Optimized allocation of defect inspection capacity with a dynamic sampling strategy

Gloria Luz Rodriguez-Verjan [a,b,*], Stéphane Dauzère-Pérès [a,1], Jacques Pinaton [b,2]

[a] Department of Manufacturing Sciences and Logistics, Center of Microelectronics in Provence, École des Mines de Saint-Etienne, CNRS UMR 6158 LIMOS, Site Georges Charpak, 880 avenue de Mimet, F-13541 Gardanne, France
[b] Department of Process Control System, STMicroelectronics, 190 Avenue Clestin Coq, Zone Industrielle, 13106 Rousset Cedex, France

ARTICLE INFO

ABSTRACT

In semiconductor manufacturing, in-line inspections are necessary to monitor processes, products and tools in order to reduce excursions and achieve high yields of final products. However, capacity is limited and inspections directly impact the cycle times of products. Sampling strategies are used to improve product yields while limiting the number of inspections, and thus the impact on the cycle times of the inspected lots. Dynamic sampling has been recently introduced and new models are required to estimate the associated inspection capacity. In this paper, we focus on micro-defect inspections and the risk on process tools in terms of Wafers at Risk (W@R), which is the number of wafers processed on a tool since its latest defect inspection. A linear programming model that estimates the required defect inspection capacity to satisfy the W@R limits on process tools is proposed. Our model can be used at different decision levels. At the tactical level, it shows if W@R limits can be satisfied when the product mix changes and/or if planned W@R reductions can be met with the available inspection capacity. At the strategic level, the model helps to justify capacity investments if the objectives in terms of W@R reduction cannot be achieved with the available inspection capacity. Numerical experiments on industrial data are performed and discussed.

1. Introduction

Semiconductor manufacturing is highly complex and very expensive. Some of the reasons are (i) molecular scope of operations, (ii) re-entrant flow of products, (iii) high level of integration and (iv) feature size of products that has continuously been reduced over the last few decades. Consequently, wafer manufacturing plants (or *fabs*) have improved the control over process parameters and reduced the sources of defects during fabrication to achieve high levels of product yields. Actually, thanks to yield enhancement and defect reduction, significant progress in semiconductor manufacturing has been possible over the last 30 years [5]. The two key areas of process control are *Metrology* and *Defect Inspection* [7]. Metrology aims at controlling the physical and electrical properties of wafers during fabrication, while defect inspection aims at monitoring the process for defect excursions and at driving the continuous improvement of product yields.

Excursions are temporal yield losses that randomly happen (in time) as a consequence of an out-of-control condition of a process tool. This condition can affect all the wafers processed on the tool until the source of the problem is detected and corrected. Therefore, improved detection mechanisms are necessary to enable an early detection of problems and limit the impact of excursions [1]. However, unnecessary controls can negatively impact product yields due to long waiting times of lots in front of metrology or inspection tools, which result in delays to take corrective actions [24]. In addition, some excursions can only be detected at the probing test. Hence, long manufacturing cycle times can have a detrimental effect on product yields [8,25]. Due to all these reasons, sampling strategies are used to improve product yields while limiting the number of inspections. Several sampling strategies exist and can be classified according to their capability to integrate the factory dynamics and variability. Nduhura Munga et al. [14] propose three categories: (i) static sampling, (ii) adaptive sampling and (iii) dynamic sampling.

Static sampling is based on rules that do not change throughout production and different static sampling strategies exist. Inspecting every constant number of lots arriving at a workstation or selecting a fixed number of lots at the beginning of their manufacturing

* Corresponding author at: Department of Manufacturing Sciences and Logistics, Center of Microelectronics in Provence, École des Mines de Saint-Etienne, CNRS UMR 6158 LIMOS, Site Georges Charpak, 880 avenue de Mimet, F-13541 Gardanne, France. Tel.: +33 4 42 61 66 76.

E-mail addresses: rodriguez@emse.fr,
gl.rodriguezverjan@gmail.com (G.L. Rodriguez-Verjan),
Dauzere-Peres@emse.fr (S. Dauzère-Pérès), jacques.pinaton@st.com (J. Pinaton).
[1] Tel.: +33 4 42 61 66 60.
[2] Tel.: +33 4 42 68 58 07.

routes are the most common strategies used in practice. When lots are selected and inspected at different stages of their manufacturing routes, the added defect density can be identified between sequential steps [5]. Based on static sampling, Nurani et al. [15] propose a cost-based sampling methodology to allocate inspection capacity. The model aims at specifying the process operations to inspect, the number of lots, the number of wafers within a lot and the percentage area of the wafer. Scanlan [21] showed the importance of using baseline monitoring to reduce the inspection cost. The main advantage of static sampling strategies is that they are relatively simple to implement compared to other strategies. However, they are not the most effective at detecting excursions as quickly as they occur [2]. The manufacturing dynamics cannot be correctly handled, which often leads to cases of lack of control and over-control of process tools [12].

