An Error Analysis Framework for Process Coupling in Atmospheric Models

PDC22: Theory of Coupling

June 2022

Chris Vogl, Carol Woodward, Hui Wan, Kai Zhang, Shixuan Zhang, Vincent Larson, Phil Rasch, Quan Bui
LLNL, PNNL, UW-Milwaukee, University of Washington, Blue River Technology
Motivation: large time-stepping errors found in EAM simulations

- Strong timestep sensitivity observed in climate simulations
  - reduced all timesteps to 1/6x (e.g., coupling from 30 min to 5 min)
  - 20%-50% change in subtropical marine low cloud amounts
  - see Wan et al. (2021, GMD) & Santos et al. (2021, JAMES)

- Found additional sensitivity to the coupling approach of cloud macrophysics and microphysics
  - no change in coupling or parameterization timesteps
  - no change in number of subcycles
  - strong impacts seen in subtropical stratocumulus decks

- Questions to be answered:
  - would a different coupling approach reduce the timestep sensitivity?
  - can it be determined a priori what the better coupling approach is?
Two coupling problems will be discussed

- Coupling between aerosol processes in EAM

- Coupling between macrophysics and microphysics in EAM
  - idealized atmosphere model
  - EAM with CLUBB and MG2
The modeling and computation of aerosols in EAM remains a challenge—
aerosol lifetime has significant impact on climate—
remains significant source of model uncertainty

Feng et al. (2021) demonstrated significant sensitivity of dust lifetime in EAMv1—
using CAM5 with 30 levels, dust lifetime is 2.6 days with a burden of 22.4 Tg—
using EAMv1 with 30 levels, dust lifetime is 2.4 days, but with a burden of 28.3 Tg—
using EAMv1 with **72 levels**, dust lifetime is **1.9 days** with a burden of 22.2 Tg

Changing the coupling restores proper dust lifecycle—
noticed dust was becoming trapped near surface—
alternative coupling removes the trapping behavior

Seek analysis to explain the improvement
The aerosol model can be described as
\[
\frac{dy}{dt} = \text{(emission)} + \text{(dry removal)} + \text{(turb. mixing)} + \ldots
\]

Pairs of processes, denoted \(A(y)\) and \(B(y)\), can be coupled with various approaches:

- Sequential splitting:

  \[
  \frac{dy_A}{dt} = A(y_A)
  \]

  \[
  \frac{dy_B}{dt} = B(y_B)
  \]

- Parallel splitting:

  \[
  \frac{dy_A}{dt} = A(y_A)
  \]

  \[
  \frac{dy_B}{dt} = B(y_B)
  \]
Which coupling method to use (sequential splitting vs parallel splitting vs ...)?

Unfortunately, the results from the literature seem to fall into one of two categories

Too general in that they apply to our problem, but do not answer the question
    — e.g., requirements for second-order methods, but feasible methods are first order
    — how do we then decide between multiple first-order schemes?

Too specific in that they only answer the question for a specific problem
    — e.g., comparison of coupling methods for a specific idealized model & time integration method
    — how much can we expect those results to generalize to our problem?

Existing work requires information on the time integration used within each process
    — parameterizations might provide the user with multiple time integration options
    — cannot analyze every combination of time integration option
    — prior work indicates that coupling error can dominate the time integration error in some regimes
Quantify the coupling error to develop *a priori* expectation of scheme performance

A semi-discrete approach (exact integration of individual processes) is used to isolate coupling error from discretization error

As an illustration, consider first a two-process problem:

\[
\frac{dy}{dt}(t) = A(y) + B(y) \quad \Rightarrow \quad \frac{dy_A}{dt} = A(y_A) \quad \Rightarrow \quad y_A(t) = y_A(0) + \int_0^t A(y_A) \, d\tilde{t}
\]

\[
\frac{dy_B}{dt} = B(y_B) \quad \Rightarrow \quad y_B(t) = y_B(0) + \int_0^t B(y_B) \, d\tilde{t}
\]

Small quantity expansion obtains the leading order error terms incurred at each \( \Delta t \):

- sequential splitting: \( -\frac{1}{2} \Delta t^2 \left( \frac{\partial B}{\partial y} A - \frac{\partial A}{\partial y} B \right) \)

- parallel splitting: \( -\frac{1}{2} \Delta t^2 \left( \frac{\partial B}{\partial y} A + \frac{\partial A}{\partial y} B \right) \)
Coupling error analysis can determine which splitting method to use

