MILP subproblem instead. The rest of the cuts corresponding to linear objective or constraint functions are added to the initial master MILP subproblem only. In order to expedite the convergence of the OA algorithm, the initial master MILP subproblem incorporates additional cuts derived from the linearization of any nonlinear objective or constraint functions at the continuous design optimum (prior to rounding).

3.3 Software Implementation

The continuous design relaxation problem and any primal NLP subproblem during the OA iterations are solved using the sparse nonlinear solver SNOPT (v7.7), [3] with both feasibility and optimality tolerances set to 10^{-7} and analytic gradients based on Eqn. (19). The master MILP subproblems are solved using the solver GUROBI (v10.0), [4] with relative and absolute convergence tolerances of 10^{-6} and 10^{-9}, respectively. The OA iterations are terminated when the relative gap between the master solution value and the incumbent is below 10^{-5}.

[2] An efficient computational implementation entails a single Cholesky decomposition of the FIM $\sum_{i=1}^{N_s} p_i \mathbf{A}(\mathbf{x}_i,\boldsymbol{\theta}_j) = \mathbf{L}_j\mathbf{L}_j^\intercal$ per model parameter scenario θ_j, followed by the solution of two triangular systems for each direction p_i, $\mathbf{L}_j\mathbf{B}_{i,j} = \mathbf{A}(\mathbf{x}_i,\boldsymbol{\theta}_j)$ and $\mathbf{L}_j^\intercal\mathbf{C}_{i,j} = \mathbf{B}_{i,j}$, and finally computing $\operatorname{tr}\mathbf{C}_{i,j}$.

[3] `https://ccom.ucsd.edu/~optimizers/solvers/snopt/`

[4] `https://www.gurobi.com/solutions/gurobi-optimizer/`

The OA algorithm is coordinated from C++, using the library MC++ (v2.3) [5] to evaluate and differentiate expression trees. Linear algebra calculations, including matrix rank and determinant, Cholesky decomposition and triangular system solve, are carried out using the library Armadillo (v12.6), [6] interfaced with MC++ through external operations.

4. CASE STUDY

4.1 Experiment Design Problem Definition

The case study is based on Kusumo et al. (2022a) and considers the exothermic esterification reaction between butan-2-ol and propionic anhydride,

$$\underset{\text{A: propionic anhydride}}{(CH_3CH_2CO)_2O} + \underset{\text{B: butan-2-ol}}{CH_3CHOHC_2H_5} \xrightarrow{H_2SO_4} \underset{\text{C: butyl propionate}}{C_2H_5CO_2(CH_3)CHC_2H_5} + \underset{\text{D: propionic acid}}{C_2H_5CO_2H}. \tag{20}$$

Experiments on a bench-scale calorimeter start with an initial charge of the anhydride (A), containing N_A^0 moles with volume V_A^0 at the given reaction conditions. The alcohol (B) is fed into the reactor at a time-varying volumetric feed rate u ($\mathrm{L\,s^{-1}}$) and concentration $C_\mathsf{B}^{\mathrm{in}}$ ($\mathrm{mol\,L^{-1}}$). Both the initial charge of A and the feed of B are pre-heated to the set reaction temperature before starting the experiments. The esterification reaction is initiated by introducing the alcohol into the reactor and runs for a given batch time τ.

Neglecting thermodynamic non-idealities in the reaction mixture, a simple mechanistic model of the calorimeter can be formulated as

$$\frac{d\xi_\mathsf{A}(t)}{dt} = \frac{r(t)V(t)}{N_\mathsf{A}^0} \quad \text{with } \xi_\mathsf{A}(0)=0 \tag{21}$$

$$\frac{dV(t)}{dt} = u(t) \qquad \text{with } V(0)=V_\mathsf{A}^0 \tag{22}$$

$$r(t) = k^\circ \exp\left(\frac{-E_\mathrm{a}}{RT}\right)\left[\frac{N_\mathsf{A}(t)}{V(t)}\right]^\alpha\left[\frac{N_\mathsf{B}(t)}{V(t)}\right]^\beta \tag{23}$$

