Institute of Scientific Computing

Control of Laser Surface Hardening by a Memory Efficient Approach of Second Order

Julia Sternberg, Andreas Griewank

TU Dresden, HU Berlin

AD Days, Shrivenham 2005

Outline

1. Optimal control problem
2. Multiple sweep algorithm
3. Nested checkpointing
4. Numerical results
5. Summary and conclusion

1. Optimal control problem

$$\min_{\mathbf{u}\in\mathcal{U}_{ad}} \phi(\mathbf{x}(T)),$$

$$\dot{\mathbf{x}} = f(\mathbf{x}(t), \mathbf{u}(t), t), \qquad \mathbf{x}(0) = \mathbf{x}_0$$

Adjoint problem

$$\dot{\bar{\mathbf{x}}} = -H_{\mathbf{x}}^T = -f_{\mathbf{x}}^T\,\bar{\mathbf{x}}, \qquad \bar{\mathbf{x}}(T) = \phi_{\mathbf{x}}^T(\mathbf{x}(T))$$

First order necessary optimality condition

$$H_{\mathbf{u}}^T = f_{\mathbf{u}}^T\,\bar{\mathbf{x}} = 0, \qquad 0 \le t \le T$$

Express $\mathbf{u}(t)$ in terms of $\mathbf{x}(t)$ and $\bar{\mathbf{x}}(t)$

$$H_{\mathbf{u}}^T = f_{\mathbf{u}}^T\,\bar{\mathbf{x}} = 0 \Rightarrow \mathbf{u}(t) = \mathbf{u}(\mathbf{x}(t), \bar{\mathbf{x}}(t))$$

$\Rightarrow$ BVP with separated BC

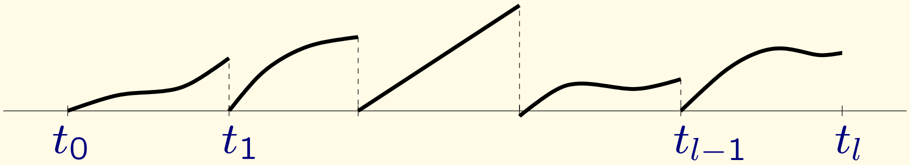
Multiple shooting technique

$$\begin{aligned} \dot{\mathbf{x}} &= f(\mathbf{x}(t), \bar{\mathbf{x}}(t)), & \mathbf{x}(0) &= \mathbf{x}_0 \\ \dot{\bar{\mathbf{x}}} &= f_{\mathbf{x}}^T(\mathbf{x}(t), \bar{\mathbf{x}}(t))\,\bar{\mathbf{x}}, & \bar{\mathbf{x}}(T) &= \phi_{\mathbf{x}}^T(\mathbf{x}(T)) \end{aligned}$$

Coose a mesh

$$0 = t_0 < t_1 < \ldots < t_{l-1} < t_l = T.$$

- Integrate IVP on each subinterval $[t_i, t_{i+1}]$
- Patch resulting solution segments over entire interval $[0, T]$
- Use separated BC and suitable decoupling



— Requires expression for $\mathbf{u}(t) = \mathbf{u}(\mathbf{x}(t), \bar{\mathbf{x}}(t))$
— Requires estimation for $\bar{\mathbf{x}}(0)$

Quasilinearization technique

Step $\delta\mathbf{x}$, $\delta\bar{\mathbf{x}}$, $\delta\mathbf{u}$ satisfy:

$$
\begin{aligned}
\delta\dot{\mathbf{x}} &= f_{\mathbf{x}}\delta\mathbf{x} + f_{\mathbf{u}}\delta\mathbf{u} && \textbf{linearized state} \\
\delta\dot{\bar{\mathbf{x}}} &= -H^T_{\mathbf{x}\mathbf{x}}\delta\mathbf{x} - H^T_{\mathbf{x}\mathbf{u}}\delta\mathbf{u} - H^T_{\mathbf{x}\bar{\mathbf{x}}}\delta\bar{\mathbf{x}} && \textbf{linearized adjoint} \\
0 &= H^T_{\mathbf{u}} + H^T_{\mathbf{u}\mathbf{u}}\delta\mathbf{u} + H^T_{\mathbf{u}\mathbf{x}}\delta\mathbf{x} + H^T_{\mathbf{u}\bar{\mathbf{x}}}\delta\bar{\mathbf{x}} && \textbf{linearized optimality}
\end{aligned}
$$

Separated boundary conditions:

$$
\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}
\begin{pmatrix} \delta\mathbf{x}(0) \\ \delta\bar{\mathbf{x}}(0) \end{pmatrix}
+
\begin{bmatrix} 0 & 0 \\ -\phi^T_{\mathbf{x}\mathbf{x}}(T) & I \end{bmatrix}
\begin{pmatrix} \delta\mathbf{x}(T) \\ \delta\bar{\mathbf{x}}(T) \end{pmatrix}
=
\begin{bmatrix} \mathbf{x}_0 - \mathbf{x}(0) \\ \phi^T_{\mathbf{x}}(T) - \bar{\mathbf{x}}(T) \end{bmatrix}
$$

$\Rightarrow$ Eliminate $\delta\mathbf{u}$ from **linearized optimality** $(det(H_{\mathbf{u}\mathbf{u}}) \neq 0)$

$\Rightarrow$ Linear BVP for $\delta\mathbf{x}$ & $\delta\bar{\mathbf{x}}$

Linear BVP

$$\begin{pmatrix} \delta\dot{\mathbf{x}}(t) \\ \delta\dot{\bar{\mathbf{x}}}(t) \end{pmatrix} = \begin{pmatrix} A^{11}(t) & A^{12}(t) \\ A^{21}(t) & A^{22}(t) \end{pmatrix} \begin{pmatrix} \delta\mathbf{x}(t) \\ \bar{\mathbf{x}}(t) \end{pmatrix} + \begin{pmatrix} q^1(t) \\ q^2(t) \end{pmatrix}$$

Introduce substitution

$$\begin{pmatrix} \delta\mathbf{x}(t) \\ \delta\bar{\mathbf{x}}(t) \end{pmatrix} = \begin{pmatrix} I & 0 \\ K(t) & I \end{pmatrix} \begin{pmatrix} \delta\mathbf{x}(t) \\ y(t) \end{pmatrix}$$

$$\begin{pmatrix} \delta\dot{\mathbf{x}}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{pmatrix} A^{11} + A^{12}K & A^{12} \\ \star & -KA^{12} + A^{22} \end{pmatrix} \begin{pmatrix} \delta\mathbf{x}(t) \\ y(t) \end{pmatrix} + \begin{pmatrix} g^1(t) \\ g^2(t) \end{pmatrix}$$

$\star = 0 \Rightarrow$ **Riccati equation**

$$\dot{K} = A^{21} - KA^{11} + A^{22}K - KA^{12}K, \qquad K(T) = \phi_{\mathbf{xx}}^T$$

$$\dot{y} = [-KA^{12} + A^{22}]y - Kq^1 + q^2, \qquad y(T) = 0$$

$\Rightarrow$ Decoupling can be realised using size$(\mathbf{u}) << $ size$(\mathbf{x})$

Explicit discretization

$$\min_{\mathbf{x},\mathbf{u}} \phi(\mathbf{x}_l),$$
$$\mathbf{x}_{i+1} = f_i(\mathbf{x}_i, \mathbf{u}_i), \quad i = 0, ..., l-1, \quad \mathbf{x}_0 \text{ fixed},$$

- Apply discrete Riccati approach
- Integration of discrete Riccati equation too expensive

Remedy: change the order of elimination

$$\delta\mathbf{u} = \delta\mathbf{u}(\delta\mathbf{x}, \delta\bar{\mathbf{x}})$$

and substitution

$$\begin{pmatrix} \delta\mathbf{x}_i \\ \delta\bar{\mathbf{x}}_i \end{pmatrix} = \begin{pmatrix} I & 0 \\ K_i & I \end{pmatrix} \begin{pmatrix} \delta\mathbf{x}_i \\ y_i \end{pmatrix}.$$

$\Rightarrow$ Pantoja approach [Pantoja]
$\Rightarrow$ AD implementation [Christianson]

Problem: Implicit discretisation is required
$\Rightarrow$ Straightforward extension of Pantoja algorithm