- Consider when the two processes form a “push-pull” system
  \[
  \frac{dy}{dt}(t) = A(y) + B(y) \text{ where } A > 0 \text{ and } B < 0
  \]

- Consider when the “push” and “pull” are both stronger with larger \(y\)
  - mathematically, this is \(\frac{\partial A}{\partial y} > 0\) and \(\frac{\partial B}{\partial y} < 0\)
  - the terms in the leading order error have signs \(\frac{\partial A}{\partial y} B < 0\) and \(\frac{\partial B}{\partial y} A < 0\)
  - because the terms have the same sign, the framework suggests sequential splitting be used

- Consider when the “push” is weaker with larger \(y\), while the “pull” is stronger
  - mathematically, this is \(\frac{\partial A}{\partial y} < 0\) and \(\frac{\partial B}{\partial y} < 0\)
  - the terms in the leading order error have signs \(\frac{\partial A}{\partial y} B > 0\) and \(\frac{\partial B}{\partial y} A < 0\)
  - because the terms have opposite sign, the framework suggests parallel splitting be used
Recall the dust model as a three-process problem:

\[
\frac{dy}{dt} = A_{\text{emission}} + B(y)_{\text{dry removal}} + C(y)_{\text{turb. mixing}}
\]

The semi-discrete error analysis is applied to this problem
- sequential splitting:
  \[
  \frac{1}{2} \Delta t^2 \left[ \frac{\partial C}{\partial y} (A + B) - \left( \frac{\partial B}{\partial y} C - \frac{\partial B}{\partial y} A \right) \right]
  \]
- parallel splitting:
  \[
  \frac{1}{2} \Delta t^2 \left[ \frac{\partial C}{\partial y} (A + B) - \left( \frac{\partial B}{\partial y} C + \frac{\partial B}{\partial y} A \right) \right]
  \]

Dry removal is a “pull” that is stronger with larger \( y \):
- \( \frac{\partial B}{\partial y} < 0, A > 0, \) and \( C < 0 \) mean that \( \frac{\partial B}{\partial y} C > 0 \) and \( \frac{\partial B}{\partial y} A < 0 \)
- because of the opposite signs of the terms, the analysis explains the benefit of parallel splitting

Note the analysis provides insight without knowledge of time integration method
Presentation Overview

- Coupling between aerosol processes in EAM

- Coupling between macrophysics and microphysics in EAM
  - idealized atmosphere model
  - EAM with CLUBB and MG2
Coupling in E3SM Atmosphere Model (EAM) is complex

- EAM’s many processes, coupled every 30 minutes, can be categorized in 2 broad groups
  - dycore + deep convection (dyn-other)
  - stratiform cloud macrophysics + microphysics (mac-mic)

- Coupling is mostly sequential: each group adds 30 minutes worth of increment to state

- The mac-mic group uses sequential subcycling to add the 30-minute increment
  - CLUBB (mac) applies 5-minute increment to state
  - MG2 (mic) applies 5-minute increment to state
  - this subcycle is repeated 6 times
“Isolated” sequential splitting is the default for subcycle coupling

- Consider the coupling between the dyn-other and the mac-mic groups
- Model as a two-process problem with all prognosed quantities grouped into $y(t)$:

$$\frac{dy}{dt} = A(y) + B(y)$$

- Denote the default coupling approach isolated sequential splitting

![Diagram showing the coupling between dyn-other and mac-mic groups]
Forcing method approach breaks isolation (but modifies equations)

- The tendency \( A^* = \frac{y_A^n - y^n}{\Delta t} \) can be applied as a forcing term alongside \( B(y) \)

Recall isolated sequential splitting:

\[
\begin{align*}
\frac{dy_A}{dt} &= A(y_A) \\
\frac{dy_B}{dt} &= B(y_B) + A^*
\end{align*}
\]
Dribbling method does not modify equations (but uses some splitting)

- The tendency $A^* = \frac{y_A^n - y^n}{\Delta t}$ can be applied in sequence with $B(y)$ (dribbled in subcycle)