$$N_\mathsf{A}(t) = N_\mathsf{A}^0\left[1-\xi_\mathsf{A}(t)\right] \tag{24}$$

$$N_\mathsf{B}(t) = C_\mathsf{B}^{\mathrm{in}}\left[V(t)-V_\mathsf{A}^0\right] - N_\mathsf{A}^0\xi_\mathsf{A}(t) \tag{25}$$

$$\dot{Q}_\mathrm{r}(t) = (-\Delta H_\mathrm{r})r(t)V(t), \tag{26}$$

where t (s) denotes time, ξ_A ($-$) the molar conversion of A, V (L) the volume of reaction mixture, N_A and N_B (mol) the respective amounts of A and B, r ($\mathrm{mol\,L^{-1}\,s^{-1}}$) the reaction rate, $\dot{Q}_\mathrm{r}$ ($\mathrm{J\,s^{-1}}$) the rate of heat generation by the chemical reaction, T (K) the reaction temperature, $\Delta H_\mathrm{r} = -62.5\ \mathrm{J\,mol^{-1}}$ the molar enthalpy of reaction, k° (L, mol, s) the pre-exponential factor, E_a ($\mathrm{J\,mol^{-1}}$) the activation energy, α and β the reaction orders for A and B, respectively, and $R = 8.314\,\mathrm{J\,K^{-1}\,mol^{-1}}$ the gas constant.

To prevent a thermal runaway if the temperature control system should fail, a safety constraint is imposed so a maximum reaction temperature of $T^{\max} = 405\,\mathrm{K}$ may never be exceeded even under adiabatic conditions,

$$T + \min\{N_\mathsf{A}(\tau), N_\mathsf{B}(\tau)\}\frac{-\Delta H_\mathrm{r}}{\rho c_p V(\tau)} \leq T^{\max} \tag{27}$$

[5] `https://github.com/omega-icl/mcpp`

[6] `https://arma.sourceforge.net/docs.html`

with $\rho = 900\,\mathrm{g\,L^{-1}}$ the density, and $c_p = 2\,\mathrm{J\,g^{-1}\,K^{-1}}$ the specific heat capacity of the reaction mixture.

The objective is to design campaigns of maximally-informative experiments using the bench-scale calorimeter to precisely estimate the parameters E_a, k°, α, and β while guaranteeing a safe operation. The kinetic parameters E_a and k° are assumed to follow independent Gaussian distributions, given by $E_\mathrm{a} \sim \mathcal{N}(8.25, 0.41^2) \times 10^4\,\mathrm{J\,mol^{-1}}$, and $k^\circ \sim \mathcal{N}(9.72, 2.92^2) \times 10^7\,(\mathrm{L, mol, s})$ whereas the reaction orders α and β are assumed to follow Bernoulli distributions with a 75% probability of taking a value of 1 and a 25% probability of taking a value of 2.

Measurements of the heat release rate $\dot{Q}_\mathrm{r}$ and the molar conversion ξ_A are available at regular 3000 s intervals. These are assumed to be independently and identically distributed, and to follow Gaussian distributions with a standard deviation of 0.01 for ξ_A and $1\,\mathrm{J\,s^{-1}}$ for $\dot{Q}_\mathrm{r}$.

The experimental controls consist of the reaction temperature $T \in [338, 348]$ (K) and butan-2-ol feed rate $u(t)$, parameterized as a piecewise-constant profile on 4 time intervals of equal duration with $u_1, u_2, u_3, u_4 \in [0.00, 2.80] \times 10^{-5}$ $(\mathrm{L\,s^{-1}})$. For the purposes of this paper, other potential experimental controls are considered to be fixed. These include the initial amount of A, $N_\mathsf{A}^0 = 2.50\,\mathrm{mol}$; initial volume of A, $V_\mathsf{A}^0 = 0.32\,\mathrm{L}$; feed concentration of B, $c_\mathsf{B}^\mathrm{in} = 10.87\,\mathrm{mol\,L^{-1}}$; and overall batch time, $\tau = 30\,000\,\mathrm{s}$.

