

# Global Optimization of a Semiconductor IC Supply Chain Network

Adar A. Kalir<sup>1</sup>, Dean I. Grosbard<sup>2</sup>

<sup>1</sup>Intel Qiriat-Gat and Ben-Gurion University, Beer-Sheva, Israel

<sup>2</sup>University of California at Berkeley, Berkeley, CA, USA

Email: kalira@post.bgu.ac.il, dean.grosbard@berkeley.edu

**How to cite this paper:** Kalir, A.A. and Grosbard, D.I. (2017) Global Optimization of a Semiconductor IC Supply Chain Network. *Journal of Service Science and Management*, 10, 338-352.

<https://doi.org/10.4236/jssm.2017.103027>

**Received:** February 2, 2017

**Accepted:** June 24, 2017

**Published:** June 27, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International

License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In this paper we develop a framework for the semiconductor manufacturing process, including front-end (fab) operations and backend (assembly/test) operations. We then propose a Quadratic Programming (QP) formulation for the profit maximizing objective with flexible demand and price ranges (per product) and fab capacities. We demonstrate the model by applying it to a case study that is based on a real-world (distorted) dataset and show how the solution varies between a local optimization of a single fab (for minimum wafer cost) and a globally optimal solution for maximum profit of a network of fabs.

## Keywords

Supply Chain Management, Optimization, Network Based Manufacturing, Non-Linear Programming, Quadratic Programming

---

## 1. Introduction

Supply chain management (SCM) is defined as a set of approaches used to efficiently integrate suppliers, manufacturers, warehouses, and stores so that products are produced and distributed at the right quantities, to the right locations, and at the right time, in order to minimize overall costs while meeting service level requirements [1]. SCM also pertains to the set of actions and decisions that attempt to synchronize demand and supply with in-process inventories, in order to ensure on-time delivery of product commitments to customers, and optimize the overall manufacturing operations from start to end.

SCM is of paramount importance especially in the semiconductor industry, as highlighted by [2]. The criticality of SCM for this industry stems from the fact that semiconductor wafer fabrication facilities represent large capital investments (usually in the range of Billions of dollars), and the assembly and test fa-

cilities are quite expensive as well (with some of the individual testers costing a few Millions of dollars). The products produced by these facilities are of high value, both in the form of wafers and in the form of integrated circuit (IC) chips. In case of processor chips, each unit can be of the order of a few hundred dollars. With these high capital investments and cost of products, it is critical for semiconductor manufacturers to maintain high utilization of the equipment with minimal inventory. Supply chain management can help in achieving these goals and provide large savings for the semiconductor industry.

Only a handful of papers were published to date that offer a mathematical formulation for the semiconductor supply-chain network planning problem. By network planning, we refer to the strategic long term planning of satisfying forecasted demand and addressing the question of what products to manufacture and where. In this paper, we formulate this problem as a quadratic programming model and provide insight to the benefits of such an optimization to the supply-chain of a semiconductor company. Before going into the details of the model, we first review the existing literature. Then we develop the notation and model formulation of the problem in Section 3, followed by a case study via a numerical example in Section 4. Conclusions and suggestions for further work are depicted in Section 5.

## 2. Literature Review

In this section, we review some of the key research work to date on SCM and the semiconductor industry in particular. Jain [2] proposes a conceptual framework for supply chain modeling and simulation that is based on the supply-chain operational reference (SCOR) model. However, it seems that the framework is too generic and does not include enough specific details for simulating supply chains of the semiconductor industry. Ayers [3], in his handbook of SCM, provides a variety of methodologies applied in order to execute the strategy and deal with root causes such as clarity, variability, design, information flow, and weak links. In semiconductor manufacturing the supply chain is characterized by long throughput times, high levels of variability, and nonlinear dynamics of the manufacturing process. As pointed out in [4], the IC supply chain has become globalized and thus more complex, with additional sources of concern. Specifically, the trustworthiness of IC supply-chain, and the ability to assess its predictability. Chien *et al.* [5] name another reason for the added complexity, caused by the fact that front-end operations are often performed in highly industrialized nations, while back-end operations are typically carried out in countries where labor rates are cheaper.

