Towards a Fair Comparison between the Nested and Simultaneous Control Co-Design Methods using an Active Suspension Case Study

Athul K. Sundarrajan\(^1\) \hspace{1cm} Daniel R. Herber\(^2\)

\(^1\)Graduate Student
\hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} Colorado State University, Department of Systems Engineering
\hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} athul.sundarrajan@colostate.edu

\(^2\)Assistant Professor, \textbf{Presenting Author}
\hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} Colorado State University, Department of Systems Engineering
\hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} daniel.herber@colostate.edu
Introduction
Control co-design (CCD) is a class of integrated design methods that **concurrently treat the dynamic system’s physical and control aspects**, overcoming some of the limitations of traditional sequential approaches\(^1\).

Various solution strategies for CCD have been studied, but only relatively few studies have made detailed comparisons.

Such studies could provide the needed early insights into the appropriate choice of strategy and implementation techniques.

The goal of this work is to make a contribution towards this end for a subset of popular CCD coordination and solution strategies on a complex CCD problem.

While guidelines applicable to broad classes of CCD problems will not be made, the existing literature-based comparisons and a detailed case study will better **demonstrate the state-of-the-art understanding and how a thorough investigation can be conducted** to yield the desired implementation insights.

---

\(^1\) Allison and Herber 2014; Fathy et al. 2001
The nested CCD formulation is a reorganization of the simultaneous CCD problem as a two-level optimization problem with an outer and inner loop hierarchy.
Comparing the Strategies
The Coordination Selection Decision

- In many studies, there is a brief statement on the coordination selection
  - In References\(^1\), the authors state that the nested approach is better suited for their problem
  - In References\(^2\), the authors assert the simultaneous approach is the better choice
- The supporting evidence is often omitted from the final published work (and it should not be compulsory that every study fully justify their decision or even try all the approaches)
- However, we still feel that there is a lack of clear, comprehensive case studies that can be used to be better inform implementation decisions
- In Reference\(^3\), the CCD problem (and the one considered in this article) was solved using the simultaneous approach, but the nested implementation in that work was incomplete not fully leveraging the problem’s structure

\(^1\) Deshmukh and Allison 2015; Chilan et al. 2017; Belvin and Park 1990; Onoda and Haftka 1987
\(^2\) Allison, Guo, and Han 2014; Azad et al. 2019; Hu et al. 2016
\(^3\) Allison, Guo, and Han 2014
Towards a Fair Comparison

- Several studies have shown quantitative comparisons between the two strategies on the same CCD problem\(^1\).

- In the paper, we provided a detailed list of potential trade-offs between the coordination strategies based on current CCD and MDO literature.

- If ineffective implementations of some parts of a coordination/solution method are used, then the comparison may be biased.
  - For example, in Ref. Herber and Allison 2019b, no “fair” comparisons were made because the simultaneous implementation did not leverage the problem’s sparsity.

- Often, function evaluations are the main computational comparison, but this is quite limited in the case of diverse optimization architectures (e.g., what about the other costs associated with solving the problem?)

- Runtime benchmarking does try to normalize the utility of the two strategies where lowered measured runtime cost (for the same outcome) is desirable.

- Solution quality is another dimension that could be argued (how accurate is the resulting solution, other factors?)

- Therefore, we should also be considering solution quality vs. computational expense.

\(^1\) Deshmukh and Allison 2015; Clarizia 2019; Reyer and Papalambros 2002
Active Suspension CCD Case Study
Active Suspension Case Study

- **Active vehicle suspension CCD problem in this work is the one from Allison, Guo, and Han 2014**

- The system consists of two masses: sprung mass $m_{s/4}$ and unsprung mass $m_{us/4}$

- The suspension is composed of a spring $k_s$ and damper $c_s$, and a force actuator $u(t)$

- $k_t$ and $c_t$ are the spring damper constants of the tire, and $z_0(t)$ is the road input

- There are seven design geometric plant design variables associated with the spring and damper
→ System Dynamics and Objective

