Applied and Computational Mathematics
2021; 10(6): 121-137
http://www.sciencepublishinggroup.com/j/acm
doi: 10.11648/j.acm.20211006.11
ISSN: 2328-5605 (Print); ISSN: 2328-5613 (Online)

An Accurate and Stable Filtered Explicit Scheme for Biopolymerization Processes in the Presence of Perturbations

Lisa Davis[1], Faranak Pahlevani[2,*], Timmy Susai Rajan[2]

[1]Department of Mathematical Sciences, Montana State University, Bozeman, United States
[2]Division of Science & Engineering, Penn State University-Abington, Abington, United States

Email address:

lisa.davis@montana.edu (L. Davis), fxp10@psu.edu (F. Pahlevani), tms6516@psu.edu (T. S. Rajan)
*Corresponding author

To cite this article:

Lisa Davis, Faranak Pahlevani, Timmy Susai Rajan. An Accurate and Stable Filtered Explicit Scheme for Biopolymerization Processes in the Presence of Perturbations. *Applied and Computational Mathematics*. Vol. 10, No. 6, 2021, pp. 121-137.
doi: 10.11648/j.acm.20211006.11

Received: August 31, 2021; **Accepted:** October 13, 2021; **Published:** November 5, 2021

Abstract: The focus of this paper is the development, numerical simulation and parameter analysis of a model of the transcription of ribosomal RNA in highly transcribed genes. Inspired by the well-known classic Lighthill-Whitham-Richards (LWR) traffic flow model, a linear advection continuum model is used to describe the DNA transcription process. In this model, elongation velocity is assumed to be essentially constant as RNA polymerases move along the strand through different phases of gene transcription. One advantage of using the linear model is that it allows one to quantify how small perturbations in elongation velocity and inflow parameters affect important biology measures such as Average Transcription Time (ATT) for the gene. The ATT per polymerase is the amount of time an individual RNAP spends traveling through the DNA strand. The numerical treatment for model simulations includes introducing a low complexity and time accurate method by adding a simple linear time filter to the classic upwind scheme. This improved method is modular and requires a minimal modification of adding only one line of code resulting in increased accuracy without increased computational expense. In addition, it removes the overdamping of upwind. A stability condition for the new algorithm is derived, and numerical computations illustrate stability and convergence of the filtered scheme as well as improved ATT estimation.

Keywords: Advection Equation, Lighthill-Whitham-Richards Model, Ribosomal RNA, RNA Polymerases, Time Filter, Traffic Flow Models, Transcription Time, Upwind Scheme

1. Introduction

This paper considers a simple one compartment model that is to be used for future work in the construction of a larger ribosome assembly model. Ribosome assembly relies on several processes including the transcription of ribosomal RNA, the transcription of mRNA of ribosomal proteins and the translation of ribosomal proteins. Each of these processes can be considered as an individual compartment that is coupled to the others through inputs and outputs. Each compartment process also relies on the availability of various amino acids as well as initiation and termination complexes and signaling molecules. The availability of many of these quantities vary depending on the bacterial response to changes in the supply of environmental nutrients. The goal of this paper is to begin with a simple compartment model to understand how small perturbations in two key parameters influence the output of the compartment, a quantity providing a crucial measure that can be viewed as a proxy for the efficiency of protein production. The focus is on using a simple continuum model that allows one to explore the accuracy of a new numerical method and to use analytical results for the associated parameter analysis.

The continuum model takes the form of a linear advection equation where the parameters of interest determine advection speed and inflow. A time filter technique is combined with the standard upwind algorithm for model simulations. It is shown that the time filter can be applied repeatedly, and the

consistency and stability requirements for these approaches are derived. Numerical results demonstrate that the filtered approach improves accuracy for the same computational expense as the original upwind scheme.

In the following subsection, we give a brief background of the relevant biological application. The remainder of the paper is organized as follows. Section 2 gives a brief derivation of the transcription model, and a rescaling provides a dimensionless form of the continuum model which is convenient for introducing the computational schemes. The algorithms used to numerically approximate solutions to the mathematical model are given in Section 4. A time filtering approach is combined with the classical upwind scheme, and consistency, stability and convergence is addressed. Section 5 illustrates the numerical results. Finally, Section 6 contains a discussion of two important parameters in the transcription model, and it is shown that the filtered scheme provides accurate numerical approximations of biological quantities of interest related to these parameters. Perturbations of these key parameters are interpreted in the context of the original transcription model.

