Multi-Parametric Programming for Design Space Identification

Simon Walsha†, Steven Sachioa†, Cleo Kontoravdia, Diogo Narcisob , Maria M. Papathanasioua\*

aThe Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

bDepartamento de Engenharia Química, Instituto Superior Técnico, Universidade de Lisboa, Lisboa 1049-001, Portugal

\*Corresponding author: maria.papathanasiou11@imperial.ac.uk
†These authors contributed equally to this work.

Abstract

Process design is often a challenging task that involves conflicting objectives. To ensure that processes operate within specifications and to accelerate their development, a design space (DSp) can be identified. The DSp is composed by combinations of operating and design decisions that yield a feasible operation. In this work, we present a model-based framework that employs multi-parametric programming for the identification of the DSp. The methodology is applied to the case study of a batch reactor design. The results are compared to the identification of the DSp via quasi-random sampling and the role of the fidelity of the process model in the identification of an accurate DSp is discussed.

**Keywords**: Design Space Identification, Multi-parametric Programming, Industry 4.0.

* 1. Introduction

Process and unit operation design is often challenged by the combination of conflicting objectives and underlying trade-offs with respect to target Key Performance Indicators (KPIs). Feasibility analysis (Grossmann et al., 2014) can be used as a tool to identify a region within which all performance and feasibility constraints are satisfied, also known as the design space (DSp). DSp identification has been an open challenge for the Process Systems Engineering (PSE) community, approached using various methods that include surrogate modelling (Geremia et al., 2023), (adaptive) sampling (Kusumo et al., 2019, Sachio et al., 2023) and probabilistic approaches (Kucherenko et al., 2020). All such approaches exhibit great potential in providing a systematic approach to explore and identify feasible design spaces, guiding process development. Most of such approaches are dependent to the generation of an informed dataset that will provide the basis for the classification of candidate operating points and, therefore, a feasible design space. This can often result into increased computational complexity, particularly in cases where nonlinear models are used.

Bansal et al. (2004) presented a methodology whereby process uncertainty can be integrated in the design and synthesis problem through flexibility analysis via parametric approaches. In this work, we harness the advantages of multi-parametric programming (mp-P) (Pistikopoulos, 2009) to develop a framework for design space identification. The capabilities of the methodology are assessed through a reactor design case study and results are compared to a previously published approach based on quasi-random sampling (Sachio et al., 2023).

* 1. The framework

The multi-parametric Design Space (mp-DSp) framework provides a rigorous methodology for model-based identification of multi-dimensional design spaces. A schematic representation of the multi-parametric Design Space (mp-DSp) framework is shown in Figure 1.



Figure 1 Multi-parametric Design Space (mp-DSp) framework

* + 1. Step 1: Process modelling

The framework relies on the use of validated, high-fidelity process models that describe the unit operation(s) of interest. Such models most commonly comprise nonlinear ordinary or partial differential and algebraic equations ((P)DAEs) and can be simulated in any high-level environment, such as Python, gPROMS® ModelBuilder and MATLAB®.

* + 1. Step 2: Model approximation and virtual experimentation

In this step, the high-fidelity model developed in Step 1 is linearised to harness the potential of linear mp-P and leverage its computational efficiency (Pistikopoulos 2009). For this, the expensive high-fidelity model is sampled within its validation region and a range of interest. In this case, quasi-random Sobol sampling is employed. The resulting dataset is used for the development of a discrete, linear state-space model. Equations 1 & 2 summarise the general formulation, where $x$, $u$, and $y$ are the states, inputs, and outputs respectively, $t$ equates to time, and $T\_{s}$ is the sample time. $A$, $B$, and $C$ denote the matrices of the state-space model.



* + 1. Step 3: Formulation and solution of the mp-P design space problem

Step 3 is focused on the formulation and the solution of the mp-P problem that will determine the feasible design space. Here, we reformulate the traditional mp-P problem (Equations 3-9) and define $θ$ as the set of design variables and/or uncertain system parameters. This will allow the mp-P solution (critical regions) to be expressed as a function of the manipulated (design) variables. This enables direct identification of the possible design space(s) as every critical region of the mp-P solution will be meeting the pre-specified constraints.