- Recall isolated sequential splitting:

\[
\begin{align*}
y^n & \quad A(y) \text{ for 30min} \\
\frac{dy_A}{dt} = A(y_A) & \\
\end{align*}
\]
\[
\begin{align*}
y^n & \quad y_A^n \\
\frac{dy_A}{dt} = A(y_A) & \\
\end{align*}
\]
\[
\begin{align*}
y_B^{n,0,*} & \quad y_B^{n,0} \\
A^* \text{ for 5min} & \\
\frac{dy_B^{n,0,*}}{dt} = A^* & \\
\end{align*}
\]
\[
\begin{align*}
y_B^{n,0} & \quad y_B^{n,0} \\
B(y) \text{ for 5min} & \\
\frac{dy_B^{n,0}}{dt} = B(y_B^{n,0}) & \\
\end{align*}
\]
\[
\begin{align*}
y_B^{n,0} & \quad y_B^{n,0} \\
B(y) \text{ for 5min} & \\
\frac{dy_B^{n,1}}{dt} = B(y_B^{n,1}) & \\
\end{align*}
\]
\[
\begin{align*}
y_B^{n,0} & \quad y_B^{n,1} \\
A^* \text{ for 5min} & \\
\frac{dy_B^{n,1}}{dt} = A^* & \\
\end{align*}
\]
\[
\begin{align*}
y_B^{n,1} & \quad y_B^{n,1} \\
B(y) \text{ for 5min} & \\
\frac{dy_B^{n,1}}{dt} = B(y_B^{n,1}) & \\
\end{align*}
\]
\[
\begin{align*}
y_B^{n,1} & \quad y_B^{n,1} \\
B(y) \text{ for 5min} & \\
\frac{dy_B^{n,2}}{dt} = B(y_B^{n,2}) & \\
\end{align*}
\]
\[
\begin{align*}
y_B^{n,1} & \quad y_B^{n,1} \\
B(y) \text{ for 5min} & \\
\frac{dy_B^{n,6}}{dt} = B(y_B^{n,6}) & \\
\end{align*}
\]
\[
\begin{align*}
y_B^{n,1} & \quad y_B^{n,1} \\
B(y) \text{ for 5min} & \\
\frac{dy_B^{n,6}}{dt} = B(y_B^{n,6}) & \\
\end{align*}
\]
Dribbling reduces time-stepping error in EAM

- Dribbling solution shows less jumps and is closer to reference solution
- Note systematic drifts in stratocumulus region

Comparison of EAM results using default coupling (CNTL, v1_CTRL) with results using mac-mic dribbling (DRIB, v1_Dribble)
Semi-discrete error analysis framework derives subcycle coupling error

- Small quantity expansion obtains the leading order error terms incurred at each $\Delta t$:
  - isolated sequential splitting: $\frac{1}{2} \Delta t^2 \left( \frac{\partial B}{\partial y} A - \frac{\partial A}{\partial y} B \right)$
  - dribbling method: $\frac{1}{2} \Delta t^2 \left( \frac{1}{M} \frac{\partial B}{\partial y} A - \frac{\partial A}{\partial y} B \right)$ where $M$ is the number of subcycles
  - forcing method: $\frac{1}{2} \Delta t^2 \left( - \frac{\partial A}{\partial y} B \right)$

- Unlike the aerosol case, $y$ now represents many highly-coupled variables
  - $A, B$ might be a “push” for a particular variable, location, and climate regime and a “pull” for others
  - cancellation of error terms across all variables, global locations, and climate regimes is not likely

- Instead, consider the effect of $M$ on the $\frac{\partial B}{\partial y} A$ term
  - dribbling method reduces the term because it has more frequent coupling from $A$ to $B$
  - forcing method removes the term entirely because it has continual coupling from $A$ to $B$
  - easier-to-implement dribbling method with $M \gg 1$ can inform on performance of forcing method
Presentation Overview

- Coupling between aerosol processes in EAM
- Coupling between macrophysics and microphysics in EAM
  - idealized atmosphere model
  - EAM with CLUBB and MG2
Idealized atmosphere model to analyze dribbling result

- Recall the idealized atmosphere model presented earlier by Hui Wan

- **Supersaturation equation:**
  \[ \frac{ds}{dt} = \delta(l)D + F - C(s) \]
  - $D$ is the rate of supersaturation produced by resolved fluid dynamics and radiation
  - $F$ is the rate of supersaturation produced by surface evaporation and turbulent transport
  - $C(s)$ is the rate of water vapor condensation
  - $\delta(l)$ captures “cloud liquid -> radiative cooling -> supersaturation” feedback mechanism

