

A Computational Study of Data Assimilation  
for a Reaction-Diffusion Equation

An Undergraduate Honors Thesis  
Submitted in Partial fulfillment of  
University Honors Program Requirements  
University of Nebraska-Lincoln by

Collin Victor, BS  
Mathematics and Computer Science  
College of Arts and Sciences

March 3, 2018

Faculty Mentors:  
Adam Larios, PhD, Mathematics

## ABSTRACT

This research project applied the Azouani-Olson-Titi data assimilation algorithm [1, 2] to the 1D Chafee-Infante equation in order to investigate the potential for different grid configurations. It was discovered that for certain  $\nu$  values convergence could be increased by replacing a uniform grid by a moving cluster of points. In addition to this, this study found a heuristic argument solving the inverse problem of determining the minimum length given the minimum number of nodes required for convergence. All simulations were conducted using a semi-implicit convex splitting numerical scheme [?] and spatial derivatives were approximated using second-order finite difference approximations.

**Key Words:** Data Assimilation, Chafee-Infante, Mathematics

DEDICATION

This research was partially supported by grant numbers... .

Official  
UCARE st  
goes here

# A COMPUTATIONAL STUDY OF DATA ASSIMILATION FOR A REACTION-DIFFUSION EQUATION

COLLIN VICTOR

## 1. INTRODUCTION

In many simulation-driven fields that involve real-world data, such as numerical prediction of weather on earth and on the sun, a central problem arises. Namely, initial data may be known at a small number of locations, such as the temperature or wind speed measured at weather stations. Without this data, evolving the system in time is highly error prone. Data assimilation is an approach that eliminates the need for complete initial data. Instead, it incorporates incoming data into simulations by penalizing deviations from the data, driving the system to the “true” solution.

Classical data assimilation is based on an algorithm known as Kalman Filtering. This technique and its modifications have a long history, but it still suffers from many drawback when applied to nonlinear problems, such as those in turbulent flows. In 2013 a new approach was developed, known as the Azouani-Olson-Titi (AOT) algorithm [1, 2]. The AOT algorithm incorporates the data at the level of the underlying partial differential equation. It acts as a continuous feedback-control mechanism, using incoming data to drive the simulation to the correct solution. The algorithm is given by:

$$\frac{d}{dt}v = F(v, t) - \mu(I_h(y) - I_h(v)), v(0) = 0.$$

Here,  $\mu \geq 0$  is a constant, and  $I_h$  is an interpolation operator for a grid,  $h$ .  $y$  is given by the following:

$$\frac{d}{dt}y = F(y, t), y(0) = y_0.$$

The number of grid points in  $h$  is a very important parameter for data assimilation. If the grid has too few points, the solution will never converge, but having more points than required adds unnecessary computational complexity. In addition to this, in real world scenarios (such as weather prediction) where these grid points are marked by sensors (or weather stations), minimizing the number of grid points reduces the financial cost of sensor production and placement. This study examines whether we can achieve the same level of convergence using fewer points and without sacrificing numerical stability by using different grid configurations.

This study examines the potential for non-uniform sensor placement as a viable strategy for data assimilation for reaction-diffusion equations. The specific equation used in this study is the 1D Chafee-Infante equation:

$$\begin{aligned} u_t - \nu u_{xx} &= u - \alpha u^3 \\ u(x, 0) &= u_0(x) \end{aligned}$$

where  $\alpha > 0, \nu > 0$  with Dirichlet boundary conditions.

Normally data assimilation uses data collected at points on a uniform grid. This study examines how the use of a moving cluster of sensors instead of a static uniform grid affects the convergence of the system to the true solution in the context of reaction-diffusion equations. By altering the placement of the nodes it is hoped that it will reduce the minimum number of grid points required for convergence to the true solution.

This study has experimentally found a bound for the number of grid points in  $h$  for the uniform grid required for convergence using periodic initial data. A heuristic argument for a relationship between the size of  $h$  required for convergence and the minimum length,  $\lambda$ , using only data assimilation was also found. In addition to this, it was found that for certain values of  $\nu$  that using a moving cluster of points does in fact require fewer grid points.

