Optimal control of re-entrant manufacturing systems

Matthias Kawski †

School of Mathematical and Statistical Sciences

Arizona State University

http://math.asu.edu/~kawski

HYP 2012 – 14$^{\text{th}}$ Internat. Conf. on Hyperbolic Problems

Based on joint work with Jean-Michel Coron (Paris VI) & Zhiqiang Wang (Fudan), K. Kempff (INTEL), D. Armbruster, C. Ringhofer, D. Marthaler, M. LaMarca (ASU).

† This work was partially supported by the National Science Foundation through the grants DMS 05-09030 and 09-08204, the University Pierre and Marie Curie-Paris VI and the Foundation Sciences Mathématiques de Paris.

Outline

- Semiconductor manufacturing
  Some background and mathematical problems

- A hyperbolic conservation law

$$\partial_t \rho(t,x) + \partial_x \left(\lambda(W(t))\, \rho(t,x) \right) = 0$$

$$W(t) = \int_0^1 \rho(t,x)\, dx, \quad \text{later specialize to} \quad \lambda(W) = \frac{1}{1+W}$$

  - Motivation
  - Intuitive properties, numerical studies
  - Analysis: existence[ZW], optimality, L^1 vs L^2, minimum-time

- More control problems under investigation

Selected immediately related references

- D. Armbruster, D. Marthaler, C. Ringhofer, K. Kempf, and T.-C. Jo, *A Continuum Model For A Re-Entrant Factory*, Oper. Res., **54** (2006), 933–950.
- M. La Marca, D. Armbruster, M. Herty, and C. Ringhofer, *Control of continuum models of production systems,* IEEE Trans. Automat. Control **55** (2010), no. 11, 2511–2526.
- J.-M. Coron, M. Kawski, and Z. Wang, *Analysis of a conservation law modeling a highly re-entrant manufacturing system*, Discrete Contin. Dyn. Syst. Ser. B **14** (2010), no. 4, 1337–1359.
- R. Colombo, M. Herty, and M. Mercier, *Control of the continuity equation with a non local flow*, ESAIM Control Optim. Calc. Var., **17** (2011), no. 2, 353–379.

- vast body of related literature

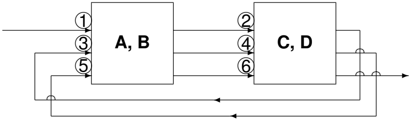
Semiconductor manufacturing: Distinguishing features

- Product: small size, high value, global supply network
- Volatile demand, difficult to predict yields (processor speed, energy consumption, ...). Stochastics everywhere
- Typical 20 day pure processing time, 60 day start2finish, compare to demand forecast horizon
- Very short product life times, never in equilibrium
- New “fab” every few years & huge capital cost, utilization
- Layered (sandwich) structure: highly re-entrant manufacturing line (e.g., 600 processes, 200 stations)

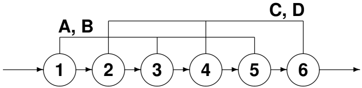
Supply chains, collaboration ASU and INTEL

- Several new fabs in Chandler, AZ
- 90s: scheduling processes/machines.
  E.g., chaos, queuing models, discrete event systems.
- Hierarchical control, inner-outer-loop, “model predictive”
- 2000: global supply network, many products
- “downbinning” – stochastic optimal control
- early 2000s: fast computation using PDE models
  compare gas dynamics, traffic flow. (K.Kempff (INTEL), D.Armbruster, C.Ringhofer, D.Marthaler, M.LaMarca, . . .)

Toy example: M sets of machines, N processing steps



The machine-based view



The process-based view

Real world

- approx $N \approx 600$ production steps total, up to 20 loops
- approx $M \approx 200$ work-stations, up to 20 parallel machines
- total processing time approximately 20 days
- total manufacturing time approximately 60 days
- several different products (common: shared initial steps)

Sample data for toy example: queues and idling

Topology M machines N process steps. $\Phi = (p_{ij})$ where $p_{ij} = 1$ if machine i can carry out process j, and $p_{ij} = 0$ else.

$$\Phi = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{200\times 600}$$

Process times τ_{ij} is the time it takes machine i to complete process step j.

$$\tau = \begin{pmatrix} 11 & 17 \\ 11 & 17 \end{pmatrix}$$

Parallel batching β_{ij} number of parts that must be batched together in machine i for process step j.