- **Cloud liquid equation:**
  \[ \frac{dl}{dt} = \chi(s)C(s) - I(l) \]
  - $I(l)$ is the rate at which cloud liquid is converted to rain
  - $\chi(s)$ captures “radiative cooling -> turbulence -> cloud liquid” feedback mechanism

- Timescales are introduced consistent with near-instantaneous condensation
  - $C(s) = s/\tau_s$, where $\tau_s = 1s$
  - $I(l) = \alpha(l/l_c)^{2.47}/\tau_l$, where $\tau_l = 1800s$
Simplified model as two-process problem:

$$\frac{d}{dt} \left[ \begin{array}{c} S \\ l \end{array} \right] = \left[ \begin{array}{c} \delta(l) D \\ F - C(s) \chi(s) C(s) \end{array} \right] + \left[ \begin{array}{c} 0 \\ -I(l) \end{array} \right]$$

Recall error terms for subcycle problem

- isolated sequential splitting: $$\frac{1}{2} \Delta t^2 \left( \frac{\partial B}{\partial y} A - \frac{\partial A}{\partial y} B \right)$$
- dribbling method: $$\frac{1}{2} \Delta t^2 \left( \frac{1}{M} \frac{\partial B}{\partial y} A - \frac{\partial A}{\partial y} B \right)$$
- forcing method: $$\frac{1}{2} \Delta t^2 \left( - \frac{\partial A}{\partial y} B \right)$$

Because dribbling approaches reference...

- $$\frac{\partial B}{\partial y} A$$ is significant relative to $$\frac{\partial A}{\partial y} B$$
- more frequent coupling needed from $$A$$ to $$B$$
- this indicates forcing method will perform well
- empirical results confirm forcing method performance

Coupling error analysis + empirical results correctly selects forcing method

Simplified Atmosphere Model Results

Cloud liquid (g/kg)

- reference
- iso. seq.
- dribbling (M=6)
- dribbling (M=12)
- dribbling (M=24)
- forcing

results showing error decrease from (i) one method to next (ii) increasing M
Overview

- Coupling between aerosol processes in EAM
- Coupling between macrophysics and microphysics in EAM
  - simplified atmosphere model
  - EAM with CLUBB and MG2
There is additional splitting in EAM for macrophysics-microphysics

- **CLUBB and MG2 are split in EAM:**
  \[ B = B_1 + B_2 \Rightarrow \frac{dy}{dt} = A(y) + B_1(y) + B_2(y) \]

- Isolated sequential split

- Dribbling method

- Forcing (mac) method
Semi-discrete error analysis framework provides error for three processes

- Leading-order error for three-process-with-subcycle problem:

\[
\frac{dy}{dt} = A(y) + B_1(y) + B_2(y)
\]

- iso. seq. split.: \( \frac{1}{2} \Delta t^2 \left( \frac{\partial B_1}{\partial y} A + \frac{\partial B_2}{\partial y} A + \frac{1}{M} \left[ \frac{\partial B_2}{\partial y} B_1 - \frac{\partial B_1}{\partial y} B_2 \right] - \frac{\partial A}{\partial y} B \right) \)

- dribbling: \( \frac{1}{2} \Delta t^2 \left( \frac{1}{M} \left[ \frac{\partial B_1}{\partial y} A + \frac{\partial B_2}{\partial y} A + \frac{\partial B_2}{\partial y} B_1 - \frac{\partial B_1}{\partial y} B_2 \right] - \frac{\partial A}{\partial y} B \right) \)

- forcing (mac): \( \frac{1}{2} \Delta t^2 \left( \frac{1}{M} \left[ \frac{\partial B_2}{\partial y} A + \frac{\partial B_2}{\partial y} B_1 - \frac{\partial B_1}{\partial y} B_2 \right] - \frac{\partial A}{\partial y} B \right) \)

- Consider the effect of \( M \) on the \( \frac{\partial B}{\partial y} = \frac{\partial B_1}{\partial y} A + \frac{\partial B_2}{\partial y} A \) terms
  - dribbling method **reduces** the terms, as expected b/c more frequent coupling from \( A \) to \( B_1 \) and \( B_2 \)
  - forcing (mac) method **removes** the \( \frac{\partial B_1}{\partial y} A \) term but only **reduces** the \( \frac{\partial B_2}{\partial y} A \) term
  - thus, forcing (mac) method will outperform dribbling **only if** \( \frac{\partial B_1}{\partial y} A \) is significant
Error analysis suggests dribbling (and maybe forcing) for EAM coupling