The linear state-space model from Step 2 is introduced in the mp-P problem as a set of equality constraints (Equations 4, 5). For this, the parameters of the state-space model are discretised over a chosen number of time steps. The density of the discretisation interval is user-specific and can be assessed based on the performance of the state-space model and system dynamics. However, it is noted that the computational complexity will increase when using a higher number of time steps. Slack variables are introduced in the system to calculate the constraint violations and are appended to the equality constraint matrices. For the state constraint matrices, the rows correspond to the variables, while the columns correspond to equations. To define the parameter space constraints, both the constraint matrix and the constant terms need to be specified. The optimisation variable is selected by assigning the linear cost term to the associated row in the cost vector. A linear objective function in terms of $x$ and $θ$ is defined. The purpose of this formulation is to get the DSp boundary projected onto the critical region. Therefore, the degrees of freedom in the mp-P formulation are equal to the size of theta (design variables). The resulting multi-parametric problem is solved via standard multi-parametric techniques, using the POP Toolbox in MATLAB® (Oberdieck et al., 2016).

* 1. The system case

The mp-DSp framework’s capabilities are tested on the exemplar case study (Kucherenko et al., 2020) of a single batch reactor.

* + 1. Step 1: Process modelling

The reactor is assumed to be of fixed volume and initially contains 1 m3 of component $a$, at a concentration of 2000 mol/m3. The objective is to produce product *b* from the following sequential reaction: $2a→b→c$. Equations 10-13 describe the dynamic model used in this case, whereby $C\_{a}$, $C\_{b}$, $C\_{c}$ are the concentrations of components $a$, $b$, and $c$ respectively, $τ$ and $t\_{f}$ are the process time and the final process time, $k\_{j}$ is the rate of reaction, $E\_{j}$ and $Ai\_{j}$ are the activation energy and the Arrhenius constant $∀$ reaction $j=1, 2$, $R$ and $T$ are the real gas constant and the process temperature, respectively.



The Ordinary Differential Equation (ODE) high-fidelity model (Equations 10-13) was simulated in MATLAB®, assuming: $E\_{1}=20786.7$ J/mol, $E\_{2}=41570.8$ J/mol, $Ai\_{1}=0.0641$, $Ai\_{2}=9938.1$. In this example, we identify two critical variables that impact the process performance; namely the temperature ($T$), and batch time ($t\_{f}$) ranging 250-300 K and 250-350 min, respectively. We aim to identify a design space that will guarantee end-point component concentrations as follows: $C\_{a}\left(t\_{f}\right)\leq 185 $mol/m3, $C\_{b}\left(t\_{f}\right)\geq 790 $mol/m3, and $C\_{c}\left(t\_{f}\right)\leq 140$ mol/m3.

* + 1. Step 2: Model approximation and virtual experimentation

The first step for the development of a linear, state space model is the generation of representative dataset. For this, the ODE model described in Step 1 was used as virtual experimentation platform for the generation of 1023 points via quasi-random Sobol sequence. Temperature and batch time were used as inputs and varied within a range of interest ($250\leq T\leq 300$ and $250\leq t\_{f}\leq 350$) and the performance of the outputs ($C\_{a}, C\_{b}, C\_{c}$) (KPIs) was monitored. The state space model followed the general format (Equations 1 and 2) and was created using the N4SID algorithm in MATLAB®. The final formulation comprises 6 states, with discretisation interval ($T\_{s}$) of 3.5 min and the matrices $(A, B, C)$ are displayed below.



The model demonstrates good performance with mean absolute percentage errors 4-13 % for the three outputs.

* + 1. Formulation and solution of the mp-P design space problem

The linear state-state space model from Step 2 is used for the formulation and solution of multi-parametric programming (mp-P) problem. In this case, the objective function is defined as the maximisation of the amount of target product $b$ at the end of the batch. The mp-P formulation consists of 915 continuous variables, 2 parameters, 6 inequality constraints, and 915 equality constraints and is solved via the POP® Toolbox (Oberdieck et al., 2016). The design space (DSp) defined by the mp-P problem is depicted by the dashed-line in Figure 2a. The space is two-dimensional and a function of the two variables considered as inputs: namely the temperature ($T$), and the batch time ($t\_{f}$). It is observed that the mp-DSp framework can identify a DSp boundary for the system. The multi-parametric solution consisted of a single critical region defined in Equation 14.



For comparison, the mp-P solution is benchmarked against the DSp produced based on the high-fidelity process model (Figure 2a, continuous black line) and following the framework published by Sachio et al. (2023). The two design spaces consider the exact same feasible bounds for the process time (250-350 min), while the mp-DSp presents a lower acceptable bound for the temperature (262 K). As observed, the DSp identified by using the state-space formulation (mp-DSp) slightly overpredicts the acceptable design space by considering 7 % of false positive points.

