

# Simultaneous selection of designs and models for optimal forecasting in possibly misspecified regression models

許湘伶<sup>1\*</sup>、羅夢娜<sup>2</sup>、銀慶剛<sup>1</sup>

<sup>1</sup> 中央研究院統計科學研究所

<sup>2</sup> 國立中山大學應用數學系

## Abstract

The classical optimal design methods have the capability of determining the designs to achieve estimation or prediction efficiency in situations where the model is correctly specified. However, it is unlikely for these designs to possess optimal properties when model is wrong. While this dilemma can be somewhat relieved by considering a set of candidate models and applying a model selection approach to choose the most appropriate one, it is still difficult for most practitioners to claim that the true model is included among the candidate models. Hence an optimal design method that takes model misspecification into account is called for. In this talk, a three-stage procedure is proposed to choose the combination of designs and models for optimal prediction. Firstly, a model selection criterion is devised to choose the model having the best prediction capability regardless of whether the true model is one of the candidate models. Secondly, a design selection criterion is given to determine the most appropriate design under the selected model. In the end, a data splitting/merging strategy is given to enhance the prediction power of the model-design combination chosen from the first two stages. The advantages of this three-stage method is illustrated via theoretical justification as well as simulations.

Keywords: optimal design, model selection, model misspecification