Many approaches to building strategic plans to effectively operate these complex supply chains have been proposed. The most sophisticated rely on some form of mathematical optimization, e.g. linear programming [6]. One such approach of a mixed integer programming model is described by [7] for optimizing IBM's semiconductor supply chain. In their model, they consider the semiconductor production material flow throughout the supply chain from raw sili-

con wafers to finished modules. They restrict the model to integers in order to account for lot-size constraints since production starts are constrained to be at least a certain minimum quantity and must be released in discrete multiples of the allowable lot size. In our model, we relax this assumption and broaden the scope of the supply chain beyond the completed modules all the way to the shipping warehouses right before the retail customers at the end of the supply chain.

Despite the extensive research on SCM, there seem to be lack of applied research pertaining to the SC network planning problem in the semiconductor industry. During 1993-1995, a team of researchers at SEMATECH developed a strategic decision support system to assist large semiconductor companies in managing their supply networks. The system included an optimization model that helped determine the configuration of the supply network and which products should be built in which fab. This work is summarized in [8]. Heath and Jackson [9] propose the following steps for the production planning problem (short to midterm): 1) Forecast demand for next M-months at the end of each month; 2) Use a deterministic LP model to derive the production plan; 3) Implement the production plan for current month; and 4) Re-forecast. They use a simulation approach for the forecast that is based on the Martingale Model of Forecast Evolution (MMFE). A decade later, Stray *et al.* [10] developed a strategic model that determines whether or not to open new fabs; more bottleneck equipment; open new assembly/test facilities; and which parts will be made in which facilities. Rastogi *et al.* [11] further extend the work and propose a model to minimize expected total cost in the SC network when the demand is uncertain. Ehm *et al.* [12] describe steps to build a supply chain simulation reference model for the semiconductor industry and identify the main building blocks of such models. They present an approach to obtain a supply chain network as a set of reduced simulation models. Aelker *et al.* [13] suggest that managing complexity in supply chains is the key and use the semiconductor SC as an example. Recently, Lowe and Mason [14] also propose a comprehensive MILP formulation in order to minimize total costs by scheduling weekly production quantities to meet forecasted demand over a six-month planning horizon. The model considers facility capacities, facility qualifications, limits on the number of facilities to use, and minimum inventory requirements. An older related paper to this problem by Habla and Mönch [15] also deals with a volume and capacity planning problem and a linear programming (LP) approach is suggested for finding the production quantities for a planning horizon of 26 weeks with weekly time buckets and an objective of maximized revenue at the same minimized production costs.

Most of this literature review has been motivated by works that pertain to the semiconductor industry and its supply chain as this type of supply chain possesses unique characteristics in terms of the behavior of product ASP's (average selling prices) over time or the breakdown of the supply chain by the components created during the process (i.e. wafers, then die). Yet, it is important to mention that there are many related works on similar problems for other indus-

tries. As an example, Kannegiesser *et al.* [16] discuss a case study of products that are considered commodities (unlike semiconductor high performance computing products) from the chemical industry where they develop an integrated model to optimize profit by coordinating quantity and price decisions.

In summary, despite all the work that has been published, the modeling and solution approach to the semiconductor supply-chain network planning problem has not been explicitly addressed to date.

### 3. Model Framework

The contribution of the current paper emanates from the fact that it addresses pertinent aspects that were not addressed integrally in previous papers thus far. First, we consider a profit maximizing objective rather than a cost minimizing objective. Although many have addressed the profit maximizing objective before, some were noted in the literature section of this paper, it was not in conjunction with the next few differentiators.

Second, we assume that there is at least some flexibility in setting the demand for each product in sold units, such that there is a range rather than a point estimate, and the final value produced for each product is constrained within this range. Third, we extend the analysis to varying ASP's (average selling prices) of any product by customer, such that different customers may be offered different ASP's for the same product. This is embedded in order to reflect different price ranges for differences in quantities sold. This covers, for example, the cases where customers are awarded discounts for purchasing higher volumes of a given product. Implied by this extension is that the mathematical formulation of the problem becomes quadratic rather than linear. Lastly, we also assume that there is some flexible capacity across the product mix such that each fab has a capacity range between a minimum and a maximum value rather than a fixed number.

