Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

APPLICATION OF DRUM RATE CONTROL AND MULTIOBJECTIVE OPTIMIZATION IN SCHEDULING SEMICONDUCTOR MANUFACTURING FACILITIES

Igor Kuvychko[1], Bennett Poganski[2], Josh Mangahas[1], and Matthew Purdy[1]

[1]INFICON Inc., Intelligent Manufacturing Systems, Syracuse, NY, USA
[2]Polar Semiconductor, Bloomington, MN, USA

ABSTRACT

Job shop scheduling is a famously difficult NP-hard problem. Scheduling of semiconductor manufacturing factories adds many additional challenges such as batching, queue timers, setup changes, and others. The size of modern semiconductor factories necessitates problem decomposition to control the combinatorial complexity. It is typical to split a factory into multiple scheduling areas with similar tools (e.g., diffusion furnaces, or implanters) to make it computationally tractable. Any decomposition introduces a problem of optimal control of local schedulers to optimize global factory KPIs. A practical production schedule must balance multiple and often conflicting performance criteria. This makes factory scheduling a multiobjective optimization problem with their additional complexity. This paper presents an approach of using local wafer out goals (drum goals) for controlling local scheduling to satisfy global fab criteria. This is combined with multi-objective optimization to find and evaluate Pareto-optimal solutions and strike a balance between conflicting objectives.

1 INTRODUCTION

Semiconductor manufacturing is exceptionally capital-intensive. The industry employs some of the most, if not the most, sophisticated manufacturing processes and equipment (production tools). Historically, lot sequencing (the decisions of what lots to run on what tools) were made either by human operators, or by a real-time dispatching (RTD) system which would assign a lot available for processing to a machine when the machine became available. Complex and hard-to-manage rules were devised to control dispatching to complex tools with batching and/or setup changes (Monch et al. 2022). Dispatch rules are "myopic" as they consider a local context for making a job placement decision. Scheduling offers better control and optimization of performance as more "global" criteria can be considered and optimized (criteria which depend on a certain length of scheduled activities). The INIFICON Factory Scheduler is a powerful semiconductor factory scheduling tool which is currently deployed in over 40 factories on several continents, both in front-end and back-end facilities. Scheduler fully supports the complexities of semiconductor manufacturing, such as batching, setup changes, reticles and test board logic, and more. Scheduler behavior is controlled by a configuration which defines the choices and priorities that a certain scheduling area has. A factory is typically split into several scheduling areas which are controlled by separate instances of Factory Scheduler.

2 POLAR SEMICONDUCTOR FACTORY SCHEDULING

Polar Semiconductor is a U.S.-based manufacturer of analog and power semiconductor devices and sensors, with a manufacturing site located in Bloomington, Minnesota. The current capacity of the site is 21,000 wafers per month (as of July 2024), which is expanding to nearly double the capacity to 40,000 wafers per month by 2026-27. This expansion is leveraging existing cleanroom space. Optimized factory performance is critical for achieving this ramp, and automated factory scheduling is one of the key enabling factors behind the growth potential. Polar's manufacturing facility is fully scheduled by INFICON Factory Scheduler. The facility is split into six scheduling groups, and a new schedule is generated at most every 5 minutes to account for changes (such as tools going down or lots going on hold). Scheduling compliance is generally over 99% for all six areas, sustained over a period of many months.

3 DRUM CONTROL

While each scheduling group is locally optimized, this may result in global inefficiency. For example, with sufficient WIP the easiest way to maximize outs is to avoid setup (or reticle) changes and run the same types of lots. While this maximizes the outs and tool utilization, such short-term over-optimization will degrade overall factory performance as some upstream tools will see high incoming WIP load (but some upstream tools will go idle). One can see that local over-optimization is likely to result in the increased variability of arrival times, impacting global factory performance such as cycle time. Local outs goals (drum rates) offer an effective approach to control local scheduling to optimize global factory performance. The drum rates are calculated based on global knowledge of factory starts and consider output objectives over a longer timescale than the shorter-term scheduling window. Drum rates can be calculated at different levels of aggregation (e.g., by technology or across technologies). This paper describes INFICON Factory Scheduler implementation of dynamic drum control at Polar Semiconductor factory and its effectiveness. The standard deviation of moves to drum target was improved by 38% by the implementation. Drum control was also shown to provide significant benefits for factory control during large starts changes.

4 MULTIOBJECTIVE OPTIMIZATION

The INFICON Factory Scheduler generates detailed records (“snapshots”) of factory state and stores them in INFICON Digital Twin. Over time, this builds a rich dataset which captures a factory in variety of states (different loading, tool states, etc.) INFICON Factory Scheduler includes a simulation functionality where a new schedule can be built and benchmarked using any available historical snapshot of the factory state. The other important input to such simulation is a configuration that includes a list of hyperparameters which define the relative importance of different criteria in building a schedule. We have developed a Python library for exploration and tuning of hyperparameters of the Scheduler configuration which incorporates a modern approach to multi-objective optimization via surrogate models. A series of input snapshots of factory state can be selected based on a chosen criteria (e.g., factory loading). Configuration hyperparameters of interest are explored using either a DOE approach (full factorial and LHC sampling are among the supported plans), or Bayesian Optimization approach (Khatamsaz et al. 2023). Scheduler simulation functionality is then used to generate candidate schedules and evaluate them using the KPIs of interest (e.g., tool processing/idle time, number of setup changes, number of wafer completes, or number of completes to drum goal). To simplify the output, the library can aggregate the results (KPI values) across input snapshots (for the same hyperparameter values). Furthermore, Pareto filtering can be performed on rounded sets of solutions to displaying only Pareto-efficient data points. This approach has been used successfully at Polar Semiconductor to explore the interplay between conflicting objectives and to select optimal hyperparameter values to achieve desired KPI tradeoffs.

REFERENCES

Pinedo, M. L. 2022. *Scheduling: Theory, Algorithms, and Systems*. 6th Ed. Springer International Publishing.

Monch, L., J. W. Fowler, and S. J. Mason. 2013. *Production Planning and Control for Semiconductor Wafer Fabrication Facilities*. Springer International Publishing.

Ovacik I. M. and R. Uzsoy. 1997. *Decomposition Methods for Complex Factory Scheduling Problems*. Springer US.

Deb, K. 2002. *Multi-Objective Optimization using Evolutionary Algorithms*. Wiley.

Khatamsaz, D., B. Vela, P. Singh *et al*. 2023. “Bayesian optimization with active learning of design constraints using an entropy-based approach”. *Npj Comput Mater* 9(49).

Jiang, H., Y. Shen, Y. Li, B. Xu, S. Du, W. Zhang, C. Zhang, and B. Cui. 2024. “OpenBox: A Python Toolkit for Generalized Black-box Optimization”. *Journal of Machine Learning Research* 25(120):1-11.