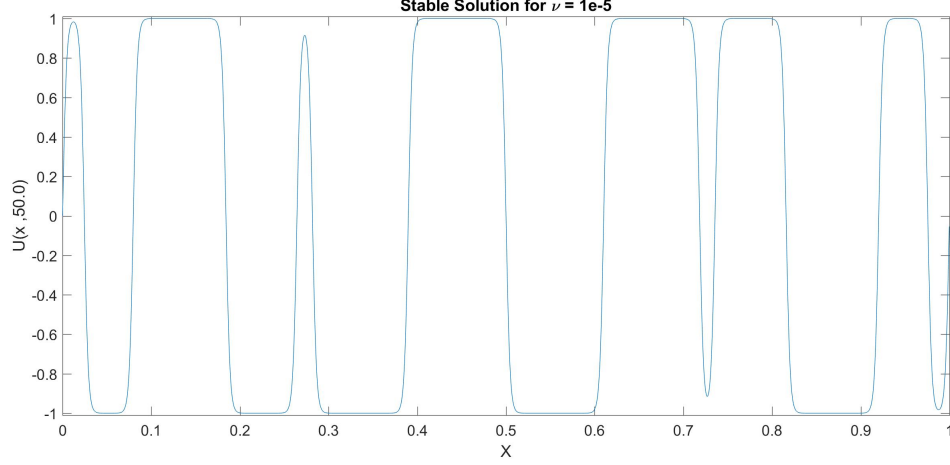


FIGURE 2.1. Example of a stable state solution for Chafee-Infante equation  $\alpha = 1$   $\nu = 1e-5$   $N = 2^{12}$

## 2. PRELIMINARIES

The Chafee-Infante equation is given by:

$$u_t - \nu u_{xx} = u - \alpha u^3,$$

$$u(x, 0) = u_0(x)$$

where  $\alpha > 0, \nu > 0$ , and with Dirichlet boundary conditions. We solved this equation on a uniform grid of  $N$  points distributed uniformly on  $[0, 1]$  using discrete time-steps of size  $dt$ . To ensure that initial data had a fully-resolved spectrum we used:

$$u_0 = \sum_{k=1}^{\frac{N}{4}} a_k \sin(2\pi kx)$$

where  $a_k$  is some integer. The Chafee-Infante equation is a good equation to use for this study because of its well-documented (?) features. Solutions to this equation have three distinct phases. The equation initially develops based on the initial data and reaches a mostly stable state. This stable state consists of large “blobs” of positive and negative activity with transition layers in between (see fig. 2.1). The system will then eventually decay to 0, although this behavior is not what we are trying to capture with data assimilation.

This study uses a semi-implicit convex splitting scheme from [?] to solve for the actual

Add Bibliography entry

solution and the simulated data assimilation solution at every time-step. The numerical scheme is given by:

$$U_k^{n+1} - U_k^n = dt(\nu D_{xx} + 1 + 2\alpha)U_k^{n+1} + dt(-2\alpha - \alpha U_k^{n2})U_k^n$$

where  $U_k^n = u(x_k, t_n)$  where  $x_k = kdx$ ,  $t_n = ndt$ .  $D_{xx}$  is the operator  $\frac{\partial^2}{\partial x^2}$ , here  $\frac{\partial^2 U_k^{n+1}}{\partial x^2}$  is given by the second-order finite difference approximation  $\frac{U_{k-1}^{n+1} - 2U_k^{n+1} + U_{k+1}^{n+1}}{dx^2}$ . This results in the below system which needs to be solved at every time-step:

$$(1 - dt(\nu D_{xx} + 1 + 2\alpha))U_k^{n+1} = (1 + dt(-2\alpha - \alpha U_k^{n2}))U_k^n$$

While this method requires us to solve a system of equations at each time-step, it is much more stable than other methods and allows us to have a much larger time-step than other methods. In addition to this, the resulting matrix is tri-diagonal, solving this system only has a time complexity of  $O(n)$ .

In order to apply data assimilation, the below system needs to be solved at each time-step:

$$(1 - dt(\nu D_{xx} + 1 + 2\alpha))V_k^{n+1} = (V_k^n + dt((-2\alpha V_k^n - \alpha V_k^{n3}) + \mu(I_h(U^n) - I_h(V^n))))$$

Two different types of data assimilation were tested in this study, each modifying  $I_h$ . The first type was the standard uniform grid,  $I_h(S)$  is the linear interpolation of  $S$  over a uniform grid of points. The second type used a moving cluster of points which we will refer to as a sweeping probe. Data assimilation via a sweeping probe uses  $I_h(S) = I_{h(t)}(S)$ . Instead of uniformly distributing  $m$  points,  $m$  consecutive points are used at each time-step.