The framework for the planning problem of this supply chain network in our paper is depicted in **Figure 1**. Demand for the portfolio of products from the various customers is aggregated by product and transformed into volumes of wafer starts in the fabs. Then, an allocation is made of how many wafers are to be started at each fab and a routing is set for how they would flow through the supply chain from the fabs to the assembly/test (AT) sites and to the warehouses and/or end customers. The aggregation is necessary in order to be able to make decisions that smooth out the required production quantities through the supply chain and fit these quantities to the desired loadings of the production sites (fabs, AT's). Changes over time are typically addressed not by changing the producing sites or the routes through which the volumes are produced along the SCN but by adjusting the quantities of the wafer starts at each fab and thus are important but addressed by lower level business processes (such as mid-term capacity planning).

Once production is completed at the fabs, the wafers are shipped for assembly, test and unit packaging, and then shipped as units, after order consolidation, to the various OEMs. **Figure 2** provides a flow chart of these seven stages.



Figure 1. The SC network planning problem.

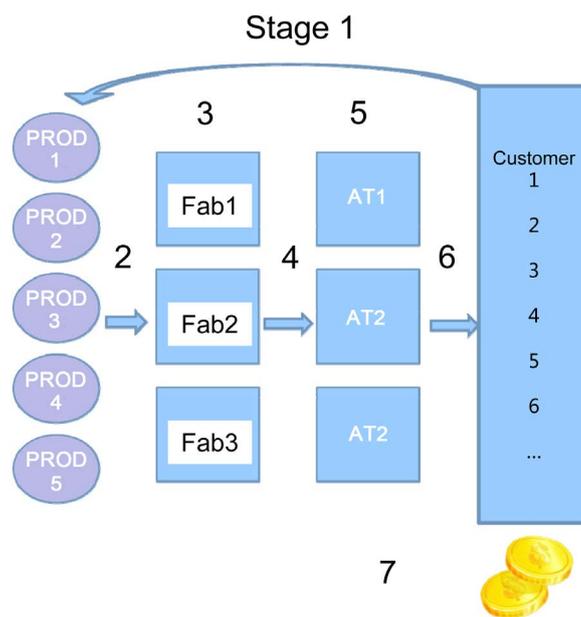


Figure 2. Flow chart of the seven stages.

### 3.1. Motivation

In order to illustrate the problem at hand, consider the following numerical example. We note here that in both the following numerical example and in the case study later on, the datasets are inspired by real fab networks. In our numerical example, there are two fabs, denoted as Fab 1 and Fab 2, are each with a

7000 weekly wafer start capacity. Demand per product and wafer cost per product per fab are given in **Table 1**. Note that the total demand for the three products (PROD 1, 2, and 3) is 14,000. Hence, total capacity is matched with total demand in this case. Given that all demand is satisfied in this simplistic case, the revenue is fixed and the profit maximization problem reduces to a cost minimization problem.

In solving this problem optimally for each of the fabs independently (i.e. to minimize that fab's cost), while satisfying the demand, the resultant solutions are shown in **Table 2**. Note that Fab 1's optimal solution leaves PROD 3 to be solely produced by Fab 2 at the expensive wafer cost of 3,000 per wafer. Similarly, Fab 2's optimal solution leaves the same product to be solely produced by Fab 1. However, when solved optimally for both fabs simultaneously, the solution is different than the previous solutions and is better by 6.25% and 1.9% respectively, as shown in **Table 3**.

**Table 1.** Numerical Example.

	Demand	Fab1	Fab2
		Wafer Cost	
PROD 1	4000	1500	2500
PROD 2	5000	1000	1500
PROD 3	5000	2250	3000

**Table 2.** Optimal solutions by fab.

Fab1 Optimal			
	Fab1	Fab2	
	Volume		
PROD 1	2000	2000	
PROD 2	5000	0	
PROD 3	0	5000	
Cost	8,000,000	20,000,000	
Total Cost	28,000,000		

Fab2 Optimal			
	Fab1	Fab2	
	Volume		
PROD 1	2000	2000	
PROD 2	0	5000	
PROD 3	5000	0	
Cost	14,250,000	12,500,000	
Total Cost	26,750,000		

**Table 3.** Global optimal solution.

	Globally Optimal	
	Fab1	Fab2
	Volume	
PROD 1	4000	0
PROD 2	0	5000
PROD 3	3000	2000
Cost	12,750,000	13,500,000
Total Cost	26,250,000	

### 3.2. Notation

The following notation is utilized. Indices:

- Products:  $i = 1, \dots, N$
- Fabs:  $j = 1, \dots, M$
- ATs:  $v = 1, \dots, V$
- Customers:  $k = 1, \dots, K$

Inputs for stages 1 and 6:

- $Dmin_{i,k}, Dmax_{i,k}$  –Min/Max of demand range in units per customer per product. Min demand must be met. Any demand between the min and the max can be supplied to increase revenue (and profit).
- $Rmin_{i,k}, Rmax_{i,k}$  –ASP per unit of product  $i$  for customer  $k$  when purchasing  $Dmin_{i,k}, Dmax_{i,k}$  units.
- $ASP_{ik}(\bar{y}_{ik}) = \alpha_{i,k} + \beta_{ik} \left( \sum_v y_{i,k,v} - Dmin_{i,k} \right)$  –ASP as function of the number of units sold above the minimum demand.
- $\alpha_{i,k} = Rmin_{i,k}$  –The intercept for the ASP as a linear function of the demand.
- $\beta_{ik} = (Rmax_{i,k} - Rmin_{i,k}) / (Dmax_{i,k} - Dmin_{i,k})$  –The rate of ASP change due to volume discount.

Inputs for stages 2 and 3:

- $CFmax_j, CFmin_j = CFmax_j \cdot MinUtilization$  –Fab capacity range, per fab, of their actual capacity.

Note that *MinUtilization* is given per fab, where it is assumed that fabs have a minimum threshold for utilization due to taxation and other governmental obligations; and this is applicable to most international companies with presence in several countries. In case this is irrelevant it can simply be set to zero.

- $Q_{i,j} \in \{0,1\}$  –Qualification index. It is set to a value of 1 if Fab  $j$  qualifies to produce product  $i$ , and a value of zero otherwise.
- $CTR_i$  –Capacity Trade Ratio per product. Since different products consume different capacities at the fab, they are normalized to a trade ratio of 1.0, and any ratio lower/higher than this value implies that the corresponding product requires less/more capacity.
- $FTA_{j,v}$  –Fab-to-AT shipping cost matrix.

Inputs for stages 4 and 5:

- $CMax_v$  –AT capacity in wafers.
- $DPW_i$  –The given die (unit) per wafer for each product.
- $WCA_v$  –Production cost per wafer
- $DCA_v$  –Production cost per K die (unit).
- $ATC_{v,k}$  –AT-to-Customer shipping cost matrix.

Decision variables:

- $x_{i,j,v}$  –The amount of wafers of product  $i$  shipped from fab  $j$  to AT  $v$ .
- $y_{i,v,k}$  –The amount of units of product  $i$  supplied from AT  $v$  to customer  $k$ .

### 3.3. Problem Formulation

Next is a Quadratic Programming (QP) problem formulation to the problem at hand. QP means that the objective function is quadratic in the decision variables and is subject to linear constraints on these variables.

$$\begin{aligned}
 Z_{\max} = & \sum_{i,k} \left( (\alpha_{i,k} + \beta_{ik} (\sum_v y_{i,v,k} - Dmin_{i,k})) (\sum_v y_{i,v,k}) \right) \\
 & - \sum_{i,j} WCF_{j,i} \cdot (\sum_v x_{i,j,v}) - \sum_{i,v} (WCA_v + DCA_v \cdot DPW_i) \cdot (\sum_j x_{i,j,v}) \\
 & - \sum_{j,v} FTA_{j,v} (\sum_i x_{i,j,v}) - \sum_{v,k} ATC_{v,k} (\sum_i y_{i,v,k})
 \end{aligned} \tag{1}$$

$$Dmin_{i,k} \leq \sum_v y_{i,v,k} \leq Dmax_{i,k}, \forall i,k \tag{2}$$

$$CFmin_j \leq \sum_i (CTR_i \sum_v x_{i,j,v}) \leq CFmax_j, \forall j \tag{3}$$

$$\sum_{i,j} x_{i,j,v} \leq CMax_v, \forall v \tag{4}$$

$$\sum_j x_{i,j,v} \cdot DPW_i \geq \sum_k y_{i,v,k}, \forall v,i \tag{5}$$

$$\sum_v x_{i,j,v} \leq Q_{i,j} \cdot CFmax_j, \forall i, j \tag{6}$$

The objective function to be maximized in Equation (1) is a quadratic profit maximizing objective that incorporates revenue from sold units with a discount factor based on the volume of purchase (first term of the equation), fab production costs (second term), AT production costs (third term) and shipping costs from Fab to AT and from AT to customers (fourth and fifth terms respectively).