- There are four states in the system \((z_s - z_{us}, \dot{z}_s, z_{us} - z_0, \dot{z}_{us})\)
- The dynamics of the system are linear with respect to \((\xi(t), u(t))\), and nonlinear with respect to \(x_p\):

\[
\dot{\xi}(t) = A(x_p)\xi(t) + Bu(t) + E\dot{z}_0(t) \tag{6a}
\]

\[
A = \begin{bmatrix}
0 & \frac{1}{m_{us}/4} & 0 & 0 \\
-k_t(x_p) & 0 & -[c_s(x_p) + c_t] & \frac{k_s(x_p)}{m_{us}/4} \\
\frac{m_{us}/4}{m_s/4} & 0 & 1 & \frac{c_s(x_p)}{m_{us}/4} \\
0 & \frac{-k_s(x_p)}{m_s/4} & -c_s(x_p) & \frac{1}{m_{us}/4}
\end{bmatrix},
B = \begin{bmatrix}
0 \\
-1 \\
0 \\
0
\end{bmatrix},
E = \begin{bmatrix}
-1 \\
c_t \\
0 \\
0
\end{bmatrix} \tag{6b-c}
\]

- The objective function is a combination of quadratic penalties on handling \((z_{us} - z_0)\), passenger comfort \(\ddot{z}_s\), and control effort \(u\):

\[
o = \int_{t_0}^{t_f} \left[ w_1 \xi_1^2 + w_2 [\dot{\xi}_4(t, \xi, u, x_p)]^2 + w_3 u^2 \right] dt \tag{7}
\]

with \(w_1 = 10^5, w_2 = 0.5, \text{ and } w_3 = 10^{-5}\) from Ref. Allison, Guo, and Han 2014

- Two design load cases \(z_0\) are simultaneously considered with a weighted sum: 1) ramp profile, 2) rough road profile

\[
\min_{\xi, u, x_p} 10^{-2} o(\xi_{\text{ramp}}, u_{\text{ramp}}, x_p) + o(\xi_{\text{rough}}, u_{\text{rough}}, x_p) \tag{8}
\]
Plant Design: Spring

- The spring physical design variables are the wire diameter $d$, helix diameter $D$, pitch $p$, and number of active coils $N_a$.
- The main intermediate parameter is the spring constant $k_s(x_p)$:
  \[
  k_s = \frac{d^4 G}{8D^3 N_a \left[1 + \frac{d^2}{2D^2}\right]}
  \]
- There are six static constraints and four dynamic constraints:

  \begin{align*}
  g_{o,1}(x_p) &= 4 - C \leq 0 \quad (10) \\
  g_{o,2}(x_p) &= C - 12 \leq 0 \quad (11) \\
  g_{o,3}(x_p) &= L_0 - 5.26D \leq 0 \quad (12) \\
  g_{o,4}(x_p) &= L_0 - 0.40 \leq 0 \quad (13) \\
  g_{o,5}(x_p) &= d + D - 0.25 \leq 0 \quad (14) \\
  g_{o,6}(x_p) &= 1.2\tau(F_s) - S_{sy} \leq 0 \quad (16) \\
  g_{i,1}(x_p, \xi) &= \max_t |\xi_3(t)| - L_0 + L_s + 0.02 + \delta_g \leq 0 \quad (17) \\
  g_{i,2}(x_p, \xi) &= 0.15 + 1 - \frac{L_0 - L_s}{\delta_g + 1.1\xi_3(t)} \leq 0 \quad (18) \\
  g_{i,3}(x_p, \xi) &= \frac{1.2\tau(F_a)}{0.24S_{ut}} + \frac{\tau(F_m)}{S_{sy}} - 1 \leq 0 \quad (19) \\
  g_{i,4}(x_p, \xi) &= \frac{1.2\tau(F_a)}{241 \times 10^6} - 1 \leq 0 \quad (20)
  \end{align*}
Plant Design: Damper

- The damper physical design variables are the valve diameter $D_o$, working piston diameter $D_p$, and damper stroke $D_s$.
- The intermediate parameter is the damper constant $c_s(x_p)$:

$$c_s = \frac{D_p^4}{8C_dC_2(D_o)D_o^2}\sqrt{\frac{\pi k_v \rho_1}{2}}$$

- There are three static and dynamic constraints:

$$g_{o,7}(x_p) = d - D + D_p + 0.022 \leq 0 \quad (22)$$
$$g_{o,8}(x_p) = 2D_s - 0.394 \leq 0 \quad (23)$$
$$g_{o,9}(x_p) = L_0 - L_s - D_s \leq 0 \quad (24)$$

$$g_{i,5}(x_p, \xi) = \frac{4c_s(D_o) \max_t |\dot{\xi}_3(t)|}{\pi D_p^2} - 4.75 \times 10^6 \leq 0 \quad (25)$$
$$g_{i,6}(x_p, \xi) = \max_t |\dot{\xi}_3(t)| - 5 \leq 0 \quad (26)$$
$$g_{i,7}(x_p, \xi) = \frac{4\pi D_o^2 c_s(D_o) \max_t |\dot{\xi}_3(t)|}{4k_v \pi D_p^2} - 0.03 \leq 0 \quad (27)$$
Results
A single complete and accurate solution to the CCD problem is presented first with $n_t = 5000$ and tight optimality/feasibility tolerances.