Protein synthesis is a complex biopolymerization process requiring an immense amount of cellular energy. Transcription and translation are two key stages of protein synthesis, and a full understanding and model characterization of these mechanisms is an active area of research for experimentalists and modelers [1, 2, 3, 4]. Transcription involves the transfer of genetic information from DNA to several types of RNA including messenger mRNA, transfer RNA, ribosomal RNA and others. Transcription begins when an RNA polymerase (RNAP) binds to a promoter sequence of a gene (*initiation*) and then proceeds through the *elongation* process. During elongation, the RNAP motors along the coding region of the gene, reading the DNA strand and generating a single-stranded RNA copy. Transcription ends at the gene's *termination* region, where the RNAP releases the nascent copy of RNA and unbinds from the DNA strand. When mRNA is produced from this transcription process, the mRNA then becomes the ribbon that is translated by a ribosome to form a chain of amino acids that will eventually become a protein. When the transcription process yields rRNA or tRNA, it is not translated but provides necessary building blocks for other RNA-protein complexes, such as ribosomes, to form. Ribosomes translate mRNA strands in order to form protein, and this activity accounts for a large amount of the energy use in the cell. A DNA strand contains several genes, each of which codes for a particular type of RNA, and the *rrn* operon codes to produce the rRNA strand. The ribosomal assembly process and subsequent protein production depends explicitly on the cell's ability to produce rRNA via transcription of the *rrn* operon.

Parameters that are key to the model control the initiation rate and the elongation rate of RNA polymerase on the *rrn* operon. While the density of polymerases on most genes is believed to be low, the *rrn* operon is a prototypical example of a gene in the bacteria *E. coli* where the density of RNAPs is observed to be very high. In such genes, transcriptional elongation of RNAPs is not uniform over the extent of the strand or over the time horizon of a cell cycle. Indeed, bacteria are capable of adjusting their growth rate according to the available resources in their environment, and they can respond quickly and efficiently to sudden changes in these resources. Environmental changes can affect both the initiation and elongation rates of the RNAPs transcribing the gene. This paper considers a simple continuum model for transcription which focuses on the crucial biological measure of average transcription time and the effects of small perturbations on this measure.

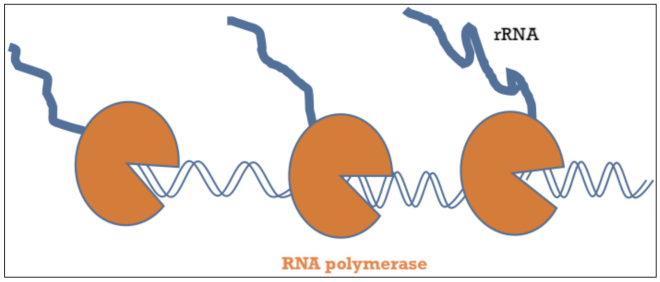


Figure 1. *Schematic diagram of the transcription process on the rrn operon. Three RNA polymerase transcribe the DNA strand, and nascent rRNA strands can be seen emanating from each polymerase.*

2. Mathematical Models

In its simplest form, the transcription process is characterized by the RNA polymerase moving in a preferred direction with some velocity on an individual gene that forms a segment of DNA strand. Many types of mathematical models have been employed to capture the overall behavior of this complex biological phenomena, and these date back as far as the 1960's and continue through present day. Such models range from the stochastic, such as the well-known Totally Asymmetric Simple Exclusion Process (TASEP) and its variants, to systems of ODEs, and finally to continuum models. Early work using stochastic approaches to model protein synthesis can be found in [5, 6] with fundamental work on Markov processes that provided the statistical theoretical foundations of TASEP done in 1970 by Spitzer [7]. As early as the 1980's, systems of ordinary differential equations were explored as a means of determining the rate-limiting step [8]. At the turn of the century, interest in the TASEP models resulted in the analysis and derivation of closed form solutions and model extensions to more biologically relevant constructions [9, 10, 11, 12, 13]. Ordinary and partial differential equations have also been explored in recent years to establish connections between the stochastic, the discrete and continuous models [14, 15, 16, 17, 18].

One simple continuum model that has been used to explore this phenomena is the classical first order traffic flow model [19, 20, 21]. The well-known Lighthill-Whitham-Richards (LWR) continuum model was proposed by researchers in the 1950's in order to analyze and quantify traffic flow, but it was soon dismissed by traffic researchers because it didn't account for driver decision-making and response times during

changing road conditions. More recently there has been debate over higher order continuum models that smooth out discontinuities in density and their ability to predict negative flows and speeds [22], and others have proposed models that ensure that the trajectory of a car is not influenced by what happens behind it on the road [23, 24]. These types of models are still being compared to assess their suitability for a variety of modeling situations such as stop-and-go phenomena as well as others [25].

The LWR equation takes the form of a nonlinear conservation law that describes the density of cars travelling on a one-lane roadway [26, 27]. The density evolves in time and space according to its relationship with the governing velocity field. Inspired by these traffic flow models, the authors use a simple model problem to focus on the fundamental relationships between flow and density, the parameter analysis and the development of a numerical scheme that provides a computationally inexpensive boost to accuracy. The fundamental property one seeks to translate from traffic flow phenomena to transcription is the relationship between the flow rate and the RNAP density on the strand. *The flow rate is interpreted as the number of RNAPs passing a specific position on the gene per unit time, and the density of RNAPs is the average number of RNAPs per length of the gene.*

The main biological quantity of interest is related to the gene's overall transcription rate and the time to transcribe the length of the strand. This quantity measures the gene's ability to be transcribed by many RNAPs simultaneously, and this is considered to be a proxy for the amount of protein that the bacteria can produce in a life cycle. This can be described mathematically in terms of the model variables and parameters. The *average transcription time per polymerase* is the amount of time an individual RNAP spends during the initiation, elongation and termination processes. For the simple case considered here, one can rely on the method of characteristics to compute trajectories of individual RNAPs on the DNA strand, and thereby compute transcription times of those individual particles using an analytical formula. Upon initiation at the left boundary, the RNAP particle traverses the domain along a path governed by the flow velocity $v(x,t)$. This same particle then exits the domain at a later time.