The constraints are reflected in Equations (2) to (6). Demand range for each product and each customer is set by Equation (2). Fab capacity range for each fab is accounted for in Equation (3). Note that to ensure that feasibility exists, the maximum capacity of the network (in units) has to exceed the minimum demand of the demand range (in units), for every product  $i$  (i.e.  $\sum_k Dmin_{i,k}$ ). This is not enforced as a constraint since there may be situations in which satisfying the minimum demand for specific products is indeed infeasible.

The upper limits on AT capacity for each AT site are considered in Equation (4). The flow balancing constraints on wafers transformed into units (die) are represented by Equation (5) and lastly fab product qualifications are set by Equation (6) which ensures that if fab  $j$  is not qualified for product  $i$  (such that  $Q_{ij} = 0$ ), then it would not produce any units for this product, but if it is qualified, then it would produce up to the maximum capacity of the fab.

## 4. Case Study

In this section, we demonstrate the proposed model discussed in the previous section by applying it to a case study inspired by a real semiconductor setting. We evaluate two cases, one in which the optimization solution is subject to having a specific fab running fully utilized (and at minimal average wafer cost), and the other in which the global optimization is attained.

### 4.1. Input Data

The dataset for the case study is depicted in **Tables 4-8**. The demand range, as expressed by minimum and maximum by product by customer (in units) is in **Table 4**. Similarly, in **Table 5** the ASP per product and customer is given. **Table 6** contains information about fab capacities, minimum utilizations, and qualifications.

**Table 7** depicts the AT capacity and production costs (in wafers and K units) and **Table 8** has the shipping costs from Fab to AT and from AT to the customer. Lastly, the CTR and DPW (die/unit per wafer) per product are given in **Table 9**. Note that in this case study, Products 1 to 3 are the baseline with a CTR of 1.0, and products 4, 5 require less/more capacity with a CTR of 0.9 and 1.2 respectively.

### 4.2. Results for the Case Study

Next, we demonstrate the usage of the model. The formulation in Section 3.3 was populated with the dataset from the previous section and executed using ILOG CPLEX Studio IDE Version 12.6.1.0. Execution time per instance is very fast

**Table 4.** Min and Max demand by product by customer (in units).

	MIN Demand in Dies				
	C1	C2	C3	C4	C5
PROD1	10,237,500	2,047,500	0	0	0
PROD2	0	10,725,000	21,092,500	0	6,077,500
PROD3	0	13,650,000	15,600,000	5,850,000	23,400,000
PROD4	35,100,000	0	0	0	23,400,000
PROD5	0	3,250,000	0	1,560,000	0

	MAX Demand in Dies				
	C1	C2	C3	C4	C5
PROD1	12,285,000	2,313,675	0	0	0
PROD2	0	14,657,500	31,460,000	0	8,937,500
PROD3	0	14,430,000	24,960,000	6,630,000	37,050,000
PROD4	53,300,000	0	0	0	36,400,000
PROD5	0	3,770,000	0	1,690,000	0

**Table 5.** ASP per product and customer.

	MAX Price Per Die				
	C1	C2	C3	C4	C5
PROD1	68	71	0	0	0
PROD2	0	127	121	0	134
PROD3	0	98	106	112	99
PROD4	5.5	0	0	0	6.5
PROD5	0	630	0	650	0

	MIN Price Die				
	C1	C2	C3	C4	C5
PROD1	57	67	0	0	0
PROD2	0	97	91	0	104
PROD3	0	94	86	100	79
PROD4	4	0	0	0	5
PROD5	0	610	0	630	0

**Table 6.** Fab capacities, minimum utilizations, and qualifications.

	Capacity Fabs		Min Util
	Week	Qtr.	
Fab1	10,000	130,000	60%
Fab2	7800	101,400	65%
Fab3	7500	97,500	70%

