

半導体生産ラインに対するWIP制御への数学的アプローチ

Mathematical Approach to WIP Control Mechanism for Semiconductor Production Line

國信 茂太

吉田 琢史

■ KUNINOBU Shigeta

■ YOSHIDA Takufumi

半導体製造では、生産ラインを通過する単位時間当たりのロット数（スループット）や生産ラインを通過するのに要する時間（TAT: Turn Around Time）が評価指標として利用される。適切なスループットとTATで生産するため、生産ライン内のロット（WIP: Work in Progress）数に上限を設けるWIP制御と呼ばれる制御方式が知られている。

東芝は、適切にWIP数の上限を設定できるようにするため、待ち行列理論を応用した数学的アプローチにより、WIP数の上限をパラメータとしたスループットとTATの近似式を導出した。対象とする生産ラインモデルは、WIP制御の仕組みと、生産ラインへのロットの到着間隔及び装置の処理時間のばらつきを考慮している。近似式の導出により、スループット及びTATが、生産ラインシミュレータを用いた場合と比較して非常に高速に計算でき、精度も同程度であることを確認した。

The throughput and the turnaround time (TAT), which are in a tradeoff relationship, are generally used as indexes to evaluate the production efficiency of a semiconductor production line. The adoption of a work-in-progress (WIP) control mechanism, which restricts the number of WIPs in a production line so as not to exceed the upper limit of WIP, k , has recently become common to facilitate production line operation while maintaining the appropriate throughput and TAT.

With this as a background, Toshiba has proposed $\phi(k)$ and $T(k)$ as approximations for throughput and TAT, respectively, in a semiconductor production line model, derived by a mathematical approach based on a queuing theory capable of determining an optimal k . This model takes into consideration the WIP control mechanism as well as variations in the inter-arrival times of lots and in the machine processing times. We have conducted evaluation experiments with a production line model consisting of about 20 processes using actual line data, and confirmed that the proposed approximations achieve higher calculation speed and almost the same accuracy in comparison with conventional simulation results.

1 まえがき

半導体製造では、生産ラインを通過する単位時間当たりのロット数（スループット）や生産ラインを通過するのに要する時間（TAT: Turn Around Time）が評価指標として利用される。スループットを最大にするためには、装置に絶え間なくロットを供給できるように装置前にロットをためておく必要がある。しかし、一方で装置前にロットをためておくと、ロットに処理待ち時間が発生するためTATが長くなる。

適切なスループットとTATで生産するため、生産ライン内のロット（WIP: Work in Progress）数に上限を設けるWIP制御と呼ばれる仕組みが知られている。一般に、WIP数の上限を高く設定するとスループットが上がってTATが長くなり、逆に低く設定するとスループットが下がってTATが短くなる。

東芝は、半導体生産ラインをモデル化し、待ち行列理論を応用した数学的アプローチにより、WIP数の上限 k とスループットとの関係 $\phi(k)$ 、及び k とTATの関係 $T(k)$ の近似式を導出した。これにより、スループット及びTATを、生産ラインシミュレータを利用して求める場合と比較して、はるかに高速に得られる。生産ラインの管理者は、 $\phi(k)$ と $T(k)$ を用いてスループットとTATとのバランスを見ながら、適切にWIP数

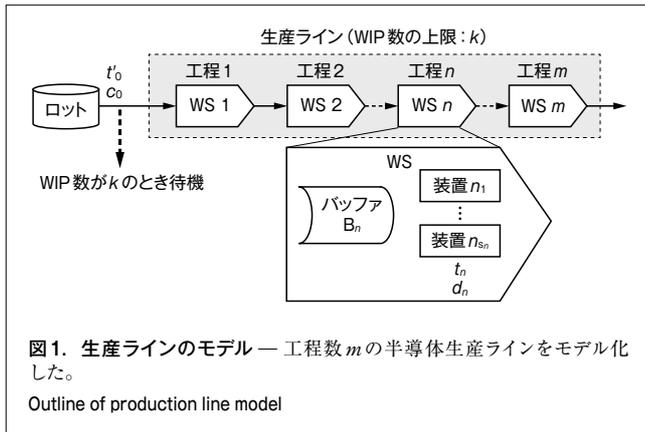
の上限を設定できるようになる。

2 生産ラインのモデル化

半導体生産ラインの全ての動きを考慮して $\phi(k)$ と $T(k)$ の近似式を導出することが理想であるが、複雑な動きの全てを考慮することは困難である。そこで生産ラインの本質を失わない範囲で、すなわちWIP制御を実運用するうえで最低限必要な生産ラインの特徴が反映されるようにモデル化する。