These results obtained using both the simultaneous and nested CCD strategies (code available on GitHub\(^1\))

More than $10 \times$ reduced computational cost with more time points (5000 here vs. 275) than the code provided by Allison, Guo, and Han 2014

\(^1\) https://github.com/danielrherber/dt-qp-project/tree/master/examples/ACC2021
• To explore the variety of implementation decisions and their trade-offs, a variety of options were tested on this CCD problem (please see the paper and code for full details)
• Simultaneous implementation (S):
  • Different variations were tested for different permutations of the number of discretization points $n_t = [200, 600, 2000]$, optimality tolerance $[10^{-3}, 10^{-5}, 10^{-7}]$, feasibility tolerance $[10^{-4}, 10^{-8}, 10^{-12}]$, and derivative methods (real-forward FD, real-central FD, complex-step FD, and symbolic)
• Nested implementation (N):
  • Different permutations of number of discretization points $n_t = [200, 600, 2000]$, outer-loop optimality tolerances $[10^{-1}, 10^{-3}, 10^{-5}]$, inner-loop optimality and feasibility tolerances $[10^{-4}, 10^{-8}, 10^{-12}]$, and outer-loop derivative methods (real-central and real-forward FD)
The results indicate that some implementation of S is the superior approach, but N is quite competitive.

However, simply increasing the number of time points does not result in an efficient or even accurate solution.
Implementations with Various Derivative Methods

- Here we see **why S was superior on the previous slide; it was using symbolic derivatives**
- Accurate derivative information is shown to be a critical factor for efficient S implementations
- **If symbolic derivatives are unavailable, then N is superior in this problem**
- Also, note the performance benefits of the complex-step method over the other real-valued based derivative approximation methods
Conclusion

• It is imperative to understand the trade-offs and goals of a particular CCD study (e.g., desired accuracy) and use sound judgment to make informed decisions on how to solve a chosen CCD problem.

• The results of the case study showed that the simultaneous strategy using symbolic derivatives was generally superior while the nested strategy was preferable when any other derivative method was used.
  • More detailed discussion is provided in the paper.

• For other CCD problems, the observed trends might be different.

• Future work will seek to understand these trade-offs and best practices in a more diverse set of CCD problem scenarios moving towards a robust set of CCD implementation guidelines.
References and Extra Slides


References and Extra Slides (continued)


Towards a Fair Comparison between the Nested and Simultaneous Control Co-Design Methods using an Active Suspension Case Study

Athul K. Sundarrajan
Colorado State University
athul.sundarrajan@colostate.edu

Daniel R. Herber
Colorado State University
daniel.herber@colostate.edu
www.engr.colostate.edu/~drherber

https://github.com/danielrherber/dt-qp-project
~/examples/ACC2021
Control Co-design Problem Formulation

- In the **simultaneous CCD formulation**, a single fixed-time horizon optimization problem is put forth\(^1\), and the optimizer simultaneously analyzes and designs the plant, state, and control variables\(^2\):

$$\begin{align*}
\min_{x=[\xi, x_c, x_p]} & \quad o = \int_{t_0}^{t_f} \ell(t, \xi, x_c, x_p) \, dt + m(\xi_0, \xi_f, x_c, x_p) \\
\text{subject to:} & \quad \dot{\xi}(t) - f(t, \xi, x_c, x_p) = 0 \\
& \quad h = \begin{bmatrix} h_o(x_p) \\ h_i(t, \xi, x_c, x_p, \xi_0, \xi_f) \end{bmatrix} = 0 \\
& \quad g = \begin{bmatrix} g_o(x_p) \\ g_i(t, \xi, x_c, x_p, \xi_0, \xi_f) \end{bmatrix} \leq 0
\end{align*}$$

where: \( \xi_0 = \xi(t_0), \xi_f = \xi(t_f) \)

- \((\xi, x_c, x_p)\) are the collections of the selected states, control design variables, and plant design variables

\(^1\) For simplicity, a free-time horizon, multiple phases, general state differential equations, and differentiating between path and boundary constraints are omitted from the formulation. No uncertainties are directly considered at this time.  
\(^2\) Allison, Guo, and Han 2014; Reyer and Papalambros 2002
List of Advantages