The probability distributions of E_a, k°, α, and β are jointly discretized using $N_\pi = 100$ Monte Carlo scenarios. Then, a set of 1080 candidate experiments are drawn from the restricted experimental space using a nested sampling algorithm (Kusumo et al., 2022a), all of which satisfy the safety constraint (27) with 95% probability. The corresponding 1080×100 atomic FIMs (5) are computed analytically, using the DAE solver IDAS part of SUNDIALS (v6.1) [7] interfaced with CasADi (v3.5.5) [8] to construct the forward sensitivity DAEs using automatic differentiation.

4.2 Experiment Design Results

Table 1 reports results for campaigns of $N_t = 5$ and 8 experiments under different optimality criteria. In all cases, the D-optimal campaigns are combinations of the same 6 candidate experiments. Except for the CVaR designs, the rounded continuous efforts for the local and average designs turn out to be suboptimal for the exact design problem. Furthermore, the optimal supports in the exact design are typically a subset of those in the rounded continuous design. A notable exception is for the local D-optimal campaign with 8 experiments where the exact design comprises Experiment 89, while neither the continuous design nor the rounded continuous design contain that experiment. It is also worth noting that the CVaR D-optimal designs has fewer supports than either their local or average counterparts. This behavior was also observed by Kusumo et al. (2022b) and is counter-intuitive insofar as one could expect to find a greater variety of supports in a design aiming to minimize risk.

Computational performance: Local Design Both local designs (left column of Fig. 1) are computed within 1 to 2 seconds on a laptop computer, [9] and the outer-approximation algorithm takes 13 iterations to converge to the exact solution (zero gap) in both cases. At iteration 0, the continuous-effort problem takes 0.46 seconds to solve for $N_t = 5$, and 0.34 seconds for $N_t = 8$. Subsequently, the cumulative time for solving the master MILP subproblems is 0.65 seconds for $N_t = 5$ and 0.29 seconds for $N_t = 8$. Both the master and primal subproblem solutions thus contribute a significant share of the total runtime.

Computational performance: Average Design The runtime needed to solve the average designs (middle column of Fig. 1) increases significantly compared to the local designs, 17 seconds for $N_t = 5$ and 25 seconds for $N_t = 8$. While the number of outer-approximation iterations does not increase compared to the local designs, it is the time needed to compute the average criterion values and their derivatives that increases by 100 times due to the 100 uncertainty scenarios. Most of the solution time (13 seconds for $N_t = 5$ and 20 seconds for $N_t = 8$) is consumed by the solution of the initial continuous-effort problem. The cumulative time for solving the master MILP subproblems is comparatively short, taking only 0.20 seconds for $N_t = 5$ and 0.34 seconds for $N_t = 8$, as a single cut is appended at every iteration.

Computational performance: CVaR Design A further increase in runtime is observed to solve the CVaR designs (right column of Fig. 1) compared to the average designs. However, solving the initial continuous-effort problem requires less time, about 11 seconds in both cases, and the number of outer-approximation iterations also decreases. It is the time needed to perform each outer-approximation iteration that increases significantly for the CVaR design, as the solution of a proper NLP subproblem is now required—instead of just evaluating the information criterion in the local and average design cases. The cumulative time for solving the master MILP subproblems also increases to 3.2 seconds for $N_t = 5$ and 9.9 seconds for $N_t = 8$, as 100 cuts are now appended at every iteration.

5. CONCLUSIONS

Through this paper, we investigated the performance of a tailored outer-approximation algorithm to compute exact designs in optimal experiment campaigns. Despite these problems being NP-hard, our results demonstrate that exact designs remain highly tractable for problems with as many as 1000 experiment candidates and 100 uncertainty scenarios. Future work will entail testing the methodology on additional case studies.