Institute of Scientific Computing

Implicit discretization

$$\min_{\mathbf{x},\mathbf{u}} \phi(\mathbf{x}_l),$$
$$F_i(\mathbf{x}_i, \mathbf{x}_{i+1}, \mathbf{u}_i) = 0, \quad i = 0, \ldots, l-1, \quad \mathbf{x}_0 \text{ fixed},$$

Unknown expression: $\mathbf{x}_{i+1} = f_i(\mathbf{x}_i, \mathbf{u}_i)$
Find

$$f_{\mathbf{x},i} = \frac{\partial \mathbf{x}_{i+1}}{\partial \mathbf{x}_i} \quad \text{and} \quad f_{\mathbf{u},i} = \frac{\partial \mathbf{x}_{i+1}}{\partial \mathbf{u}_i}$$

using Implicit Function Theorem:

$$\frac{\partial \mathbf{x}_{i+1}}{\partial \mathbf{x}_i} = -\left(\frac{\partial F_i}{\partial \mathbf{x}_{i+1}}\right)^{-1} \frac{\partial F_i}{\partial \mathbf{x}_i} = H_{\bar{\mathbf{x}}\mathbf{x}}^\top,$$

$$\frac{\partial \mathbf{x}_{i+1}}{\partial \mathbf{u}_i} = -\left(\frac{\partial F_i}{\partial \mathbf{x}_{i+1}}\right)^{-1} \frac{\partial F_i}{\partial \mathbf{u}_i} = H_{\bar{\mathbf{x}}\mathbf{u}}^\top.$$

$\Rightarrow$ Straightforward extension of Pantoja algorithm

2. Multiple sweep algorithm

Input: $\mathbf{u}_i$ for $0 < i < l$, $\mathbf{x}_0$

$\Rightarrow$ **Original sweep:** $i : 0 \rightarrow l-1$

Solve $F_i(\mathbf{x}_i, \mathbf{x}_{i+1}, \mathbf{u}_i) = 0$ for $\mathbf{x}_{i+1}$

Adjoint initialisation:

$\bar{\mathbf{x}}_l = \phi_{\mathbf{x}}^\top(\mathbf{x}_l)$, $K_l = \phi_{\mathbf{xx}}^\top(\mathbf{x}_l)$ and $y_l = 0$

$\Leftarrow$ **Adjoint sweep:** $i : l-1 \rightarrow 0$

$$\bar{\mathbf{x}}_i = \left(\frac{\partial \mathbf{x}_{i+1}}{\partial \mathbf{x}_i}\right)^\top \bar{\mathbf{x}}_{i+1}$$
$$K_i = K(K_{i+1}, \mathbf{x}_i, \bar{\mathbf{x}}_{i+1}),\ y_i = y(y_{i+1}, \mathbf{x}_i, \bar{\mathbf{x}}_{i+1}, K_{i+1})$$

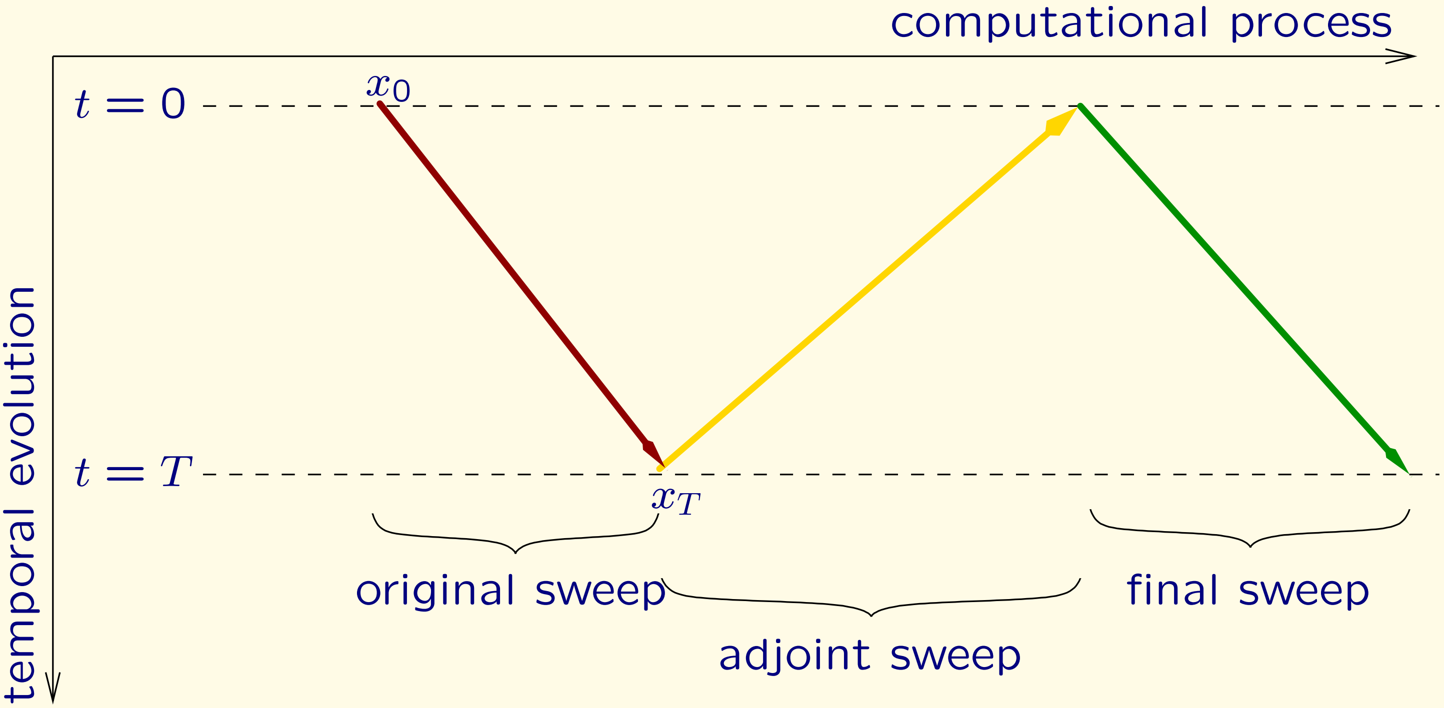
Final initialisation:

$\delta \mathbf{x}_0 = 0$

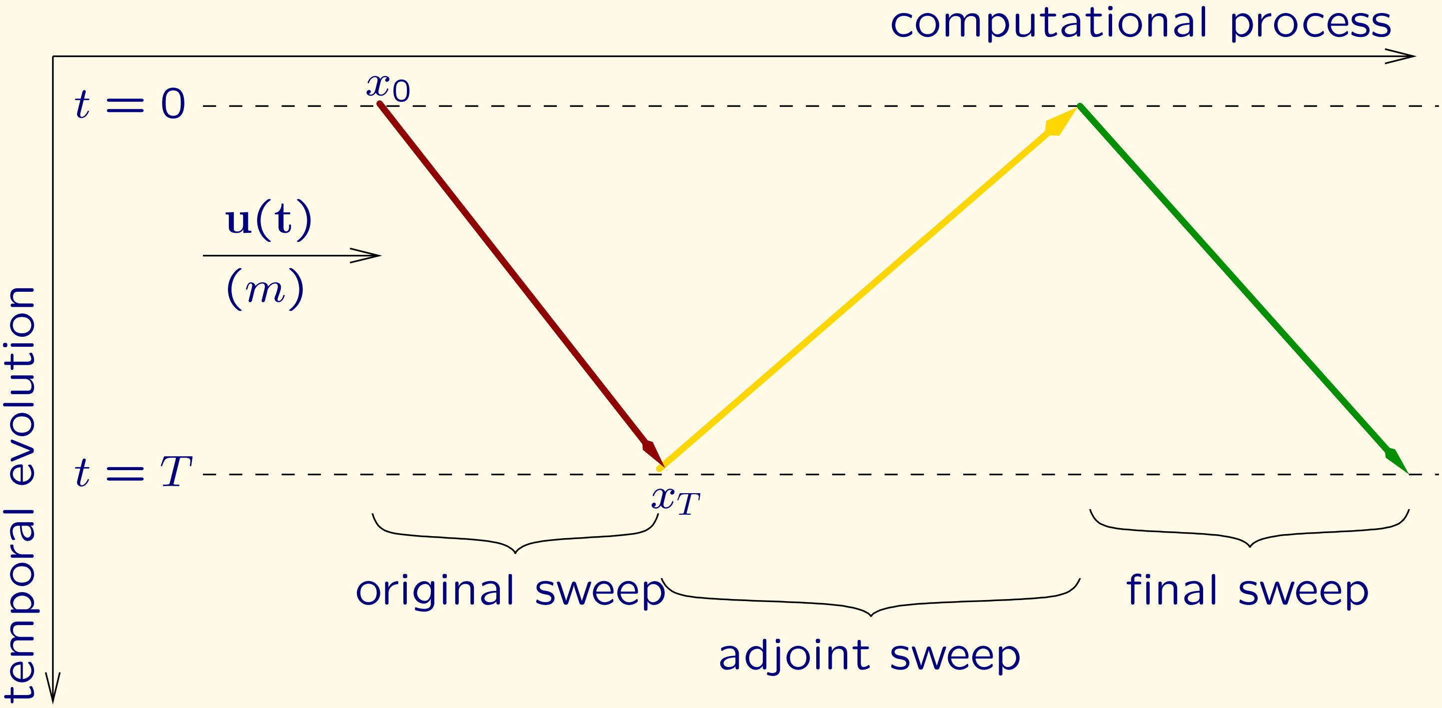
$\Rightarrow$ **Final sweep:** $i : 0 \rightarrow l-1$