**

Figure 2 (a) *Comparison of the mp-DSp solution with the design space of the high-fidelity model and (b) Parity plot comparing the high-fidelity model outputs against the state-space model outputs.*

Although a 7 % error is not numerically high, there exist processes with low error tolerance. To further investigate the source of the observed discrepancy, the state-space/high-fidelity parity plot was constructed, as a function of the model outputs ($C\_{a}, C\_{b}, C\_{c}$). For an accurate state-space model, the majority of the points would lie close or on the $x=y$ line. In this case, (Figure 2b) the state-space model presents inaccuracies particularly in the predictions of the concentrations for $b$ and $c$. Specifically, the state-space model overestimates $C\_{b},$ while underestimating $C\_{c}$ - particularly towards the end of the process. This can be translated into the concentration of *b* being predicted higher than its true value. The mismatch between the state space and the high-fidelity process model plays a key role in the formulation and solution of the mp-P and can be the main source for the mismatch observed between the DSps calculated with the two approaches (Figure 2a). Nonetheless, the mp-DSp presents great potential towards the identification of general representation of high-dimensional design spaces.

* + 1. Design space uniqueness

In any model-based analysis it is critical to understand and confirm the trustworthiness of the solution, particularly in cases where parametric uncertainty is high. As discussed in Sachio et al. (2023), design spaces that are defined using computational geometry rely on the density of the point cloud generated up to a certain extent. Figure 3 compares six design spaces generated using quasi-random Sobol sampling on the validated high-fidelity model directly (Sachio et al., 2023). The DSps are defined based on different point cloud densities that range from 512 to 4096 points. It can be observed that as the data resolution decreases, the DSps decrease in size. This is expected, as fewer points are generated by the sampling method and, therefore, there is less information available to be utilised for the identification of the acceptable space boundaries. When comparing this methodology to the mp-DSp framework presented here, it is noted that, in the case of the latter, the DSp identified is unique. Critically, the mp-DSp is independent of any data generation as it is defined via the solution of an explicit optimisation problem that focuses solely on the identification of the bounds.



Figure 3 Assessment of design spaces identified using 128, 256, 512, 1024, 2048 and 4096 quasi-random samples of the high-fidelity model and the design space identified via mp-DSp.

* 1. Conclusions

In this work, we present a framework based on mp-P for the identification of process DSp. The methodology harnesses the computational efficiency of mp-P to identify the DSp bounds, without the need for expensive data generation. The capabilities of the presented approach are demonstrated through an established example of a reactor, whereby the DSp for product maximisation needs to be identified. The results indicate that the mp-DSp framework identifies the DSp satisfactorily, presenting a relatively low mismatch (7 %) to the DSp as identified using quasi-random sampling on the high-fidelity process model. The presented mismatch is attributed to the linear state space formulation required for the construction and solution of the mp-P problem. Current work is focusing on the assessment of different approximation methods to improve the solution of the mp-DSp and solving more complex design problems with larger number of design variables.

References

V. Bansal, J. D. Perkins, and E. N. Pistikopoulos, 2004, Flexibility analysis and design using a parametric programming framework, AIChE Journal, 48, 2851-2868

M. Geremia, F. Bezzo and M. G. Ierapetritou, 2023, A novel framework for the identification of complex feasible space, Comp. & Chem. Eng., 179

I.E. Grossmann, B.A. Calfa, and P. Garcia-Herreros, 2014, Evolution of concepts and models for quantifying resiliency and flexibility of chemical processes, Comp. & Chem. Eng., 70, 22-34

S. Kucherenko, D. Giamalakis, N. Shah and S. García-Muñoz, 2020, Computationally efficient identification of probabilistic design spaces through application of metamodeling and adaptive sampling, Comp. & Chem. Eng., 132

K. P. Kusumo, L. Gomoescu, R. Paulen, S. García Muñoz, C. C. Pantelides, N. Shah, et al., 2019, Bayesian Approach to Probabilistic Design Space Characterisation: A Nested Sampling Strategy, Ind. & Eng. Chem. Res., 59, 2396-2408

R. Oberdieck, N. A. Diangelakis, M. M. Papathanasiou, I. Nascu and E. N. Pistikopoulos, 2016, POP - Parametric Optimisation Toolbox, Ind. & Eng. Chem. Res., 55, 8979-8991

E. N. Pistikopoulos, 2009, Perspectives in multiparametric programming and explicit model predictive control, AIChE Journal, 55, 1918-1925

S. Sachio, C. Kontoravdi and M. M. Papathanasiou, 2023, A model-based approach towards accelerated process development: A case study on chromatography, Chem. Eng. Res. & Des., 197, 800-820