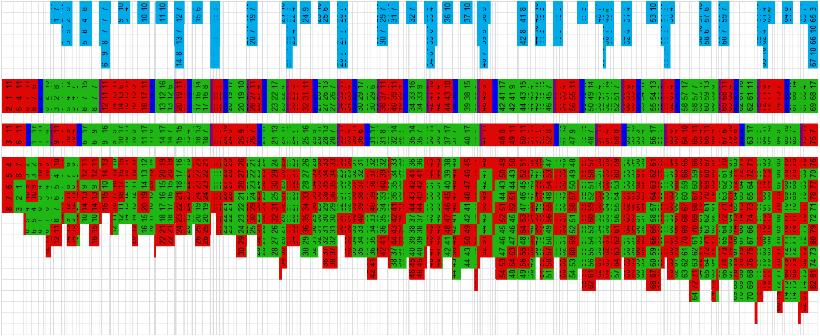
$$\beta = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$$

Set-up times For each machine i let $\sigma^i = (\sigma^i_{jk})$ set-up time after process step j before process step k may start

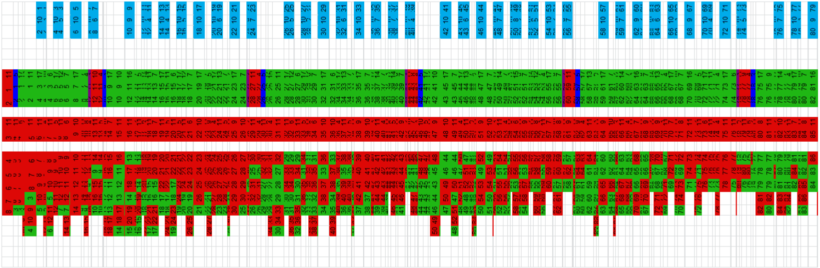
$$\sigma^1 = \sigma^2 = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$$

Playground (and careers!), DES simulation

A FIFO, B FIFO (“first in first out”):
Here, start rate below capacity, but queues grow due to *chaotic* switchings, requiring too many set-ups



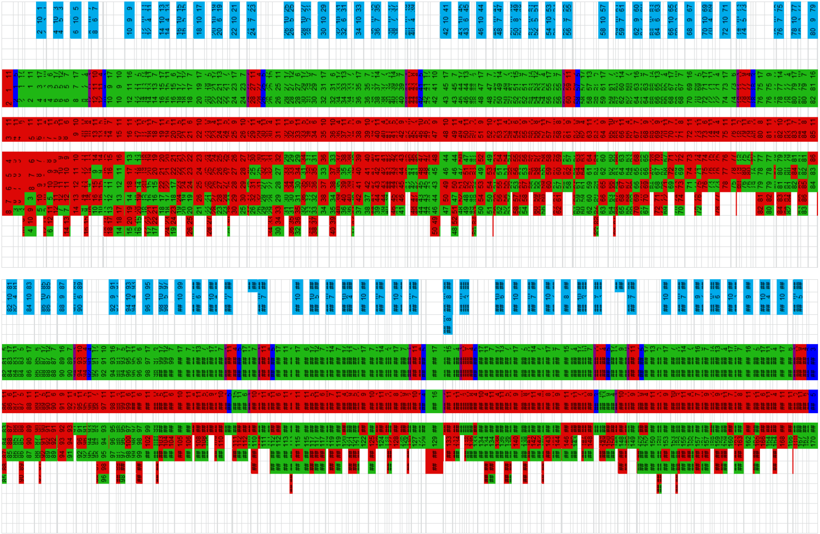
DES simulation: A “PUSH”, B “PULL”



popular policies: if *oven door is open*, work on (if available)