$h(t_n) = \{x_k | k \in [nm \pmod{N}, nm + m - 1 \pmod{N}] \subset \mathbb{Z}\}$  Initially the first  $m$  points are used, then at the next time-step, the next  $m$  points are used. When the probe reaches the end of the domain, it wraps back around to the other side.

Simulations run for this study have two distinct phases, an initial ramp up phase allows the actual solution to transition to a stable state, after this the actual solution and the simulated AOT algorithm solution run in tandem using some grid configuration of  $m$  points. After a set number of time-steps, the error is measured and the second phase is repeated

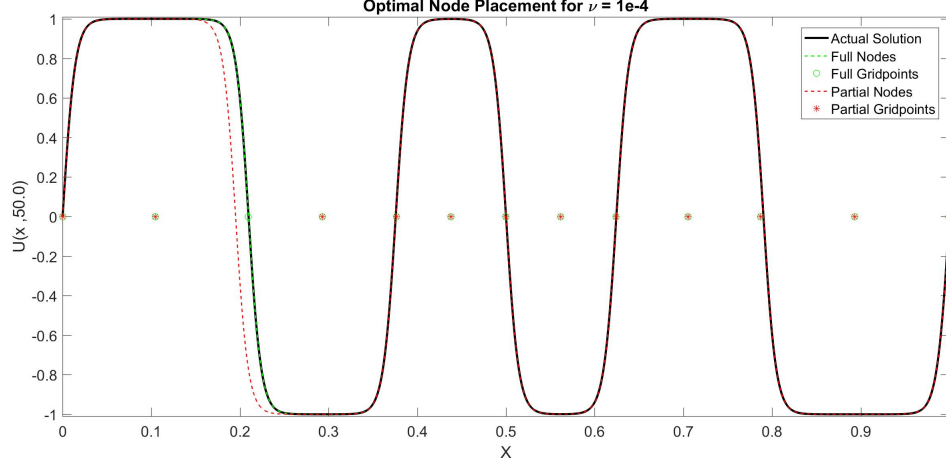


FIGURE 3.1. Data Assimilation can converge or diverge based on one grid point  $\alpha = 1$   $\nu = 1e - 4$   $N = 2^{12}$

using a binary search algorithm to determine the minimum value of  $m$  for which there is sufficient convergence. The binary search is started on the interval  $[2, N]$ , this interval will always contain a minimum value of  $m$  as  $m = N$  will always converge and the end points ( $m = 2$ ) are required nodes for the uniform grid.

### 3. PARAMETER ESTIMATION

In this section detail how to find the minimum length  $\lambda$  using nothing but data assimilation.

Through experimental trials we have found that the only requirement for consistent convergence to the true solution is to have at least one grid point in each transition layer and one in each “blob”. It can be seen from fig. 3.1 that if a transition layer does not have a grid point, the solutions may never converge. In fact by using a postereori examination of the solution, we found that if we knew where the transition layers were going to form, we can achieve convergence using this grid configuration. It seems that the most important parameter in determining the minimum number of nodes required for convergence,  $M$ , is the minimum length of the “blobs”,  $\lambda$ .

From repeated experiments we discovered that  $M \propto \frac{1}{4}\nu^{-\frac{1}{2}}$  in the worst-case scenario (see 4.1). The parameter  $\alpha$  affects the amplitude of the solutions, but  $\nu$  is directly related to the number of “blobs” in the solution. The most nodes (using an optimal placement) will

Search literature for formula for minimum length



be required when all the “blobs” have the same size  $\lambda$ . In this case, for there being  $n_b$  “blobs” we get  $\lambda = \frac{L}{n_b}$ , where  $L$  is the length of the domain. We also know  $M = 2n_b + 1$  or equivalently  $n_b = \frac{M-1}{2}$ .

$$\begin{aligned}\lambda &= \frac{L}{n_b} \\ &\approx \frac{2L}{M} \\ &\approx \frac{8L}{\nu^{-\frac{1}{2}}} \\ &\approx 8L\sqrt{\nu}\end{aligned}$$

Using only principles of data assimilation we discovered a solution for the inverse problem  $\lambda \approx 8L\sqrt{\nu}$ .