ここでは、半導体生産ラインを図1のようにモデル化する。WIP制御を実運用するうえで最低限必要な仕組みと、生産ラインへのロットの到着間隔及び装置の処理時間のばらつきを考慮できるモデルになっている点が特長である。

図1は工程数が m の生産ラインを表している。工程 n ($n=1, \dots, m$)は装置が空き待ち用バッファ B_n と s_n 台の装置から成るワークステーション（WS）で構成される。バッファ B_n の容量に制限はないと仮定するが、WIP数の上限は k であり、 $k+1$ 個目のロットは生産ライン直前で待機させる。待機ロットが存在する状態では新たにロットは生産ラインに到着しないと仮定する。ここで、生産ラインに到着するロットの平均時間間隔を t_0 、その変動係数（標準偏差を平均値で割った値）を c_0 とす



る。また、工程 n の装置1台がロットを処理するのに要する平均時間を t_n 、その変動係数を d_n とする。 t'_0 、 c_0 、 t_n 、 d_n が入力パラメータであり、これらを用いて $\phi(k)$ 、 $T(k)$ を導出する。

3 $\phi(k)$ 及び $T(k)$ の導出

待ち行列理論の分野では、様々なモデルに対する厳密式や近似式が知られている。特に、変動係数が1に限定されたモデルや、工程数が一つのモデル、各工程の装置台数が1台に限定されたモデル、WIP数の上限のないモデルに関しては、厳密式や多くの近似式が知られている⁽¹⁾。ここでは、既存の研究では知られていない、図1に示した生産ラインモデルに対する近似式 $\phi(k)$ 及び $T(k)$ を、次の三つのアイデアで導出する。

- (1) WIP数の上限が k のとき、生産ライン直前で待機ロットが存在する確率(待機率) $P_B(k)$ を導出することで、WIP数の上限に対応する。
- (2) 工程数が1で、到着するロットの平均時間間隔に対する変動係数が1に限定されたモデルを、任意の変動係数とれるように拡張する。
- (3) (2)で得られる各工程に対する近似式をモンテカルロ法により足し込むことで、複数工程が直列につながったモデルに対応する。

3.1 $P_B(k)$ と $\phi(k)$

$\phi(k)$ を導出するにあたり、 $P_B(k)$ の導出を考える。生産ラインを通過するロットの $\phi(k)$ は $P_B(k)$ の分だけ減るため、式(1)が成り立つ。すなわち、 $P_B(k)$ が導出できれば $\phi(k)$ が求まる。

$$\phi(k) = \frac{(1 - P_B(k))}{t'_0} \quad (1)$$

$P_B(k)$ は k が十分に大きい場合には0となることは明らかである。そして、 k が小さくなるにつれ $P_B(k)$ の値が増え始め、 $k=0$ では、生産ライン内にロットが入れない状態、すなわち $P_B(k)=1$ となる。

初めに、 k が十分に小さい場合の $P_B(k)$ を考える。図1のよ

うに複数工程が直列につながっているモデルでは、一般にロットの到着間隔に比べ、装置が処理に要する時間の和は十分に長いと考えられる。すなわち、式(2)が成り立つ。

$$t'_0 \ll \sum_{n=1}^m t_n \quad (2)$$

そうすると、生産ラインからロットが一つ出ていく間に、生産ライン直前に待機ロットが発生するような状況が発生する。すなわち、WIP数は k で維持されることになる。更に、 k が小さい場合には、ロットがバッファで待たされることも少ないため、式(3)と考えることができる。

$$T(k) = \sum_{n=1}^m t_n \quad (3)$$

リトルの公式⁽¹⁾から、式(4)が求まる。

$$\phi(k) = \frac{k}{\sum_{n=1}^m t_n} \quad (4)$$

式(1)と式(4)から、 k が十分に小さいときの $P_B(k)$ が式(5)で求まる。

$$P_B(k) = 1 - \frac{t'_0}{\sum_{n=1}^m t_n} k \quad (5)$$

次に、 k が大きい場合の $P_B(k)$ を考える。 k を ∞ から小さくしていくとき、どの程度 k が小さくなると $P_B(k)$ が増え始めるかを考えるため、WIP数の上限が存在しない場合の生産ラインモデルに対する状態確率分布 $P(w)$ を求める。 $P(w)$ はWIP数が w である確率を表す。 $P(w)$ を求める手順は次のとおりである。ここで、工程 n のWS内のロット数を w_n とし、工程 n の対する状態確率分布を $P_n(w_n)$ とする。