- **EAM as three-process problem:**
  \[
  \frac{dy}{dt} = A(y) + \underbrace{B_1(y) + B_2(y)}_{\text{other}} + \underbrace{B_1(y)}_{\text{(CLUBB)}} + \underbrace{B_2(y)}_{\text{(MG2)}}
  \]

  **Recall...**

  iso. seq. split.:
  \[
  \frac{1}{2} \Delta t^2 \left( \frac{\partial B_1}{\partial y} A + \frac{\partial B_2}{\partial y} A + \frac{1}{M} \left[ \frac{\partial B_1}{\partial y} B_1 - \frac{\partial B_1}{\partial y} B_2 \right] - \frac{\partial A}{\partial y} B \right)
  \]

  dribbling:
  \[
  \frac{1}{2} \Delta t^2 \left( \frac{1}{M} \left[ \frac{\partial B_1}{\partial y} A + \frac{\partial B_2}{\partial y} A + \frac{\partial B_1}{\partial y} B_1 - \frac{\partial B_1}{\partial y} B_2 \right] - \frac{\partial A}{\partial y} B \right)
  \]

  forcing (mac): 
  \[
  \frac{1}{2} \Delta t^2 \left( \frac{1}{M} \left[ \frac{\partial B_1}{\partial y} A + \frac{\partial B_1}{\partial y} B_1 - \frac{\partial B_1}{\partial y} B_2 \right] - \frac{\partial A}{\partial y} B \right)
  \]

- **Because iso. seq. splitting improves slightly but retains jumps...**
  - frequent coupling between CLUBB & MG2 \((\frac{\partial B_2}{\partial y} A \& \frac{\partial B_1}{\partial y} B_2)\) is beneficial
  - more coupling needed elsewhere \((\frac{\partial B_1}{\partial y} A, \frac{\partial B_2}{\partial y} A, \frac{\partial A}{\partial y} B)\) to decrease jumps

- **Because dribbling quickly smooths jumps...**
  - frequent coupling from EAM to CLUBB-MG2 \((\frac{\partial B_1}{\partial y} A \& \frac{\partial B_2}{\partial y} A)\) reduces jumps
  - forcing (mac) may be beneficial if EAM affects CLUBB more than MG2

- **Dribbling should be used at 30 min timesteps for better solution**
Semi-discrete error analysis provides insight into various EAM results

- Developed semi-discrete analysis approach to focus on coupling error
  - for uniform “push-pull” systems, analysis can determine which coupling method to use
  - for general systems, analysis & empirical testing provides insight

- Parallel splitting improvement in dust results explained
  - error analysis indicates that parallel splitting has less coupling error
  - no information about process integration needed

- Forcing method advantage predicted in idealized mac-mic model

- Dribbling method advantage explained in full mac-mic model
  - sequential split **only focuses** on frequent coupling between CLUBB & MG2
  - dribbling method **incorporates more coupling** from EAM to CLUBB & MG2
  - forcing method **may** be beneficial if “rest of EAM” impact on CLUBB > MG2

- Can reduce/remove jumps and have consistent mean state with dribbling method
This work was supported by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research (BER), and Office of Advanced Scientific Computing Research (ASCR) via the Scientific Discovery through Advanced Computing (SciDAC) program.

Disclaimer
This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.
Coupling error analysis provides insight to three process empirical results

- Simplified model as three-process problem: \[ \frac{d}{dt} \left[ S \right] = \begin{bmatrix} \delta(l) D \\ F - C(s) \\ 0 \end{bmatrix} + \begin{bmatrix} A(y) \\ \chi(s) C(s) \\ -I(l) \end{bmatrix} \]

- Because iso. seq. does not approach reference...
  - \( \frac{\partial B_2}{\partial y} B_1 \) & \( \frac{\partial B_1}{\partial y} B_2 \) are not significant relative to other terms

- Because dribbling does approaches reference...
  - \( \frac{\partial B_1}{\partial y} A \) & \( \frac{\partial B_2}{\partial y} A \) are significant relative to \( \frac{\partial A}{\partial y} B \)
  - forcing (mac) performance depends on if \( \frac{\partial B_1}{\partial y} A \gg \frac{\partial B_2}{\partial y} A \)
  - empirical results confirm that this is indeed the case