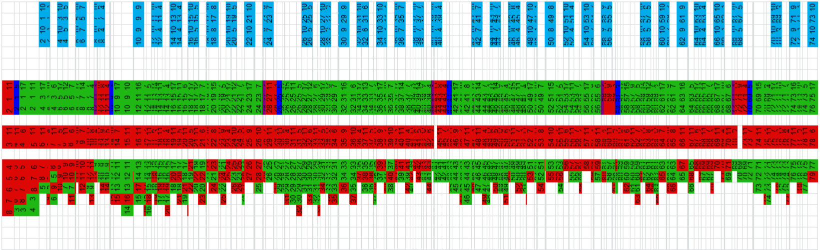
product that is least finished [PUSH], and
product that is closest to being finished [PULL]

DES simulation: A “PUSH”, B “PULL”



DES simulation: idling is better!

machine A: step 1 or **wait**, machine B: PULL



Real simulation: 600 steps, 400 machines, multiple products

Flow models

- popular: model the flow of parts through the fab by a PDE justified by very large number of parts and process stages supposedly superior numerical algorithms for simulation
- **First**: single product, single fab, $M = 1$ machine, $N = \infty$ reentries. speed depends on total load ("work in progress")

$$\partial_t \rho(t,x) + \partial_x \left(\lambda(W(t))\, \rho(t,x) \right) = 0, \quad W(t) = \int_0^1 \rho(t,x)\, dx,$$

- speed model $\lambda(W) = \frac{1}{1+W}$ supported by fab data after much heated discussion. $W = 0$ possible ???

Mathematical model (INTEL, Armbruster, Kempff)

- "Observed": Any increase in the total load (at any stage) slows down the entire fabrication line.
- Hyperbolic conservation law

$$\partial_t \rho(t,x) + \partial_x \left(\lambda(W(t))\, \rho(t,x)\right) = 0$$

- $0 \leq t \leq T$ time, $0 \leq x \leq 1$ degree of completion
- $\rho \colon [0,1] \times [0,T] \mapsto \mathbb{R}$ WIP-density ("work in progress")
- $W(t) = \int_0^1 \rho(t,x)\, dx$ total load (WIP)
- $\lambda(W)$ global speed, later specialize to $\lambda(W) = \frac{1}{1+W}$

Control, outputs, practical objectives

- Control: influx $u(t) = \rho(t,0)\lambda(W(t))$
  later: PUSH-PULL-point, allow varying local speeds

- Key objective: track desired output $y_d(t)$ by the outflux
  $y(t) = \rho(t,1)\lambda(W(t))$ [perishable demand]

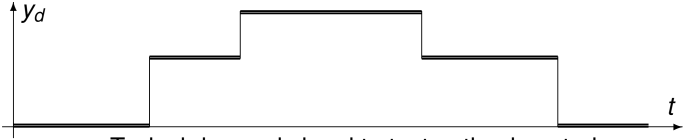
- Usual: backlogs allowed but penalized
  track cumulative demand $Y(t) = \int_0^t y_d(\tau)\,d\tau$
  by accumulated outflux $Y(t) = \int_0^t y(\tau)\,d\tau$

- controllability: What demands y_d or Y_d can be tracked?
  Initial load $\rho(0,x)$ matters (reversed time system?) [ZW]

- optimal control: minimize tracking error
  $\int_0^T |y(\tau) - y_d(\tau)|^p\,d\tau$ or $\int_0^T \left| \int_0^t (y(\tau) - y_d(\tau))d\tau \right|^p dt$

Earlier simulation results (M.LaMarca, D. Marthaler)

- For piecewise constant demand y_d to be tracked numerically calculate piecewise constant optimal input