	PROD1	PROD2	PROD3	PROD4	PROD5
Fab1					
Fab2					
Fab3					

**Table 7.** AT capacity and production costs (in wafers and K units).

	Capacity AT	
	Week	Qtr.
AT1	8500	110,500
AT2	10,000	130,000
AT3	9000	117,000

	Cost AT	
	Wafer	K Dies
AT1	110	120
AT2	100	90
AT3	115	115

**Table 8.** Shipping costs: Fab-to-AT and AT-to-Customer.

	Fab-to-AT wafer shipping Cost		
	AT1	AT2	AT3
Fab1	30	30	30
Fab2	75	75	75
Fab3	50	55	50

	AT-to-Customer Cost K die				
	C1	C2	C3	C4	C5
AT1	53	51	60	58	55
AT2	53	52	56	54	54
AT3	58	56	56	52	51

**Table 9.** Capacity Trade Ratios and DPW (die/unit per wafer).

PRODUCT	CTR	DPW	
		PRODUCT	DPW
PROD1	1	PROD1	630
PROD2	1	PROD2	550
PROD3	1	PROD3	600
PROD4	0.9	PROD4	2000
PROD5	1.2	PROD5	100

(1.25 seconds on an Intel Core i5 5200U CPU at 2.20 GHz.) **Table 10** contains the details of the optimal solution in two cases:

Case (a): Fab-specific optimal solution

The first case that is evaluated is the case where the optimization is solved subject to the requirement that a specific fab (in this case study, Fab2) would achieve minimum wafer cost and chooses each subsequent variable following a greedy algorithm for the specific stage of the supply chain be fully utilized (i.e., 100% utilization).

Case (b): Globally optimal solution

The second case that is evaluated is the case of global optimization without any additional constraints such as in case (a).

As can be depicted by the left-hand-side in **Table 10**, by enforcing a minimum wafer cost at 100% utilization on Fab2 as in case (a), the resultant total fab production costs are higher and the revenues drop. Consequently, the profit is lower. On the other hand, when allowing the model to find the global optimum, depicted by the right-hand-side in the table, it identifies a solution at which none of the fabs is fully utilized but the total profits are significantly higher (by 3.4%).

**Table 10.** Optimal solutions for the case study.

	Fab specific Optimal					Globally Optimal				
	F1		F2		F3	F1		F2	F3	
Fab Utilization	1.00%		1.00%		1.00%	100.00%		100.00%	100.00%	
Avg Wfr Cost	4141.16		3842.97		4151.23	3863.30		4018.45	4107.00	
Fab Production Cost	550,463,550		389,677,600		359,413,080	502,228,661		407,470,440	400,432,500	
FabToAT Shipping Costs	3,987,750		7,605,000		4,329,000	3,939,000		4,143,750	4,647,500	
			15,921,750					12,730,250		
AT Utilization	AT1		AT2		AT3	AT1		AT2	AT3	
	52.29%		100.00%		75.56%	94.53%		100.00%	73.89%	
			78.18%					89.47%		
AT Die Cost	13,637,520		17,582,760		16,056,300	11,490,050		13,000,000	9,941,750	
AT Wfr Cost	10,703,550		13,000,000		10,764,000	4,138,194		11,686,538	7,848,750	
AT Production			81,744,130					58,105,282		
ATTo Customer Shipping Costs	3,134,534		4,106,148		3,824,470	1,773,294		5,445,684	3,502,752	
			11,065,152					10,721,730		
Cost			1,408,285,262					1,391,688,863		
	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5
Revenue	889,200,000	4,971,070,000	5,009,420,000	1,878,500,000	3,464,719,514	889,204,094	5,191,162,039	4,929,698,053	1,719,901,188	3,955,006,953
			16,212,909,514					16,684,972,326		
Profit			14,804,624,253					15,293,283,464		

### 4.3. Sensitivity Analysis

The results of the case study demonstrate the importance of searching for the global optimal solution for the network. In this section, we extend upon the case study that was presented and vary the inputs, to reflect different network sizes. Specifically, we executed the model for several values as follows:

- Number of fabs: 3, 5, and 7.
- Number of ATs: 3, 5, and 7.
- Number of customers: 5, 10, and 20.
- Number of products: 20, 50, and 100.