- In the paper, we provided a detailed list of potential trade-offs between the coordination strategies based on current CCD and MDO literature
- For our complete list, please see the paper

Some advantages (+) of the simultaneous strategy:
- Can find the solution quickly by letting the optimizer explore regions that are infeasible
- Naturally handles bidirectional coupling between the plant and control design variables
- It supports fine-grained parallelization and advanced derivative methods

Some advantages (+) of the nested strategy:
- Each subproblem’s structure is simplified and the size is reduced from the original simultaneous formulation
- Tailored optimization algorithms and tolerances can be used in the different subproblems that can leverage the simplified problem structure (e.g., QP or LQR) or outer-loop global search
- Results from intermediate iterations are feasible

1 Martins and Lambe 2013  
2 Allison, Guo, and Han 2014; Herber and Allison 2019b  
3 Allison, Guo, and Han 2014  
4 Herber and Allison 2019b; Belvin and Park 1990; Onoda and Haftka 1987; Chilan et al. 2017; Herber and Allison 2019a
List of Disadvantages

- Some disadvantages (−) of the simultaneous strategy
  - Large problem size\(^1\)
  - Requires many function calls to potentially expensive problem elements such as \(f(\cdot)\)\(^2\)
  - Results from intermediate iterations are not guaranteed to be feasible\(^3\)

- Some disadvantages (−) of the nested strategy
  - Issues when the inner-loop problem is not defined at a particular \(x_p\)\(^4\)
  - Challenging to compute accurate outer-loop derivatives (e.g., accuracy of the subproblem impacts accuracy of outer-loop derivative information\(^5\))
  - Only supports coarse-grained parallelization\(^6\)

---

1. Martins and Lambe 2013
2. Deshmukh and Allison 2015; Chilan et al. 2017; Herber and Allison 2019a
3. Martins and Lambe 2013
4. Herber and Allison 2019b
5. Martins and Lambe 2013
6. Allison, Guo, and Han 2014
Here we formalize the analysis of a CCD problem to determine if the problem is an LQDO-amenable CCD problem:

- A problem where you have a linear-quadratic dynamic optimization problem for fixed plant design.

- The dependencies are categorized as linear, quadratic, and nonlinear, and a specialized qualifier is added for the case when the plant design is fixed (denoted $x_p^*$).

- It also facilitates the partitioning of the inner- and outer-loop constraints.

### Dependency Matrix for an LQDO-amenable CCD Problem

<table>
<thead>
<tr>
<th></th>
<th>$\xi$</th>
<th>$x_c$</th>
<th>$x_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$o(.)$</td>
<td>I</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>$f(.)$</td>
<td>III</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>$h_i(.)$</td>
<td>X</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>$g_i(.)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h_o(.)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$g_o(.)$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Legend:**
- Linear
- $x_p^*$ Linear
- Quadratic
- $x_p^*$ Quadratic
- Nonlinear

Increasing Complexity
Detailed Problem Formulation Analysis of the CCD Problem (continued)

- There are a variety of implementation insights
- All $g_i(\cdot)$ can be transformed into linear inequality constraints (some needing two)
- Inner-loop constraints $g_{i,1}(\cdot) - g_{i,4}(\cdot)$ have the same mathematical structure for a fixed plant design so only one can be active!
  - We can easily determine which one might be active before solving the inner-loop problem, removing 3/4 inequality constraints
- Similarly analysis on $g_{i,5}(\cdot) - g_{i,7}(\cdot)$
- $D_s$ doesn’t change the inner-loop problem
- Several linear $g_o(\cdot)$ are linear constraints
- All of these insights are factors that improve the solution implementation for both coordination strategies
Nested Implementation Tolerance Insights

**Outer-loop Optimality Tolerance**

- These only show N, highlighting the different inner- and outer-loop tolerances.
- Outer-loop tolerance study shows a clear trade-off between objective value accuracy and runtime.
- Loss in outer-loop accuracy is commonplace when the inner-loop tolerance is not small enough.
- Balancing these tolerances can permit efficient implementations that achieve the desired accuracy without extra computational cost.
A fundamental assumption for equivalence between the nested and simultaneous formulations is consistent inner-loop feasibility.

However, uniform random sampling of plant designs indicate 44% have infeasible inner loops, and approximately 0.033% are infeasible with respect to all constraints.

So the nested implementation must start from a point with a feasible inner-loop or use an algorithm that can handle this issue (e.g., genetic algorithm).

Implementing the linear constraints appropriately was quite helpful in finding better populations and search directions because many optimization algorithms can directly avoid linear constraint infeasibility regions.