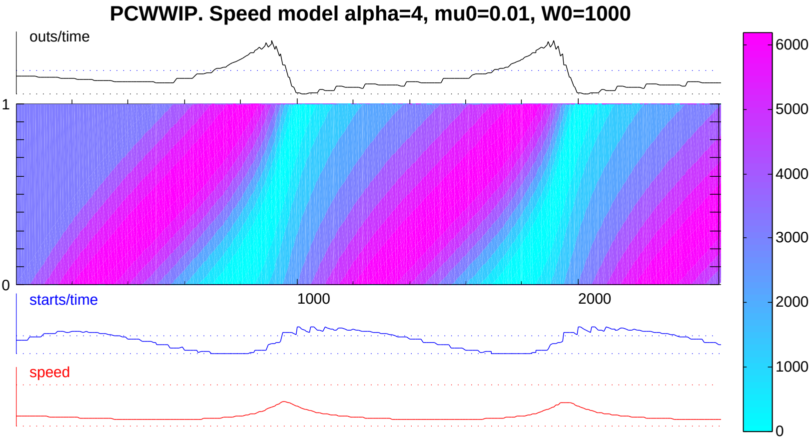


Typical demand signal to test optimal control

- formulate max principle, numerically solve adjoint equation
- results appear reasonable and useful for strategic planning (very outer loop), especially *inverse response*: Increasing start rate initially decreases output rate!

Simulation (MK, tracking characteristics)

pcw const density. Note: increased steepness, but NO shocks



J.-M Coron, MK, and Zhiqiang Wang (DCDS 2010)

Provide analytic foundations for simulations

- For L^1 data prove existence of unique weak L^1 solution
- For L^2 data & objective prove existence of optimal control
- For minimum-time transfer between equilibria using L^1 control prove optimality of natural candidate.

The hyperbolic conservation law with BC [ZW]

$$
\begin{aligned}
0 &= \partial_t \rho(t,x) + \partial_x \left(\lambda(W(t))\,\rho(t,x)\right) \\
W(t) &= \textstyle\int_0^1 \rho(t,x)\,dx,
\end{aligned}
\qquad (1)
$$

on $[0, T] \times [0, 1]$ or $[0, \infty) \times [0, 1]$.

Assume that $\lambda(\cdot) \in C^1([0, +\infty); (0, +\infty))$.

$$
\begin{aligned}
\rho(0,x) &= \rho_0(x), \quad \text{for} \quad 0 \le x \le 1, \quad \text{and} \\
\rho(t,0)\lambda(W(t)) &= u(t), \quad \text{for} \quad t \ge 0.
\end{aligned}
\qquad (2)
$$

Weak solution - standard definition [ZW]

Definition

Let $T > 0$, $\rho_0 \in L^1(0,1)$ and $u \in L^1(0,T)$ be given.
A weak solution of the Cauchy problem (1) and (2) is a function $\rho \in C^0([0,T]; L^1(0,1))$ such that,
for every $\tau \in [0,T]$ and every $\varphi \in C^1([0,\tau] \times [0,1])$ such that

$$\varphi(\tau, x) = 0, \forall x \in [0,1] \quad \text{and} \quad \varphi(t,1) = 0, \forall t \in [0,\tau], \tag{3}$$

one has

$$\begin{aligned} &\int_0^\tau \int_0^1 \rho(t,x)(\varphi_t(t,x) + \lambda(W(t))\varphi_x(t,x))dxdt \\ &\qquad + \int_0^\tau u(t)\varphi(t,0)dt + \int_0^1 \rho_0(x)\varphi(0,x)dx = 0. \end{aligned} \tag{4}$$

Existence of weak L^1-solution [ZW], (measures ?)

Theorem

If $\rho_0 \in L^1(0,1)$ and $u \in L^1(0,T)$ are nonnegative almost everywhere, then the Cauchy problem (1) *and* (2) *admits a unique weak solution $\rho \in C^0([0,T]; L^1(0,1))$, which is also nonnegative almost everywhere in $Q = [0,T] \times [0,1]$.*

Proof strategy:

- First prove the existence of weak solution for small times
- Strategy: Analyze the characteristic curve $\xi = \xi(t)$ with $\xi(0) = 0$, use it to construct solution to the Cauchy problem.
- Tool: Contraction mapping
- Key step: Change order of integration.

Contraction for characteristics [ZW], (measures ?)