#### 4. SWEEPING PROBE DATA ASSIMILATION

The main finding of this study was in the potential for sweeping data assimilation. In general it was found that for  $\mu$  large enough for which the system is stable that in general far fewer nodes are required for data assimilation via a sweeping probe for sufficiently small  $\nu$  (see fig. 4.1). Of course when  $\nu$  is large it is more efficient to use a uniform grid, but there is a threshold where the methods are roughly equivalent. For  $\mu = 100$  this value occurs at  $\nu \approx 2e - 4$ . For smaller  $\nu$  it is more effective to use the sweeping probe while for  $\nu$  larger it will be more effective to use the uniform grid. Moreover the convergence of the sweeping probe is not sensitive to the number of nodes, adding extra nodes past the minimum only serve to speed up the convergence. This is not always the case for a uniform grid as adding more points can shift the positioning of every point on the grid, meaning that convergence is not guaranteed when adding nodes. Convergence is only guaranteed for nested sets of nodes, however this matters increasing little as  $M$  increases.

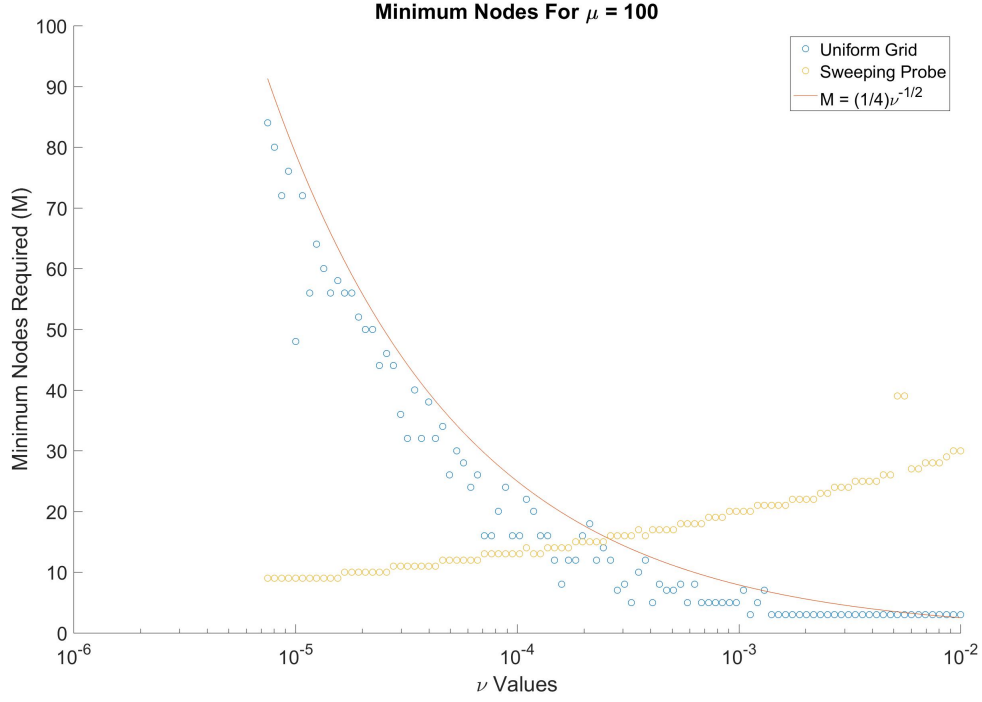


FIGURE 4.1. Minimum Nodes for Uniform Grid and Sweeping Probe required for convergence of  $5e-15$ ,  $t = 50$  units after ramp up period  $\alpha = 1$   $N = 2^{12}$   $\mu = 100$

## 5. CONCLUDING REMARKS

This study has established data assimilation as a viable alternative to the uniform grid. This requires far fewer points for convergence for certain  $\nu$  values and doesn't require nodes to be placed in a special configuration. More work is required to determine the viability of this type of data assimilation in higher dimensions and its usability for other equations.

## REFERENCES

1. Abderrahim Azouani, Eric Olson, and Edriss S. Titi, *Continuous data assimilation using general interpolant observables*, J. Nonlinear Sci. **24** (2014), no. 2, 277–304. MR 3183055
2. Abderrahim Azouani and Edriss S. Titi, *Feedback control of nonlinear dissipative systems by finite determining parameters—a reaction-diffusion paradigm*, Evol. Equ. Control Theory **3** (2014), no. 4, 579–594. MR 3274649

(Collin Victor) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NEBRASKA–LINCOLN, LINCOLN, NE 68588-0130, USA

*E-mail address*, Collin Victor: `collinzacharyvictor@gmail.com`