$$\delta \mathbf{u}_i = \delta \mathbf{u}(\delta \mathbf{x}_i, \delta \bar{\mathbf{x}}_{i+1}, \mathbf{x}_i, \bar{\mathbf{x}}_{i+1}),\ \delta \mathbf{x}_{i+1} = \frac{\partial \mathbf{x}_{i+1}}{\partial \mathbf{x}_i}\delta \mathbf{x}_i + \frac{\partial \mathbf{x}_{i+1}}{\partial \mathbf{u}_i}\delta \mathbf{u}_i$$
$$\mathbf{u}_i^{j+1} = \mathbf{u}_i + \delta \mathbf{u}_i$$

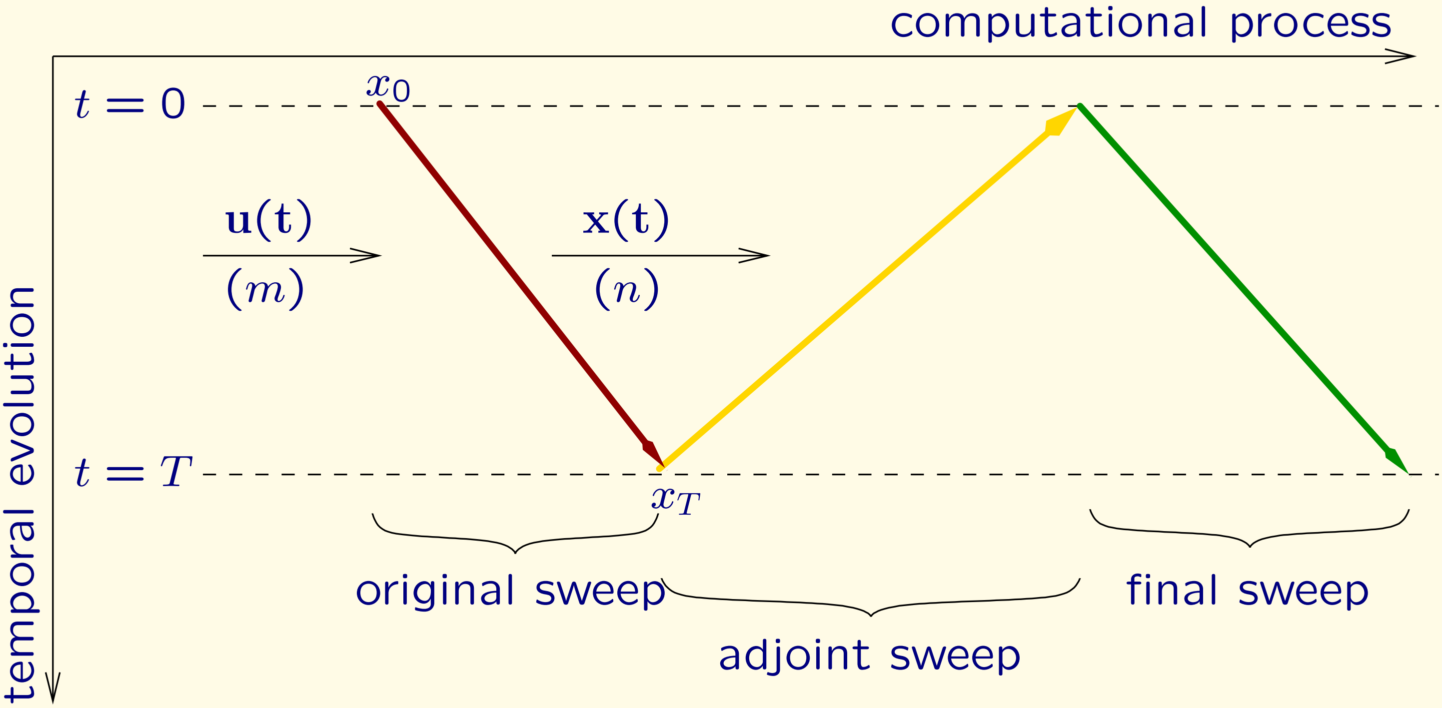
Information flow by Riccati/Pantoja computation of Newton step



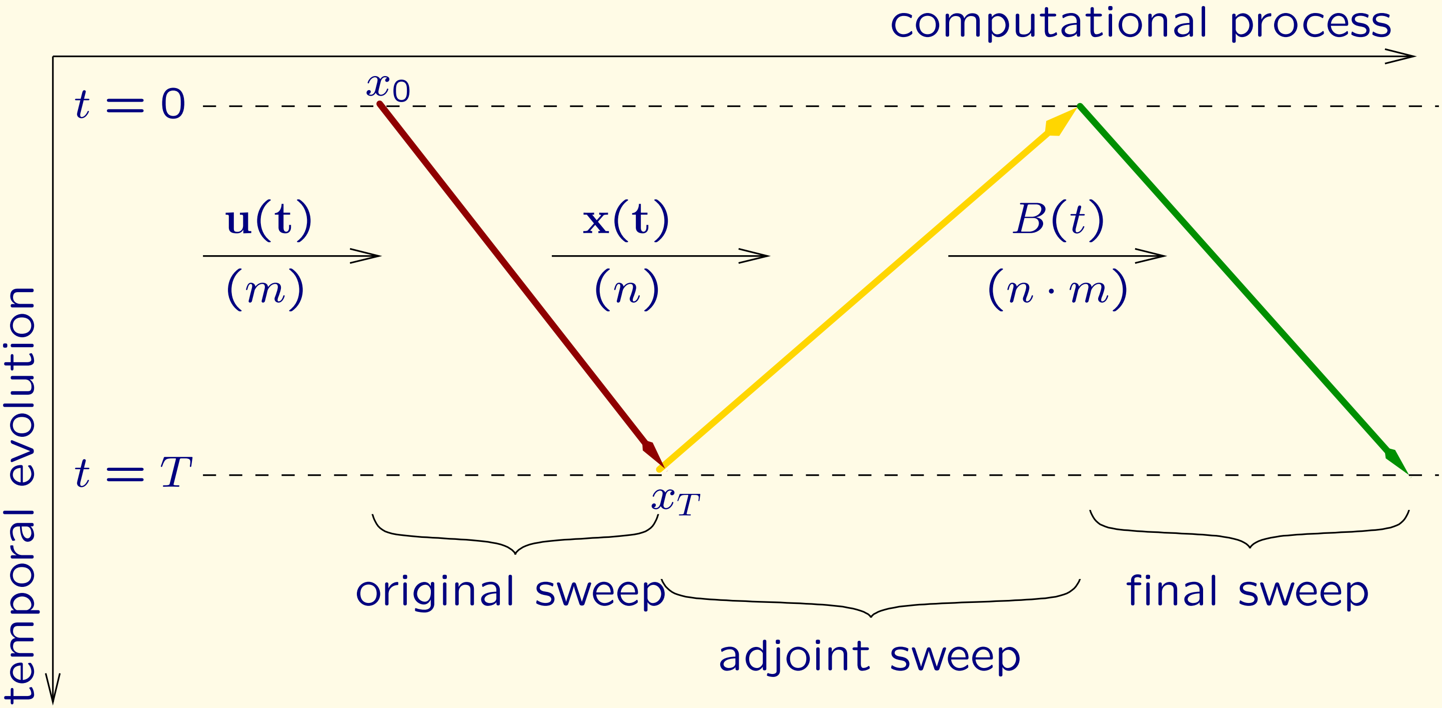
Information flow by Riccati/Pantoja computation of Newton step