Define $F : \Omega_{\delta,M} \to C^0([0,\delta]), \xi \mapsto F(\xi), \forall \xi \in \Omega_{\delta,M}, \forall t \in [0,\delta]$ by

$$F(\xi)(t) = \int_0^t \lambda\Big(\int_0^s u(\sigma)d\sigma + \int_0^{1-\xi(s)} \rho_0(x)dx\Big)ds.$$

Prove that, for $\delta > 0$ small, F is a contraction on

$$\Omega_{\delta,M} = \Big\{\xi \in C^0([0,\delta]) \colon \xi(0) = 0, \widetilde{\lambda}(M) \leq \frac{\xi(s)-\xi(t)}{s-t} \leq \overline{\lambda}(M), \forall s,t \in [0,\delta]\Big\}$$

with C^0 norm and $M = \|u\|_{L^1(0,T)} + \|\rho_0\|_{L^1(0,1)}$.
Key step: change order of integration.
Build candidate solution ρ from ξ, verify it is weak solution, and extend to large times. Note L^p and hidden regularity.

L^2-optimal control for demand tracking problem

For $y_d \in L^2_+(0,T)$ and $\rho_0 \in L^2_+(0,1)$,
let $\rho \in C^0([0,T], L^2(0,1)) \cap C^0([0,1], L^2(0,T))$
be the unique weak solution of the Cauchy problem (1) and (2),
write $y(t) = \rho(t,1)\lambda(W(t))$, and define $J\colon L^2_+(0,T) \mapsto \mathbb{R}$ by

$$J(u) = \int_0^T |u(t)|^2 dt + \int_0^T |y(t) - y_d(t)|^2 dt.$$

Theorem

There exists $u^* \in L^2_+(0,T)$ that minimizes J over $L^2_+(0,T)$

$$J(u^*) = \inf_{u \in L^2_+(0,T)} J(u). \tag{5}$$

Proof of existence of optimal L^2-solution

- Consider minimizing sequence $\{u_n\}_{n=1}^{\infty} \subseteq L^2_+(0,T)$
- Use uniform boundedness to extract converging subsequence $\{u_{n_k}\}_{k=1}^{\infty}$
- From corresponding sequence of characteristics $\{\xi_{n_k}\}_{k=1}^{\infty}$ extract converging subsequence with limit ξ_∞
- Construct associated weak solution ϱ_∞ of Cauchy problem
- Prove that $y_n(t) = \lambda(W_n(t))\rho_n(t,1)$ converges weakly in $L^2(0,T)$ to $y_\infty(t) = \lambda(W_\infty(t))\rho_\infty(t,1)$
- Verify $u\infty$ minimizes J in $L^2_+(0,T)$ and note $u_n \longrightarrow U_\infty$ strongly in $L^2_+(0,T)$.

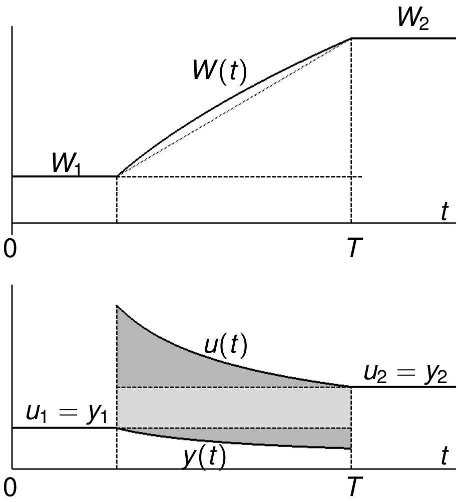
Time-optimal transition between equilibria

- Special case $\lambda(W) = \frac{1}{1+W}$
- Suppose $\rho_1 \geq \rho_0 \geq 0$ are constant, and for $x \in [0,1]$
  $\rho(0,x) = \rho_0$
  desired: $\rho(T,x) = \rho_1$ and some minimal $T > 0$.
- Note, backlog is not considered here.
- A natural choice $\rho(t,0) = \rho_1$ for $t \geq 0$. This determines $u(t) = \rho_1 \lambda(W(t))$ and $y(t) = \rho_0 \lambda(W(t))$, where W is a solution of

$$W'(t) = \frac{\rho_1 - \rho_0}{1 + W(t)}, \quad W(0) = \int_0^1 \rho(0,x)\, dx = \rho_0.$$

Easy calculations yield transfer time $T = 1 + \frac{\rho_0 + \rho_1}{2}$.