- (1) 工程数が1で $c_0=1$ に限定されたモデルに対する状態確率分布⁽²⁾を拡張し、任意の c_0 で利用可能な状態確率分布 $P_n(w_n)$ を求める。
- (2) 全工程について $P_n(w_n)$ を求め、それらをモンテカルロ法⁽³⁾を用いて足し込むことで、生産ラインモデルに対する状態確率分布 $P(w)$ とする。

$P_n(w_n)$ は $c_0=1$ に限定されたモデルにおける状態確率分布⁽²⁾を拡張し、式(6)のように導出した。

$$P_n(w_n) = \begin{cases} \frac{(s_n u_n)^{w_n}}{w_n!} P_n(0) & (w_n < s_n) \\ b_n^{w_n - s_n} (1 - b_n) \frac{(s_n u_n)^{s_n}}{s_n! (1 - u_n)} P_n(0) & (w_n \geq s_n) \end{cases} \quad (6)$$

$$P_n(0) = \left(\sum_{i=0}^{s_n-1} \frac{(s_n u_n)^i}{i!} + \frac{(s_n u_n)^{s_n}}{s_n! (1 - u_n)} \right)^{-1}$$

$$u_n = t_n / s_n t_0 \quad (u_n < 1)$$

$$b_n = \frac{(c_n^2 + d_n^2) u_n}{2 - (2 - c_n^2 - d_n^2) u_n}$$

ここで、 u_n は工程 n の装置の稼働率、 c_n は工程 n へ到着するロットの平均時間間隔に対する変動係数である。

c_n は、式(7)のリンクング方程式⁽⁴⁾により求めることができる。ただし、 $c_1=c_0$ と仮定する。

$$c_{n+1}^2 = 1 + (1 - u_n^2)(c_n^2 - 1) + \frac{u_n^2}{\sqrt{s_n}}(d_n^2 - 1) \quad (7)$$

w が十分に大きく $P(w)$ が0に近い領域では、 $P_B(w)$ も0に近いはずである。また、 w を十分に大きい値から小さくしていくと $P(w)$ が増え始め、そこで $P_B(w)$ も増え始めると考えられる。そこで、 k が十分に大きい領域から $P(k)$ が増え始める領域までの $P_B(k)$ を、式(8)とする。

$$P_B(k) = P(k) \quad (8)$$

最後に、 k が小さくも大きくもない領域における $P_B(k)$ は、 k が小さい領域の式と k が大きい領域の式とを直線でつなぐことで構成する。このとき $P_B(k)$ は単調減少関数とし、 $P_B(k) \geq P(k)$ となるように、 k が小さくも大きくもない領域の $P_B(k)$ を、式(9)とした。

$$P_B(k) = ak + (P(k_2) - ak_2) \quad (9)$$

ここで、 k_2 は $P(k)$ における傾き最小の点を与える k であり、 a はそのときの傾きである。

これまで述べた生産ラインモデルに対する $P_B(k)$ をまとめると、図2のようになる。ここで、 k_1 は式(1)と式(9)との交点となる k で、 k が十分に小さい領域の上限値である。

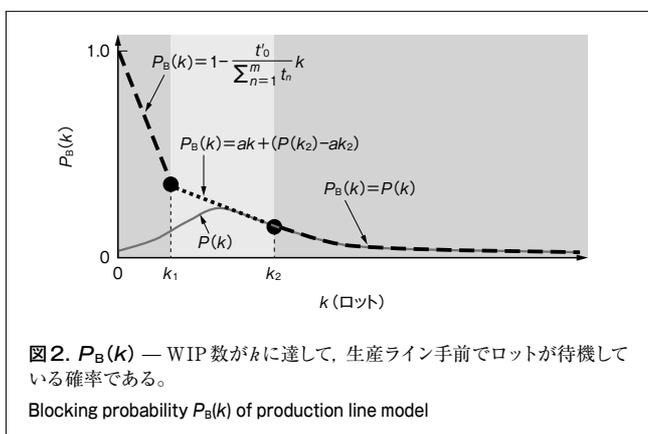
$P_B(k)$ が計算できると、 $\phi(k)$ は式(1)により計算できる。