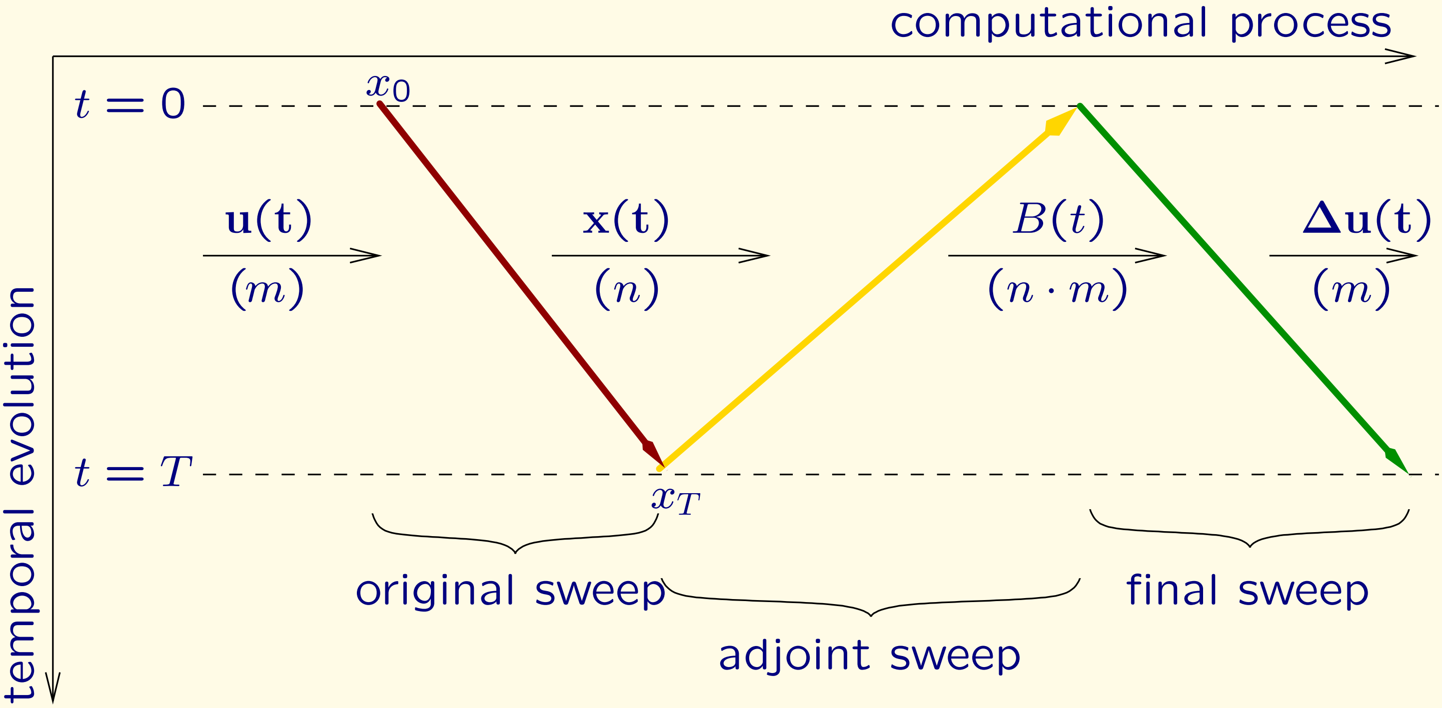
Information flow by Riccati/Pantoja computation of Newton step



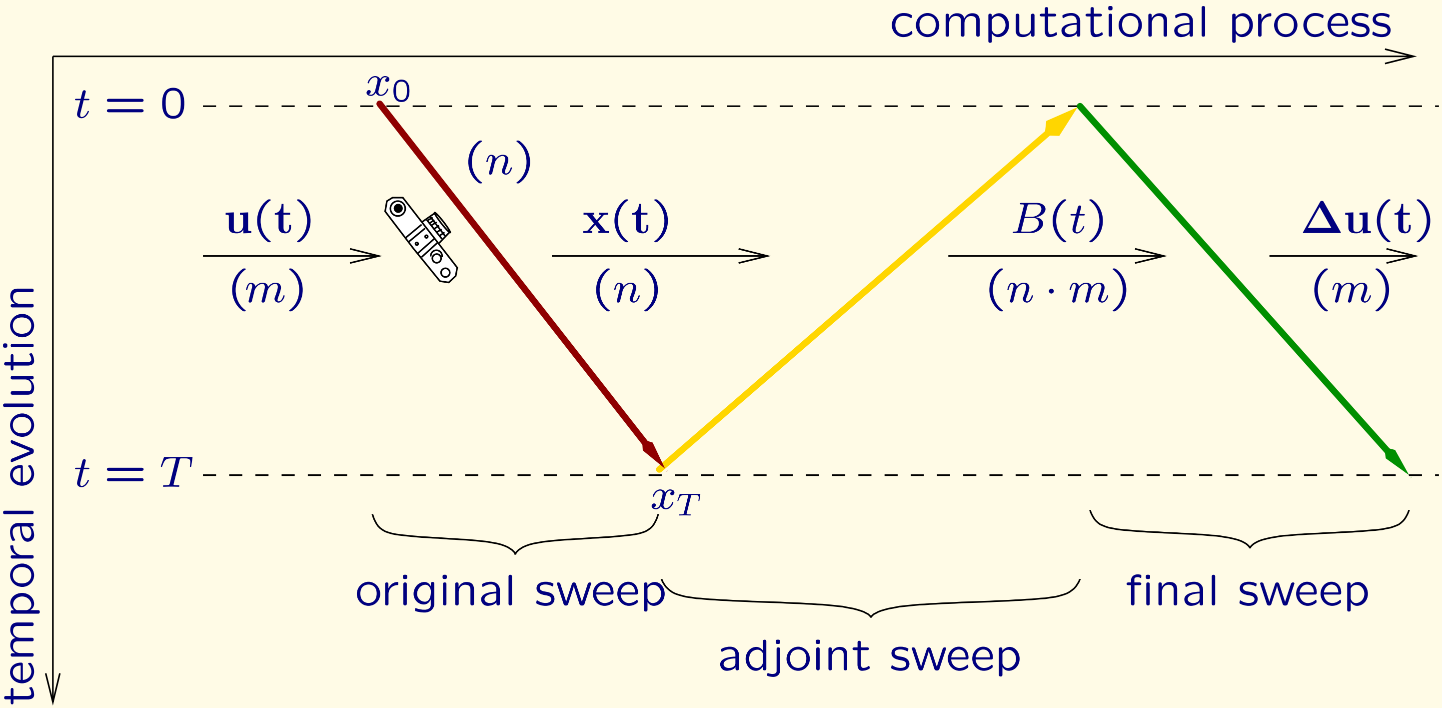
Information flow by Riccati/Pantoja computation of Newton step