REFERENCES

Atkinson, A.C., Donev, A.N., and Tobias, R. (2007). *Optimum experimental designs, with SAS*. Oxford University Press.

Duran, M.A. and Grossmann, I.E. (1986). An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Math Program*, 36(3), 307–339.

Fedorov, V.V. and Leonov, S.L. (2014). *Optimal Design for Nonlinear Response Models*. CRC Press.

[7] `https://computing.llnl.gov/projects/sundials`
[8] `https://web.casadi.org/`
[9] Lenovo ThinkPad X1 Carbon Gen 10 with 12^{th} Gen Intel® Core™ i7-1260P × 16, 32.0 GiB memory, Ubuntu 22.04 operating system

Table 1. Comparison of local, average and CVaR exact designs for D-optimal campaigns with $N_t = 5$ and 8 experiments. Each of the 1080 experiment candidates are assigned an experimental ID number. The left-most column reports experimental supports only, with the corresponding optimal efforts in columns C , R and E referring to the continuous design, rounded continuous design and exact design, respectively.

	$N_t = 5$ experiments									$N_t = 8$ experiments								
	local			average			CVaR			local			average			CVaR		
ID#	C	R	E	C	R	E	C	R	E	C	R	E	C	R	E	C	R	E
54	2.41	2	2	2.24	2	2	2.01	2	2	3.86	3	4	3.59	3	4	3.21	3	3
65	0.26	1		0.25	1	1	1.31	1	1	0.41	1	1	0.40	1		2.09	2	2
89				0.36								1	0.57	1	1			
259	1.52	1	2	1.56	1	2	1.69	2	2	2.43	2	2	2.50	2	2	2.70	3	3
460	0.68	1	1							1.09	1							
523	0.13			0.59	1					0.22	1		0.94	1	1			