In adaptive sampling, the sampling rate is adjusted according to the production state [22,2,11]. When a problem occurs, the sampling rate increases in order to control more lots and correct the process variations. When the process is under control or the risk is not significant, the sampling rate decreases to better use the inspection capacity. However, managing inspection capacity is more complex due to variations of the workload of metrology and inspection tools during production. In dynamic sampling, the selection of lots is done in real time and according to the manufacturing state (e.g. information obtained by inspecting lots, workload of inspection tools, and levels of risk in the fab) [17,23,3]. This strategy is more suitable for modern fabs that want to increase product yields while limiting the impacts on cycle times [14].

Several studies consider that inspection can be performed immediately after the process station [18,9]. However, the characteristics of the type of controls considered in this work are different. In general, defect inspections are located between critical processes of the product and based on their capability to detect defects. Therefore, this type of controls cannot be performed after each process operation. In addition, several process tools can be controlled with one inspection operation, and this is defined when defect inspection control plans are designed. Nduhura Munga et al. [13] propose a Mixed Integer Linear Program to determine the optimal values of key parameters that help to reduce the risk on process tools when lots are dynamically selected. Their model is aggregate and does not consider the details of the manufacturing routes.

The risk considered in this work is evaluated in terms of *Wafers at Risk* (W@R) on process tools. The W@R refers to the number of wafers processed on a tool since the last defect inspection performed. In this paper, we propose a model to determine whether a set of predefined W@R limits on process tools can be satisfied and, if not, the additional inspection capacity that is required to keep the W@R on process tools below the limits. Because the W@R refers to the number of wafers processed on a process tool since the last inspected lot, the lots that are selected for inspection do not require "pre" and "post" inspection results. We focus on defect inspection operations because this type of controls addresses all process tools of the fab, and the associated sampling strategy has changed from static sampling to dynamic sampling in the Rousset site of STMicroelectronics. Our model

takes into account all the process operations in the manufacturing routes, the qualifications of process and inspection tools, and the design of defect inspection control plans.

The structure of the paper is as follows. The problem is described in Section 2. Section 3 introduces the mathematical model. Section 4 is devoted to the discussion of the numerical experiments performed on real data. Finally, conclusions and perspectives are presented in Section 5.

2. Problem description

When fabs use static sampling strategies to select lots for inspection, the defect inspection capacity is allocated in advance via the sampling rates of lots. This is because a fixed number of lots are selected at the beginning of their manufacturing route and systematically inspected in all the inspection operations in their routes. By inspecting the same wafers at different inspection operations, it is possible to identify whether defects were recently added or not since the last inspection operation. However, static sampling strategies are less suitable for modern fabs because the dynamics of the fab are not correctly handled. This often leads to cases of control operations being performed without real added value and unexpected levels of risk. These drawbacks have motivated the implementation of dynamic sampling strategies. The sampling system for defect inspection recently changed from static sampling to dynamic sampling in the Rousset site of STMicroelectronics. By using dynamic sampling strategies, the risk on process tools can be better controlled but the defect inspection capacity to be allocated is not known in advance. The reason is that lots are selected in real time at any of the inspection operations of their manufacturing routes.

In order to efficiently control the W@R on process tools, the following factors need to be considered: manufacturing routes of products, defect inspection control plans, qualifications of process tools and inspection tools, W@R limits, product mix. These factors are introduced in the remaining of this section.

The manufacturing route refers to the sequence of process operations that are necessary to obtain the final product and it depends on the specifications of the technology to produce (i.e. type of devices). The wafers of a given product follow the same manufacturing route and are grouped in lots (i.e. batches of at most 25 wafers). Defect inspection operations are placed throughout the manufacturing routes of products. The position and coverage of each inspection operation is defined in the "Defect Inspection Control Plan". Fig. 1 shows a small portion of the manufacturing route of products of "Technology A". In this example, all lots have to be processed on operation 1200 and then on operations 1210–1290. Only some lots will be inspected in defect inspection operations 1202 and 1330. The lots that are going to be inspected in these inspection operations are not defined in advance, because they are dynamically selected according to the production state. The *Coverage Block* refers to the process operations that can be controlled with a given inspection operation. The process tools that are qualified to perform operation
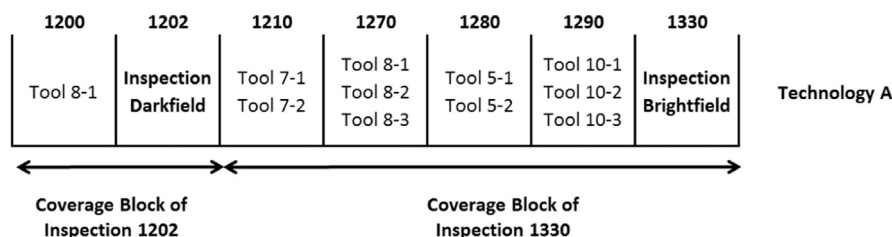

Fig. 1. Small portion of the manufacturing route for technology "A".