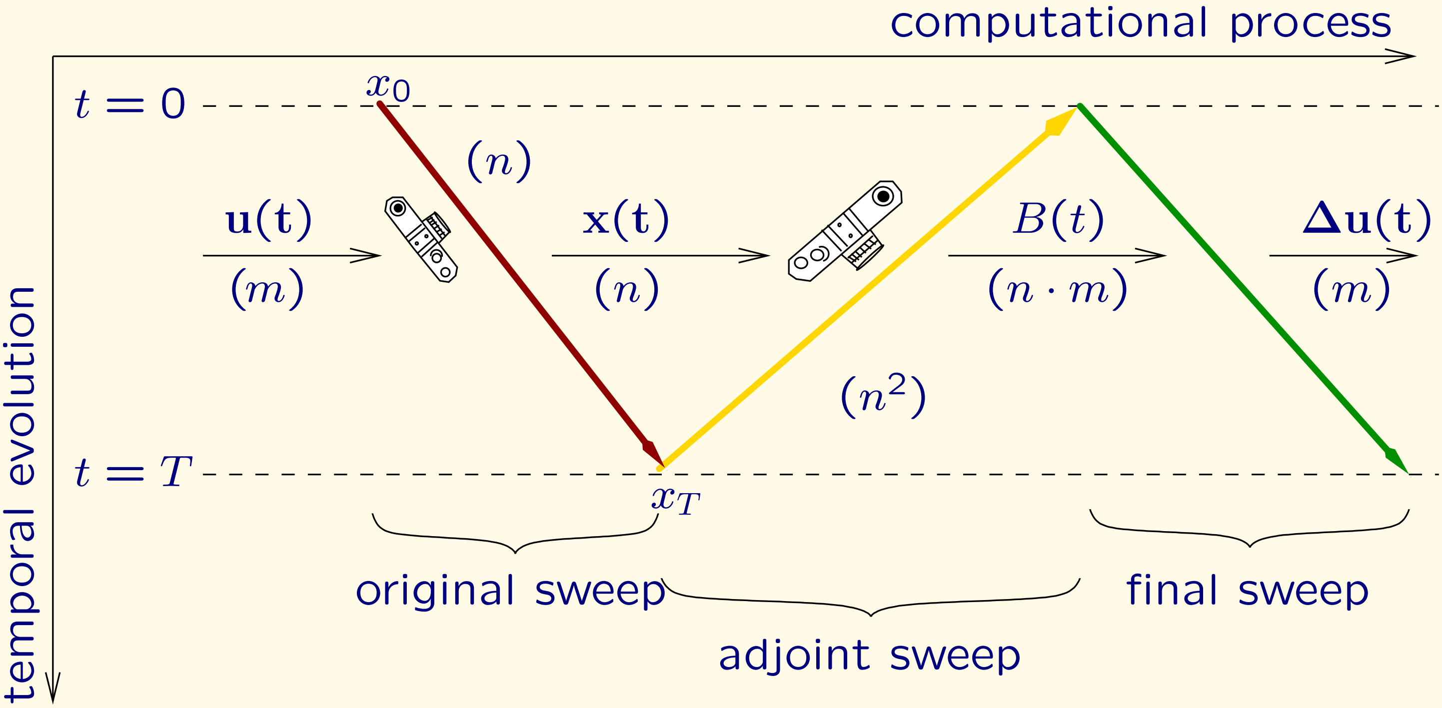
Information flow by Riccati/Pantoja computation of Newton step



Information flow by Riccati/Pantoja computation of Newton step

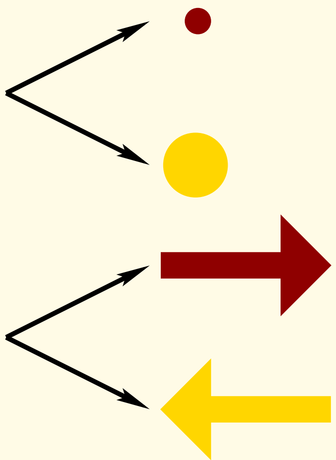


Information flow by Riccati/Pantoja computation of Newton step



3. Nested checkpointing

checkpoints



thin $d = \mathcal{O}(n)$

fat $\bar{d} = \mathcal{O}\left(n^2\right)$

step cost

original $t = \mathcal{O}\left(n^2\right)$

adjoint $\bar{t} = \mathcal{O}\left(n^3\right)$

Available memory is given as number C of fat checkpoints

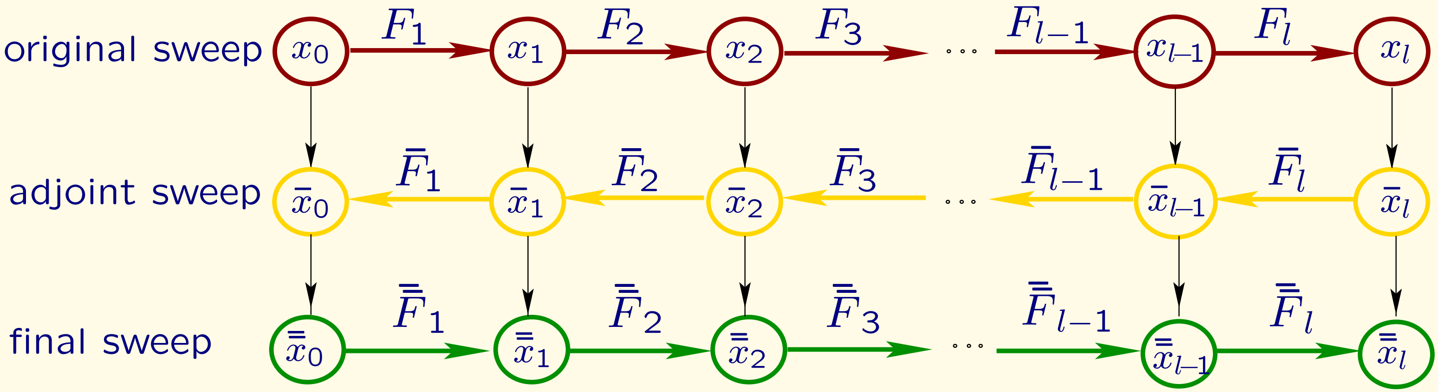
$$\#thin = c \times \#fat = c \times C$$

Find an optimal nested reversal by minimizing

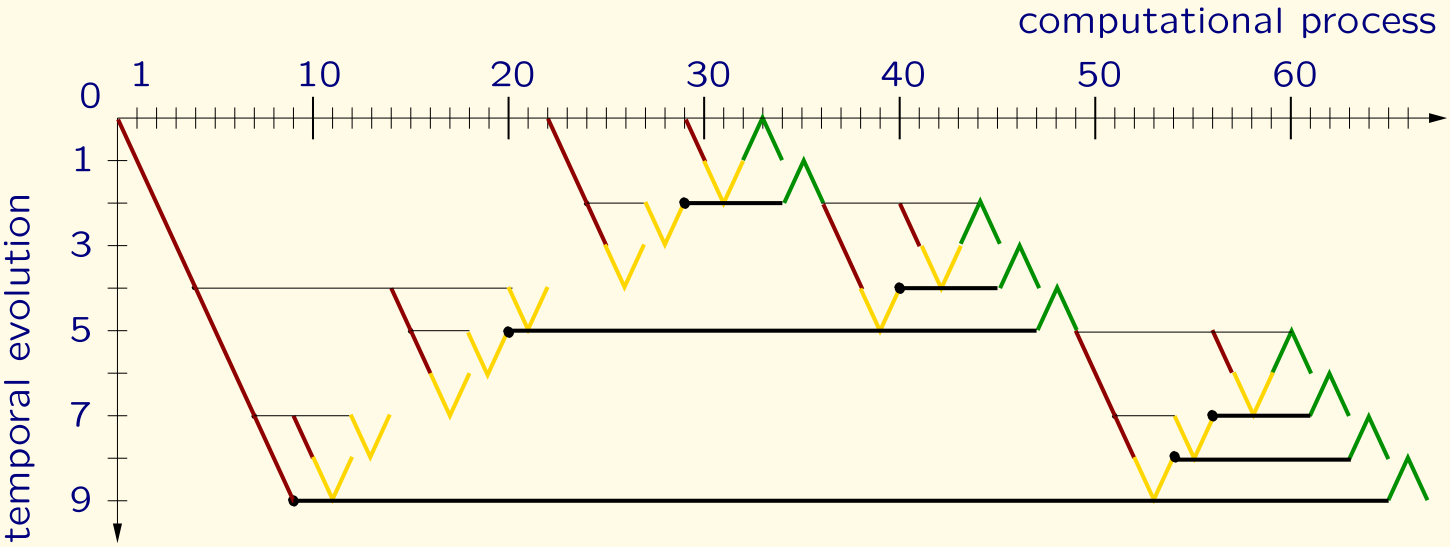
$$\mathbf{T}_{min}(l, C, c, t, \bar{t})$$

over all possible checkpoint distributions

Multiple sweep evolution $\mathcal{E}_{3\times l}$



Example of optimal nested reversal $\mathbf{S}_{min}(9,2)$ with $c = 1$, $t = 1$, and $\bar{t} = 5$