Minimum time transfer between equilibria

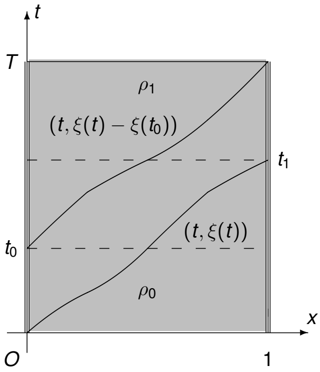


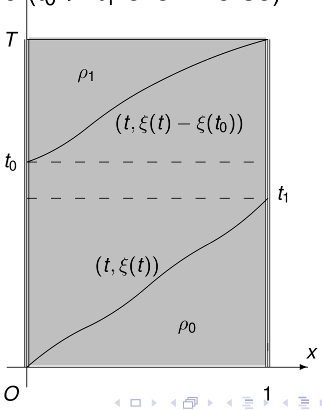
Proposition

The minimum time to transfer between equilibria $\rho(0,\cdot) = \rho_0$, *and* $\rho(T,\cdot) = \rho_1 > \rho_0$, *using* $u \in L^1([0,\infty),[0,\infty))$ *is* $T = 1 + \frac{\rho_0+\rho_1}{2}$.

Time optimal transfer between equilibrium states

Exists $t_0 \in [0, T]$ such that $\rho(t, 0) = \rho_1$ for $t_0 < t < T$.
Characteristic $\xi'(t) = \lambda(\int_0^1 \rho(t, x)\, dx)$ through ξ through $\xi(0) = 0$ defines $t_1 \in (0, T]$ such that $\xi(t_1) = 1$.
Various neat estimates for bounds on speed ξ'.
Two cases: show best choice $t_0 = 0$ ($t_0 > t_1$ even worse).





Time-optimal transfer between equilibria w/o backlog

- Suppose $\rho_1 \geq \rho_0 \geq 0$ constant, ($\rho_0 \geq \rho_1$ is different) still special case $\lambda(W) = \frac{1}{1+W}$ and $\rho(0,x) = \rho_0$
- desired: $\rho(T,x) = \rho_1$ and some minimal $T > 0$ and
$$Y(T) = \int_0^{t_{sw}} (y_t - y_0)\, dt + \int_{t_{sw}}^{T} (y_t - y_1)\, dt = 0.$$
- Numerical simulations and heuristics suggest there does not exist a minimizing L^1 control. Optimal control starts w/ pulse (Dirac delta), causing maximal initial slow-down but minimum time transition w/ zero backlog
- Note: negative pulses (negative start rates) are not admissible, hence an optimal step-down is different!
- Alternatives: Consider only L^1-inputs with bounded start rate $u(t) \leq \overline{u}$, or allow Borel measures as inputs (analogous existence proof for weak solns appears to work)

Problems: L^1, Cascades, several products, ...

- Optimality for L^1-cost (and inputs in L^1 or Borel measures)

$$J(u) = \int_0^T |y(t) - y_d(t)| dt. \quad \text{or} \quad J(u) = \int_0^T |Y(t) - Y_d(t)| dt.$$

- Cascade of systems with shared capacity
  added control: allocation of capacity to front/back
  (PDE model for PUSH / PULL policies ?)
  speed not constant (in x), but still depends on total load
- Location of push-pull-points as control
- Multiple products (vector valued HCL)

Selected immediately related references

- D. Armbruster, D. Marthaler, C. Ringhofer, K. Kempf, and T.-C. Jo, *A Continuum Model For A Re-Entrant Factory*, Oper. Res., **54** (2006), 933–950.
- M. La Marca, D. Armbruster, M. Herty, and C. Ringhofer, *Control of continuum models of production systems,* IEEE Trans. Automat. Control **55** (2010), no. 11, 2511–2526.
- J.-M. Coron, M. Kawski, and Z. Wang, *Analysis of a conservation law modeling a highly re-entrant manufacturing system*, Discrete Contin. Dyn. Syst. Ser. B **14** (2010), no. 4, 1337–1359.
- R. Colombo, M. Herty, and M. Mercier, *Control of the continuity equation with a non local flow*, ESAIM Control Optim. Calc. Var., **17** (2011), no. 2, 353–379.

- vast body of related literature

Simulation (MK, tracking characteristics)

piecewise constant density