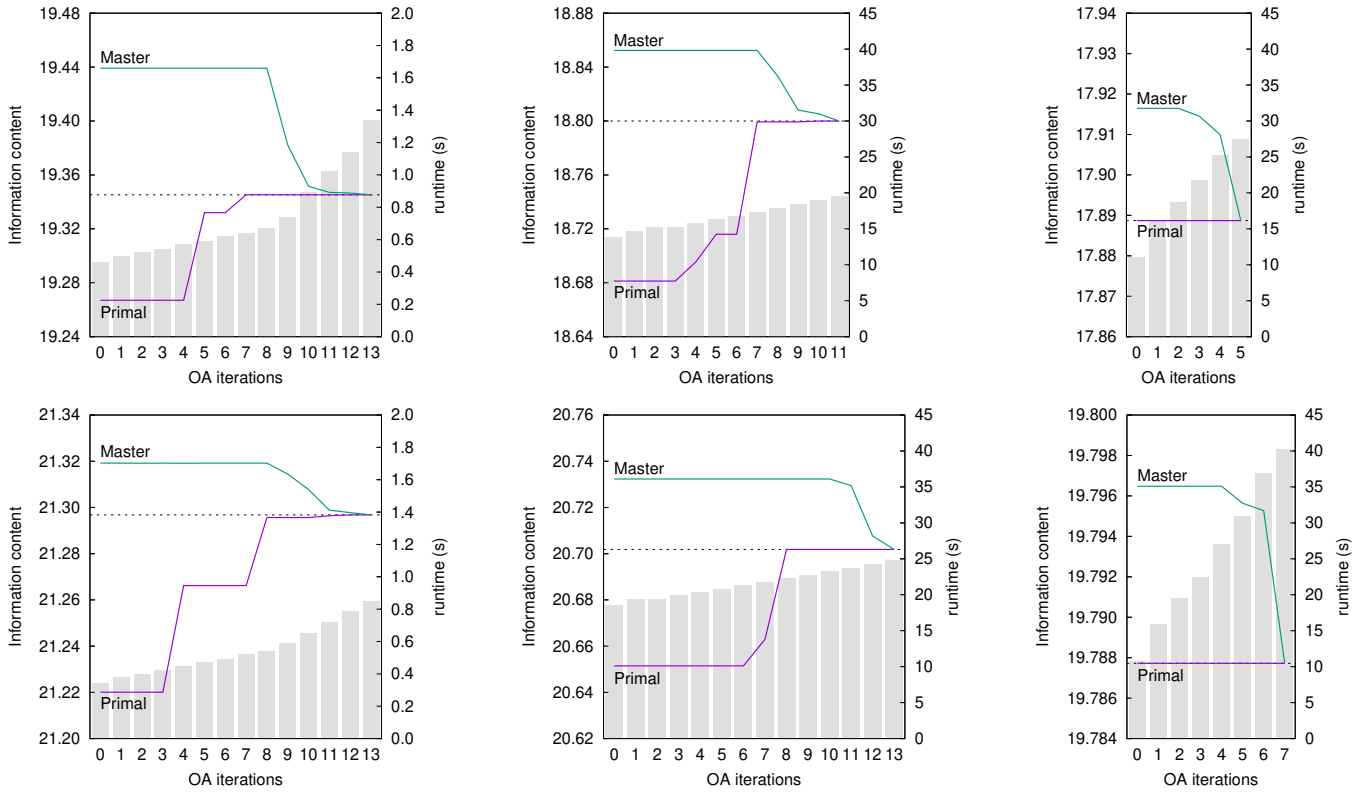

Fig. 1. Convergence profile (color lines) and corresponding runtime (shaded boxes) against iteration count of the outer-approximation algorithm. **Top:** D-optimal campaign with $N_t = 5$ experiments. **Bottom:** D-optimal campaign with $N_t = 8$ experiments. **Left**: local exact design. **Middle**: average exact design. **Right**: CVaR exact design.

Fletcher, R. and Leyffer, S. (1994). Solving mixed integer nonlinear programs by outer approximation. *Math Program*, 66(1), 327–349.

Franceschini, G. and Macchietto, S. (2008). Model-based design of experiments for parameter precision: State of the art. *Chem Eng Sci*, 63(19), 4846–4872.

Kusumo, K.P., Kuriyan, K., García-Muñoz, S., Shah, N., and Chachuat, B. (2021). Continuous-effort approach to model-based experimental designs. *Comput Aided Chem Eng*, 50, 867–873.

Kusumo, K.P., Kuriyan, K., Vaidyaraman, S., García Muñoz, S., Shah, N., and Chachuat, B. (2022a). Probabilistic framework for optimal experimental campaigns in the presence of operational constraints. *React Chem Eng*, 7, 2359–2374.

Kusumo, K.P., Kuriyan, K., Vaidyaraman, S., García-Muñoz, S., Shah, N., and Chachuat, B. (2022b). Risk mitigation in model-based experiment design: A continuous-effort approach to optimal campaigns. *Comput Chem Eng*, 159, 107680.

Pukelsheim, F. and Rieder, S. (1992). Efficient rounding of approximate designs. *Biometrika*, 79(4), 763–770.

Rockafellar, R.T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *J Risk*, 2(3), 21–41.

Sagnol, G. and Harman, R. (2015). Computing exact D-optimal designs by mixed integer second-order cone programming. *Ann Stat*, 43(5), 307–339.

Sobol', I.M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Comput Math & Math Phys*, 7(4), 86–112.

Vanaret, C., Seufert, P., Schwientek, J., Karpov, G., Ryzhakov, G., Oseledets, I., Asprion, N., and Bortz, M. (2021). Two-phase approaches to optimal model-based design of experiments: how many experiments and which ones? *Comput Chem Eng*, 146, 107218.

Welch, W.J. (1982). Branch-and-bound search for experimental designs based on D optimality and other criteria. *Technometrics*, 24, 41–48.