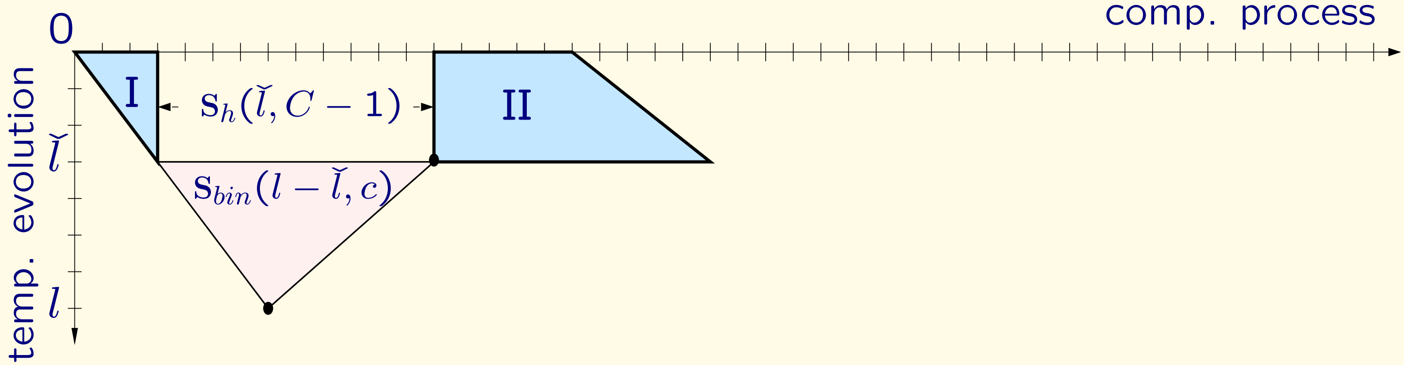


$\Rightarrow$ 23 original steps, 13 adjoint steps, 9 final steps

Institute of Scientific Computing

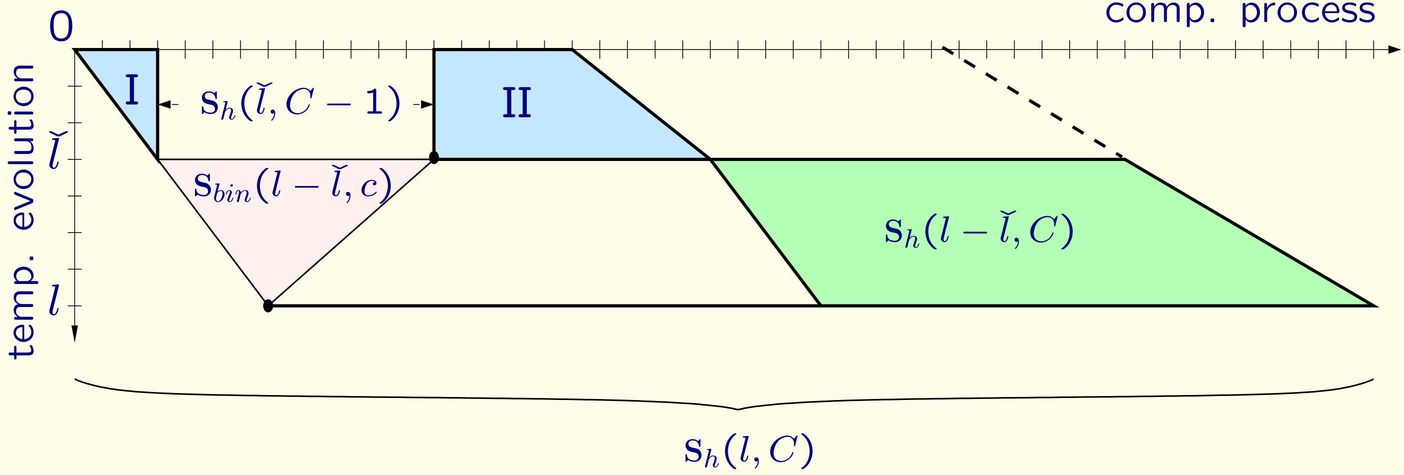
Heuristic

Decomposition of nested reversal schedule $\mathbf{S}_h(l, C)$



Heuristic

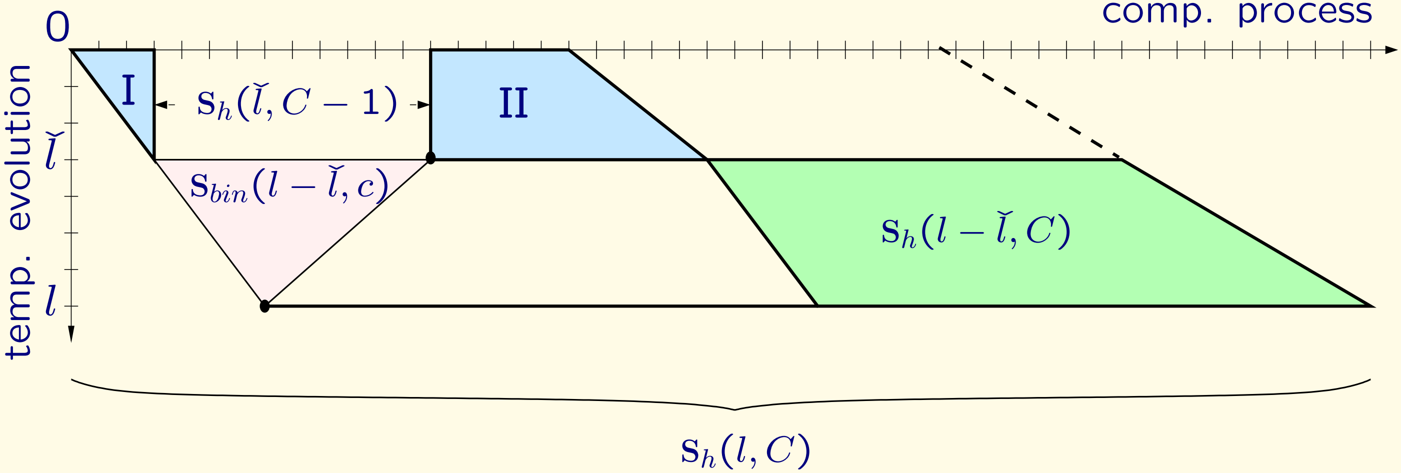
Decomposition of nested reversal schedule $\mathbf{S}_h(l, C)$



Institute of Scientific Computing

Heuristic

Decomposition of nested reversal schedule $\mathbf{S}_h(l, C)$



Find $\check{l}$ by minimizing

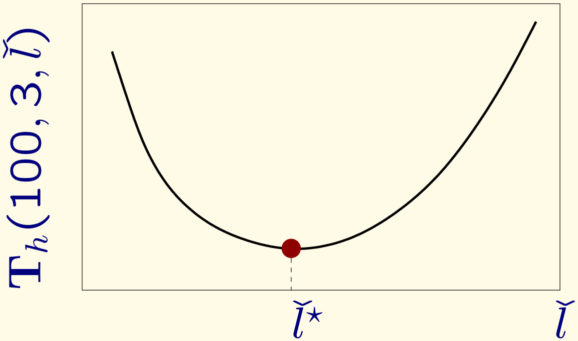
$$\mathbf{T}_h(l, C) = \min_{0<\check{l}<l} \mathbf{T}_h(l, C, \check{l})$$

Use **Linear method**

	Exhaustive search	Linear method
Memory	$\approx \mathcal{O}(l^C)$	$\mathcal{O}(l \times C)$
Run-Time	$\approx \mathcal{O}(l^3)$	$\mathcal{O}(l \times C)$

Linear method

$\mathbf{T}_h(l, C, \check{l})$ convex in dependence of $\check{l}$



$$\mathbf{T}_h(l, C, \check{l}) \searrow \; for\, \check{l} \le \check{l}^\star$$

$$\mathbf{T}_h(l, C, \check{l}) \nearrow \; for\, \check{l} > \check{l}^\star$$

- Find where $\mathbf{T}_h(l, C, \check{l}) - \mathbf{T}_h(l, C, \check{l} - 1)$ changes its sign first time