A note on the computation time before we proceed is appropriate. An expected non-linear increase has been observed in the computation time as the problem size grows, but even with the large scale problems of 100 products, the solutions were obtained in approximately 400 seconds, a reasonable time by all means. For the smaller scale problems, computation time decreased drastically, with about 15 seconds for the 20 product scenarios and 60 seconds for the 50 product scenarios. However, for each QIP solution we compared an equivalent

rounded QP solution, where the decision variables were considered to be continuous in the solution process and the final solution was the rounded down to the nearest integer. Across 24 different instances, 3 for each scenario depicted in **Table 11** we found a maximum gap of  $3.55e-09$  % between the optimal solution of the QIP and the rounded QP solution. The cause for this negligible gap is due to the large volumes of the integer units in the decision variables. Having observed such minimal differences between the rounded QP and QIP solutions, we chose to solve the instances as rounded QPs since the run time for such a rounded QP solutions remained approximately 3.4 seconds for even the large models. We selected a subset of all possible combinations as this would suffice to provide insight. 20 instances were generated for each combination that was selected. For each instance two algorithms were tested. A greedy algorithm that first maximizes the revenue (considering only feasibility of the demands at the fabs and ATs) and then minimizes the supply chain costs. This algorithm was compared with the results from the proposed algorithm discussed in section 3.3, which maximizes profit.

The results are depicted in **Table 11**. The Profit Gain column shows the average increase in profit across the 20 instances generated when using the proposed profit maximizing algorithm instead of a greedy, revenue maximizing, algorithm while the Revenue Loss column shows decrease in revenue. The Ratio column provides the quotient of the Profit Gain/ Revenue Loss. A common practice in the industry today is to first focus on generating maximum sales (revenue) and let the companies' supply chain attempt to minimize the subsequent costs of fulfillment. However, the results of **Table 11** indicate that a "holistic" approach which considers both the revenues and the costs may considerably increase the profit with almost no impact on the overall revenue, generating between a \$126 - \$700 of profit for each dollar of revenue loss. Furthermore, the results indicate that the benefit of such an approach increases with the size and complexity of the network.

**Table 11.** Sensitivity analysis of the solution.

Fabs	ATs	Customers	Products	Instances	Profit Gain (avg)	Std. Dev.	Revenue Loss (avg)	Ratio
3	3	5	20	20	13,517,075	3,088,135	107,132	126.2
3	3	5	50	20	14,507,085	3,427,495	91,649	158.3
3	3	5	100	20	17,717,953	2,970,469	73,302	241.7
5	5	10	20	20	18,638,989	3,955,901	124,985	149.1
5	5	10	50	20	19,200,775	3,513,044	93,823	204.6
5	5	10	100	20	18,623,342	3,427,650	46,307	402.2
7	7	20	50	20	21,030,469	3,035,081	29,599	710.5
7	7	20	100	20	21,918,830	3,728,617	73,158	299.6

## 5. Concluding Remarks

In this paper we develop a framework for the key components of the supply chain of the semiconductor manufacturing process, including front-end (fab) operations and backend (assembly/test) operations. Within this framework we consider the conversion from wafers to units. Then we propose a QP formulation for the profit maximizing objective function with flexible demand ranges per product and fab capacities. We demonstrate the model by applying it to a case study that is based on an industry dataset and show how the solution varies between a local optimization of a single fab (for minimum wafer cost) and a globally optimal solution for the network.

Our model extends on previous work in several respects; most notable is the consideration of varying ASP's of any product by customer, to reflect segmentation in pricing for differences in quantities sold. Other extensions pertain to the flexibility in setting the demand for each product, and in setting capacity across the product mix such that each fab has a capacity range rather than a fixed number.

On top of these extensions, there are still opportunities for further work. One that immediately comes to mind is the consideration of a different objective function to compare with the proposed quadratic objective function. Other directions include the incorporation of stock points within the supply chain, as they are used in practice to mitigate changes in the demand, and the explicit consideration of sub-problems of the SC problem within the model framework, for example the capacity planning problem for each of the fabs and/or the AT sites.