3.2 WIP数と $T(k)$

WIP数の上限を k としたときのWIP数を $W(k)$ とすると、 $T(k)$ はリトルの公式⁽¹⁾により、式(10)で計算できる。

$$T(k) = \frac{W(k)}{\phi(k)} \quad (10)$$

ここでは、 $W(k)$ の導出について述べる。WIP数の上限が



ない($k=\infty$)場合、工程 n のWS内に存在するロット数 W_n を計算する近似式は既に知られている⁽⁵⁾。 $W(k)$ は各工程内に存在するロット数の和とする。すなわち、式(11)となる。

$$W(k) = \sum_{n=1}^m W_n \quad (11)$$

しかし、文献(5)では $k=\infty$ を前提としており、WIP数の上限が考慮されていない。そこで、 W_n を計算する際に利用するスループロットとして、式(1)の $\phi(k)$ を利用することで、 k を考慮した、すなわち k を含む式として利用する。

4 シミュレーション結果との比較

式(1)及び式(10)の近似式の精度を評価するため、生産ラインシミュレータを用いたシミュレーションの結果と比較した。今回、検証対象としたデータは、実データを基に作成した20工程から成るデータで、 $t'_0=100$ 、 $c_0=1$ の他、各工程の装置データは表1のとおりである。また、 $P_B(k)$ の計算におけるモンテカルロ法のサンプリングは3,000,000回行っている。計算プログラムはJava^(*) Version1.7を用いて実装した。

生産ラインシミュレータはNetLogo⁽⁶⁾ Version5.0.4を用いて実装した。シミュレーションでは、実データを分析した結果を基に、 t'_0 は指数分布、 t_m はガンマ分布に従うとした。また、使用した実行環境は、Intel[®] Core[™] i5-2400, 3.1 GHz, 4 GバイトRAMの計算機である。

$\phi(k)$ 及び $T(k)$ をシミュレーション結果と比較して、それぞれ図3及び図4に示す。いずれの場合も誤差が5%以内に収まっていることが確認できた。

生産ラインシミュレータでは、WIP数の一つの上限値を例えば $k=K$ として、 $\phi(K)$ 及び $T(K)$ を計算するのに1~数分間程度の時間が必要であった。一方、導出した近似式から計

表1. 工程 n の装置データ

Machine data at n -th process

n	1	2	3	4	5	6	7
t_n	2,472.0	175.3	8,082.1	1,723.0	2,301.8	1,829.0	75.4
d_n	0.53	0.29	0.50	0.30	0.29	0.35	0.37
s_n	26	2	85	18	24	19	1

n	8	9	10	11	12	13	14
t_n	773.3	438.0	1,408.5	523.3	127.8	917.5	5,511.0
d_n	0.50	0.80	0.30	0.90	0.74	0.60	0.20
s_n	8	5	15	6	2	10	56

n	15	16	17	18	19	20
t_n	9,882.0	127.8	2,461.7	1,123.0	826.5	2,221.0
d_n	0.70	0.40	0.80	0.15	0.46	0.33
s_n	102	2	26	12	9	23

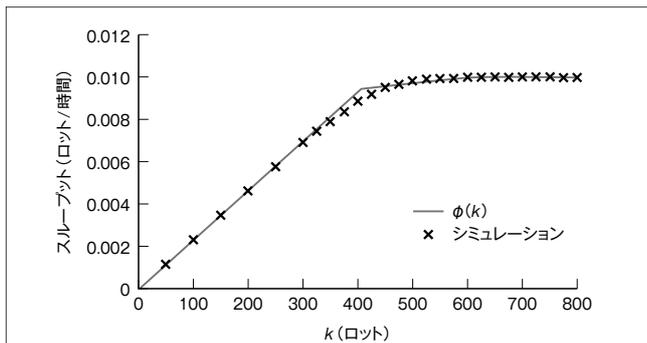


図3. $\phi(k)$ とシミュレーション結果との比較 — $\phi(k)$ を用いて計算したスループットとシミュレーションによって得たスループットを比較した結果、誤差は5%以内である。

Comparison of results of throughput profiles obtained by approximation $\phi(k)$ and simulation