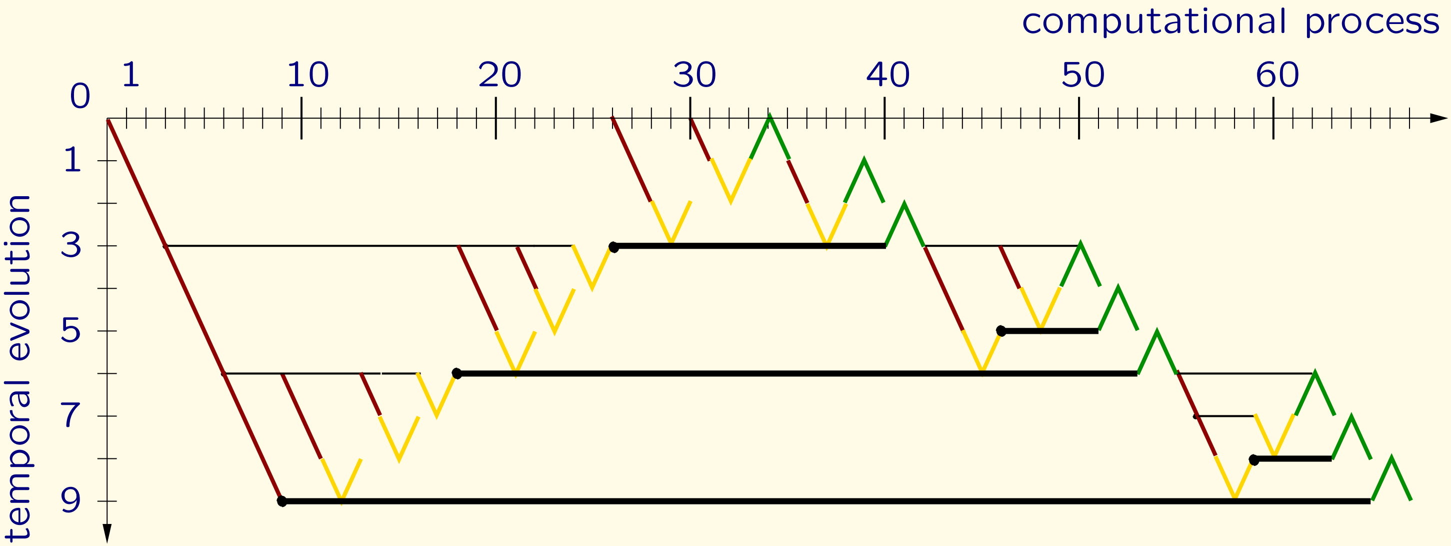
- Find $\mathbf{T}_h(l, C)$ using recursion over l and C

Use:

$$\check{l}^\star = \check{m}^\star \vee \check{l}^\star = \check{m}^\star + 1, \qquad \text{with } m = l - 1$$

$$\mathbf{F}_h(l, C) \equiv \mathbf{T}_h(l, C) - \mathbf{T}_h(l - 1, C)$$

Example of nested reversal $\mathbf{S}_h(9,2)$ with $c = 1$, $t = 1$, and $\bar{t} = 5$

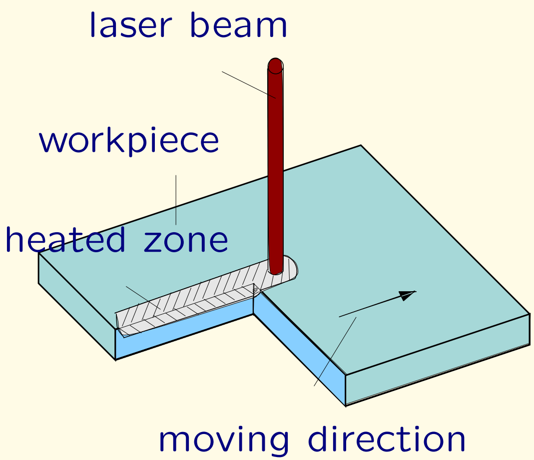


$\Rightarrow$ 24 original steps, 13 adjoint steps, 9 final steps

4. Numerical results

Laser surface hardening [Hömberg, Volkwein]

$Q = \Omega \times (0,T)$, $\Sigma = \partial\Omega \times (0,T)$, $\Omega = [0,5] \times [-1,0]$, $T = 5.25$



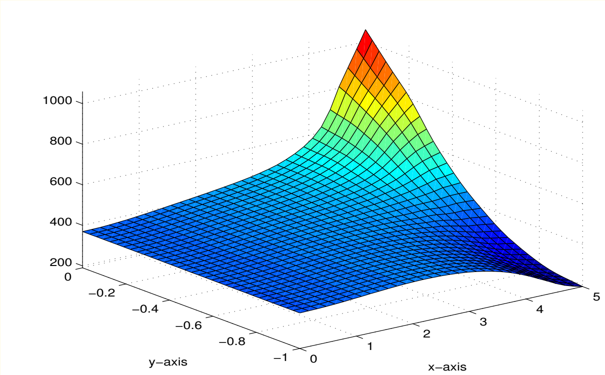
State equations

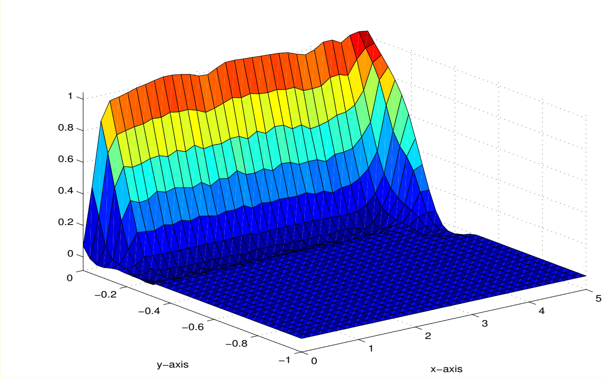
$$
\begin{aligned}
a_t &= \frac{1}{\tau(\theta)}[a_{eq}(\theta) - a]_+, \text{ in } Q \\
a(0) &= 0, \text{ in } \Omega \\
\rho c_\rho \theta_t - k\triangle\theta &= -\rho L a_t + u\alpha, \text{ in } Q \\
\frac{\partial\theta}{\partial\nu} &= 0, \text{ on } \Sigma \\
\theta(0) &= \theta_0, \text{ in } \Omega
\end{aligned}
$$

Cost functional

$$J(a,\theta,u) = \frac{\beta_1}{2}\int_\Omega (a(x,T) - a_d(x))^2 dx + \frac{\beta_2}{2}\int_0^T \int_\Omega [\theta - \theta_m]_+^2 dx\, dt + \frac{\beta_3}{2}\int_0^T u^2 dt$$

FE snapshot for temperature and volume fraction of austenite at $t = T$ $(n = 147)$





Desired state a_d

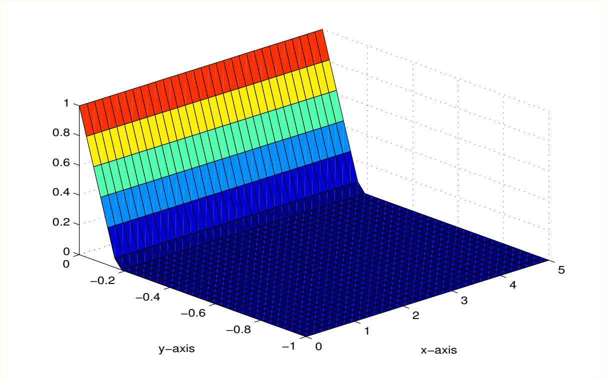
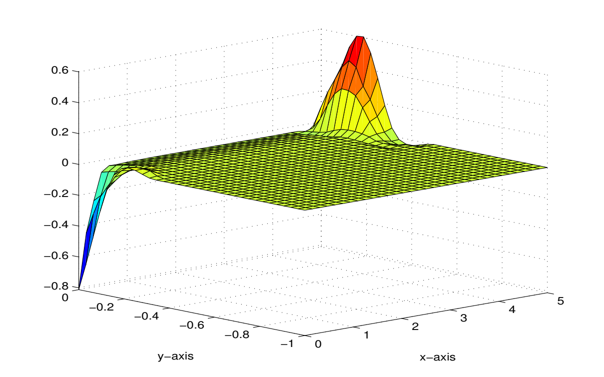