## References

- [1] Simchi-Levi, D., Kaminsky, P. and Simchi-Levi, E. (2004) *Managing the Supply Chain: The Definitive Guide for the Business Professional*. McGraw-Hill, New York.
- [2] Jain, S. (2006) A Conceptual Framework for Supply Chain Modelling and Simulation. *The International Journal of Simulation and Process Modelling*, **2**, 154-167. <https://doi.org/10.1504/IJSPM.2006.012644>
- [3] Ayers, J.B. (2006) *Handbook of Supply Chain Management*. 2nd Edition. CRC Press, 640 p.
- [4] Guin, U., Huang, K., DiMase, D., Carulli, J.M., Tehranipoor, M. and Makris, Y. (2014) Counterfeit Integrated Circuits: A Rising Threat in the Global Semiconductor Supply Chain. *Proceedings of the IEEE*, **102**, 1207-1228. <https://doi.org/10.1109/JPROC.2014.2332291>
- [5] Chien, C.F., Dauzere-Peres, S., Ehm, H., Fowler, J.W., Jiang, Z., Krishnaswamy, S., Mönch, L. and Uzsoy, R. (2011) Modeling and Analysis of Semiconductor Manufacturing in a Shrinking World: Challenges and Successes. *European Journal of Industrial Engineering*, **5**, 254-271. <https://doi.org/10.1504/EJIE.2011.041616>
- [6] Chopra, S. and Meindl, P. (2001) *Supply Chain Management: Strategy, Planning, and Operation*. Prentice-Hall, Upper Saddle River, NJ.
- [7] Denton, B.T., Forrest, J. and Milne, R.J. (2006) IBM Solves a Mixed-Integer Pro-

- gram to Optimize Its Semiconductor Supply Chain, *Interfaces*, **36**, 386-399. <https://doi.org/10.1287/inte.1060.0238>
- [8] Padillo, J.M., Ingalls, R. and Brown, S. (1995) A Strategic Decision Support System for Supply Network Design and Management in the Semiconductor Industry, *Computers and Industrial Engineering*, **29**, 443-447. [https://doi.org/10.1016/0360-8352\(95\)00114-G](https://doi.org/10.1016/0360-8352(95)00114-G)
- [9] Heath, D.C. and Jackson, P.L. (1994) Modeling the Evolution of Demand Forecasts with Application to Safety Stock Analysis in Production/Distribution Systems. *IIE Transactions*, **26**, 17-30. <https://doi.org/10.1080/07408179408966604>
- [10] Stray, J., Fowler, J.W., Carlyle, W.M. and Rastogi, A.P. (2006) Enterprise-Wide Strategic and Logistics Planning for Semiconductor Manufacturing, *IEEE Transactions on Semiconductor Manufacturing*, **19**, 259-268. <https://doi.org/10.1109/TSM.2006.873399>
- [11] Rastogi, A.P., Fowler, J.W., Carlyle, W.M., Araz, O.M., Maltz, A., and Büke, B. (2011) Supply Network Capacity Planning for Semiconductor Manufacturing with Uncertain Demand and Correlation in Demand Considerations. *International Journal of Production Economics*, **134**, 322-332. <https://doi.org/10.1016/j.ijpe.2009.11.006>
- [12] Ehm, H., Wenke, H., Mönch, L., Ponsignon, T. and Forstner, L. (2011) Towards a Supply Chain Simulation Reference Model for the Semiconductor Industry. *Proceedings of the 2011 Winter Simulation Conference*, Phoenix, AZ, 11-14 December 2011, 2119-2130. <https://doi.org/10.1109/wsc.2011.6147925>
- [13] Aelker, J., Bauernhansl, T. and Ehm, H. (2013) Managing Complexity in Supply Chains: A Discussion of Current Approaches on the Example of the Semiconductor Industry. *Procedia CIRP*, **7**, 79-84.
- [14] Lowe, J.J. and Mason, S.J. (2016) Integrated Semiconductor Supply Chain Production Planning. *IEEE Transactions on Semiconductor Manufacturing*, **29**, 116-126.. <https://doi.org/10.1109/TSM.2016.2544202>
- [15] Habla, C. and Monch, L. (2008) Solving Volume and Capacity Planning Problems in Semiconductor Manufacturing: A Computational Study. *Simulation Conference, WSC2008. Winter*, Austin, TX, 7-10 December 2008, 2260-2266.
- [16] Kannegiesser, M., Günther, H.O., van Beek, P., Grunow, M. and Habla, C. (2009) Value Chain Management for Commodities: A Case Study from the Chemical Industry. *OR Spectrum*, **31**, 63-93. <https://doi.org/10.1007/s00291-008-0124-9>



**Submit or recommend next manuscript to SCIRP and we will provide best service for you:**

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact [jssm@scirp.org](mailto:jssm@scirp.org)