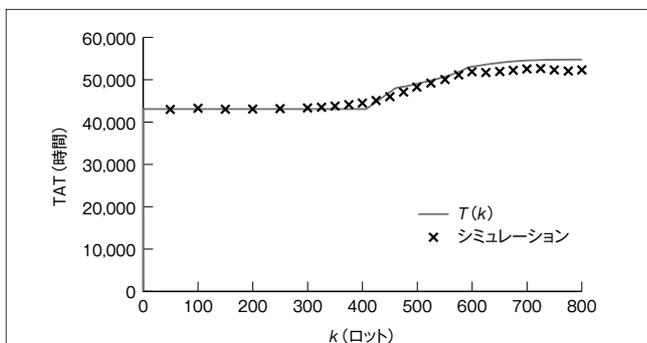


図4. $T(k)$ とシミュレーション結果との比較 — $T(k)$ を用いて計算したTATとシミュレーションによって得たTATを比較した結果、誤差は5%以内である。

Comparison of results of TAT profiles obtained by approximation $T(k)$ and simulation

算すると7秒間程度で $k=1 \sim 800$ までの $\phi(k)$ 及び $T(k)$ を計算できた。ここで、 $k=800$ は、 $\phi(k)$ 及び $T(k)$ の値がほとんど変化しなくなる値として計算された値である。

5 あとがき

半導体生産ラインにおいてWIP制御を行う際、最適な k を決定するために、 $\phi(k)$ 及び $T(k)$ を待ち行列理論を用いた数学的アプローチにより導出した。生産ラインの管理者は、これらの関係を参照し、スループットは下がらないが、TATが短くなるような k を設定できるようになる。

生産ラインシミュレータでは一つの $k=K$ に対して、 $\phi(K)$ 及び $T(K)$ を計算するのに1～数分間程度の時間が必要であるのに対し、導出した近似式から計算すると7秒間程度で $k=1 \sim 800$ までの $\phi(k)$ 及び $T(k)$ を計算でき非常に高速である。また精度も、シミュレーションの結果との差は5%未満であり、十分な精度であることを確認した。

この研究成果は、半導体生産ラインを対象として構成した生産ラインモデルに対する近似式 $\phi(k)$ 及び $T(k)$ を導出したことである。待ち行列理論の分野における一つの成果ではあるが、対象とした生産ラインモデルは、実際の半導体生産ラインにおける全ての動きを完全にモデル化してはいない。したがって、この研究の結果をそのまま実運用することはできないが、別途、運用面及び技術面で様々な工夫もしており、実運用での利用を可能にして半導体の生産性向上に役だてる。

文献

- (1) 大石進一. 待ち行列理論. 東京, コロナ社, 2003, 143p.
- (2) 近藤 浩. "M/G/s 待ち行列の定常状態分布の近似(1)". CUSCUS工場統計力学 (建設中!). 2011-03-01. <<http://d.hatena.ne.jp/CUSCUS/20110301>>, (参照 2015-12-02).
- (3) 津田孝夫. モンテカルロ法とシミュレーション. 三訂版. 東京, 培風館, 2002, 267p.
- (4) Adan, I.J.B.F. et al. "Analysis of Manufacturing Systems using Chi 3.0". Technische Universiteit Eindhoven, Department of Mechanical Engineering Systems Engineering Group. 2012-10. <<http://www.win.tue.nl/~iadan/4t400/lecture-notes-ams.pdf>>, (accessed 2015-12-14).
- (5) Kimura, T. "Interpolation approximations for the mean waiting time in a multi-server queue. Journal of the Operations Research Society of Japan. 35, 1, 1992, p.77 - 92.
- (6) Wilensky, U. "NetLogo". NetLogo. <<http://ccl.northwestern.edu/netlogo/>>, (accessed 2015-12-02).

- Javaは、Oracle Corporation及びその子会社、関係会社の米国及びその他の国における登録商標。
- Intel, Intel Coreは、Intel Corporationの米国又はその他の国における商標。



國信 茂太 KUNINOBU Shigeta, Ph.D.

研究開発統括部 研究開発センター システム技術ラボラトリー主任研究員, 博士 (工学)。最適化理論に関する研究に従事。電子情報通信学会会員。System Engineering Lab.



吉田 琢史 YOSHIDA Takufumi, Ph.D.

研究開発統括部 研究開発センター システム技術ラボラトリー主任研究員, 博士 (理学)。最適化理論に関する研究に従事。日本オペレーションズ・リサーチ学会, 日本物理学会会員。System Engineering Lab.