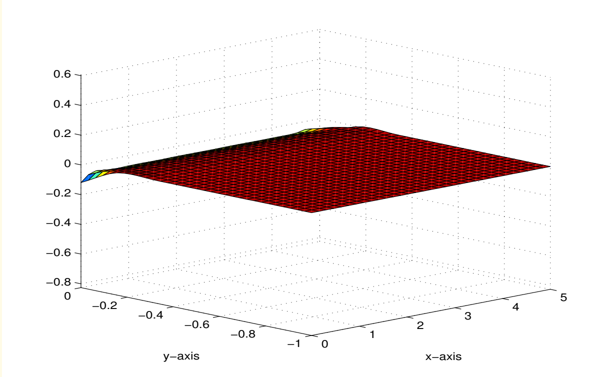
Table 1: Program data

i	λ^i	$J(\mathbf{u}^i)$	$\|\nabla J(\mathbf{u}^i)\|_2$	$\nabla J(\mathbf{u}^i)\delta\mathbf{u}^i$	$\|a^i(T) - a_d\|_{L^2(\Omega)}$	$\|\theta^i\|_{L^\infty(Q)}$	$\Lambda(i) = \frac{\|\delta\mathbf{u}^i\|}{\|\delta\mathbf{u}^i+\mathbf{u}^i\|}$
0	0.000000	670.096247	2.197499	−0.000000	0.158144	1148.594458	0.000000
1	0.001130	665.493535	2.198253	−4275.090081	0.157792	1148.527231	0.019460
2	0.005000	650.324921	2.194708	−3257.726444	0.156671	1148.370625	0.032751
3	0.000034	647.678880	2.194745	−82044.859253	0.156744	1148.433767	0.014622
4	0.000001	647.678880	2.166298	782.845599	0.155249	1147.414625	0.000561
5	0.000000	647.678880	2.136994	2524.100850	0.153778	1146.409009	0.000554
6	0.100000	589.317913	1.864956	−543.282251	0.140428	1141.066647	0.088242
7	0.134720	576.179109	1.824791	−409.596505	0.144286	1134.496470	0.088891
8	0.100000	534.398496	1.378520	−668.869910	0.123036	1201.254915	0.236620
9	0.100000	522.596740	1.229784	−209.076937	0.119931	1185.524519	0.094729
10	0.100000	507.601359	1.066479	−192.808499	0.114103	1178.850090	0.130864
11	1.000000	458.536134	0.463360	−212.394660	0.069952	1219.525080	0.391770
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
30	1.000000	355.305487	0.019172	−2.436999	0.045926	987.141029	0.052059
31	1.000000	355.106094	0.004853	−0.396550	0.045660	986.614861	0.019979
32	1.000000	355.047319	0.001508	−0.103916	0.045595	986.605644	0.008947
33	1.000000	355.043523	0.000163	−0.008023	0.045628	986.605683	0.003201
34	1.000000	355.043510	0.000006	−0.000185	0.045634	986.605683	0.000582
35	1.000000	355.043510	0.000000	−0.000000	0.045634	986.605683	0.000000

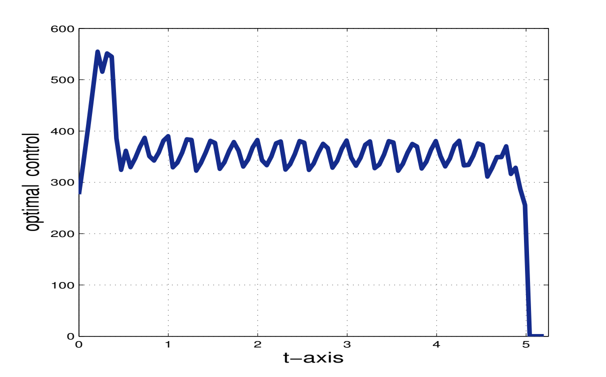
Institute of Scientific Computing

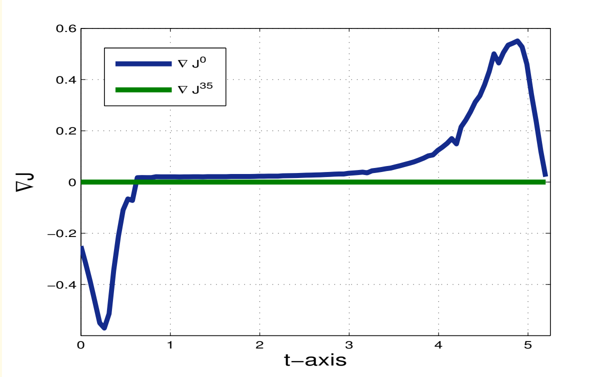
Differences $a^0(T) - a_d$ and $a^{35}(T) - a_d$ for the initial iterate a^0 and for the optimal state a^{35}



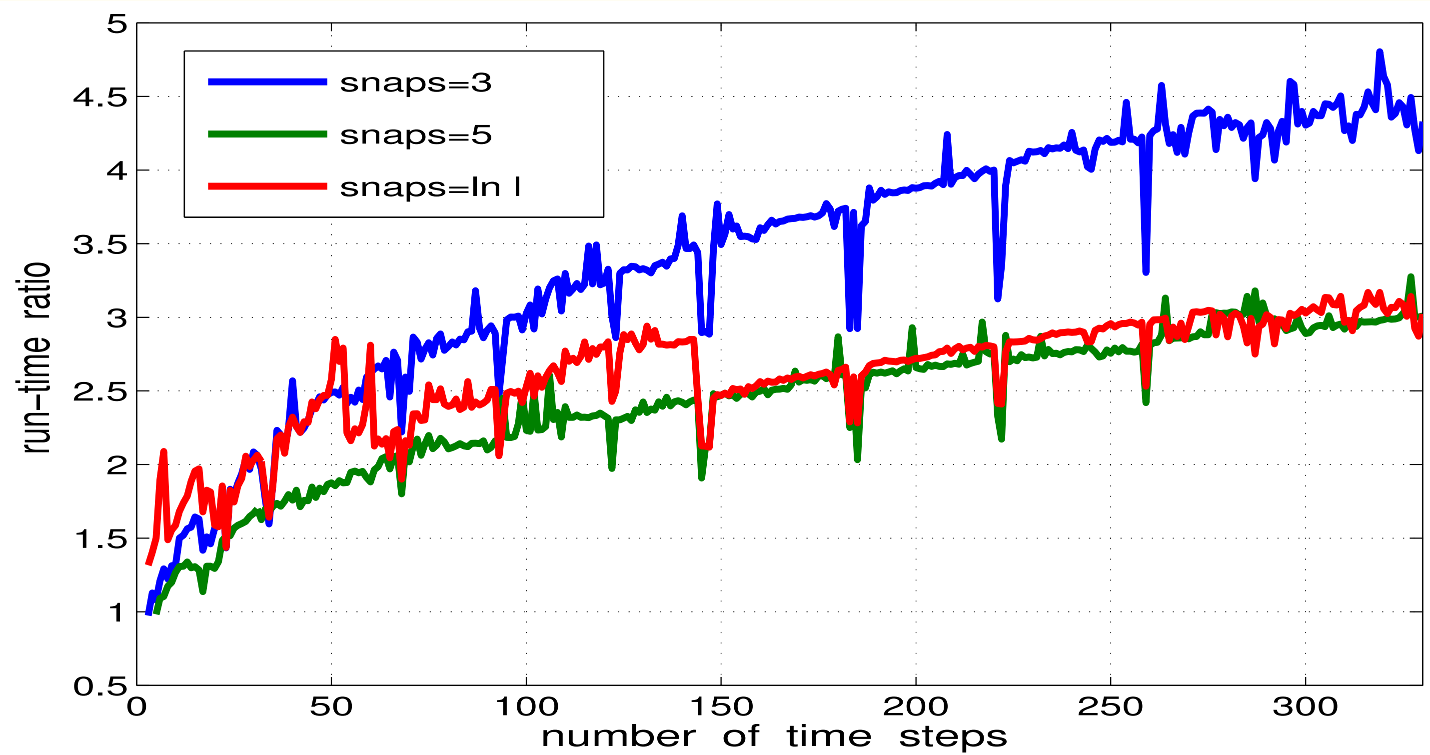


Optimal control $\mathbf{u}^{35}$ and the gradients ∇J^0 and ∇J^{35} of the cost function





Run-time & memory trade-off



5. Summary and conclusion

- Optimal control problems:

  Multiple shooting method

  - Requires expression for $\mathbf{u}(t) = \mathbf{u}(\mathbf{x}(t), \bar{\mathbf{x}}(t))$
  - Requires estimation for $\bar{\mathbf{x}}(0)$

  Newton/Riccati/Pantoja method

  Requires triple sweep

  Enormous memory requirement for its realization

- Nested checkpointing:

  Optimal strategies & heuristic

  Drastic memory reduction

  Run-time $\approx \mathcal{O}\left(l \ln^2 l\right)$