3. A Simple Continuum Model

We seek to describe the elongation of RNAPs on a DNA strand in the same abstract manner as cars travelling around a roadway. This involves the simplification that the DNA strand is *unwound* and identified with a long, straight roadway on which a velocity field is defined. In the general conservation law form, the velocity field may depend on the density of the RNAPs as a function of space and time, but we consider a simple case here in order to focus on the parameter analysis and the numerical schemes.

The classical PDE model in conservation law form is given as

$$\rho_t + [v(x,t)\rho]_x = 0 \quad x \in (0,1), \quad T > 0 \tag{1}$$

with boundary and initial conditions given by

$$f(\rho(0,T)) = f_\ell, \qquad \rho(x,0) = \rho_0 \tag{2}$$

where the flow or "flux" is defined as

$$f(\rho) = v(x,T)\rho(x,T) \tag{3}$$

and $v(x,T)$ is a velocity field. The inflow boundary is at $x = 0$ with $x = 1$ as outflow. The inflow boundary represents the initiation site or the promoter site for the gene under consideration, and the outflow boundary corresponds to the gene's termination site. The variable $\rho(x,t)$ denotes the density of the RNAPs at the spatial point x along the strand at time t. The well-known LWR model is nonlinear and uses the elongation velocity field $v(x,t) = v_{max}(x,t)(1 - \rho(x,t))$, which describes the speed at which the RNAP is travelling on the strand. Low density corresponds to RNAPs elongating near a maximum speed $v_{max}(x,t)$ while a high density of RNAPs on the strand results in a lower elongation velocity. Even using constant initial and inflow conditions in (2) that result in an equilibrium state for the density, $\rho(x,t)$, in the presence of velocity perturbations, one can encounter shock and rarefaction waves using the classical LWR model. Here we begin with a linear model with discontinuous solutions to explore parameter analysis in a setting where the discontinuities in the solution are completely understood.

As a model problem, we consider the simplest choice for the velocity field in the case where it is constant $v(\rho) = v_{max}$. Using this expression for velocity in equation (1) leads to a flux function that is defined to be $f(\rho) = v_{max}\ \rho$, and the linear PDE given by

$$\frac{\partial \rho}{\partial T} + \frac{\partial}{\partial x}\left(v_{max}\rho\right) = 0, \ \ 0 < x < 1, T > 0 \tag{4}$$

with a boundary condition and an initial condition given by

$$\rho(x,0) = \rho_0 \tag{5}$$

$$f(\rho(0,T)) = v_{max}\rho(0,T) = f_\ell \tag{6}$$

The spatial variable is scaled so that $x = 1$ corresponds to the length of the gene on the DNA strand. The constant v_{max} is measured as the fraction of the strand that the RNAP transcribes per second. The time variable T is measured in seconds. The density ρ is measured with units of the total number of RNAPs per strand, and with the flux used in (4), the inflow condition means that $v_{max}\rho(0,T) = f_\ell$. The PDE model in (4) is known as the linear advection equation, and it can be scaled to a convenient form. Many quantities can be calculated analytically for this equation, and numerical methods for solving this equation are well understood. Rescaling the equation to a dimensionless form allows for a thorough parameter analysis and testing of the numerical scheme. The details of the rescaling are given in Section 7, and in the sections that follow, we simply work with the dimensionless form of the model. Connections are made

to the original biological model once the numerical results are presented.

The parameters v_{max}, f_ℓ and ρ_0 in (4) - (6) are important for model analysis. Some parameters may be dealt with through dimensional analysis; for example, by proper scaling of the original dimensioned PDE, v_{max} is set to equal 1. However, perturbations of this parameter are important for the biological application, and Section 6 explores this.

3.1. Linear Model Problem

Here we focus on the linear model and introduce a time filter approach to improve the numerical approximations obtained through the classical upwind scheme. After nondimensionalizing in Section 7, the linear problem has the form

$$\rho_t + \bar{v}\rho_x = 0 \tag{7}$$
$$\rho(x,0) = \rho_0 \tag{8}$$
$$f((0,t)) = \bar{v}\rho(0,t) = \rho_I(t) \tag{9}$$

The parameter $\bar{v}$ represents the linear advection speed, and it has a nominal value of $\bar{v} = 1$ from Section 7. In the sections that follow, we consider small perturbations to this parameter as we construct the filtered numerical scheme, and we introduce it here for the sake of consistency. Note that the parameters appearing in the initial and boundary condition above, ρ_0 and ρ_I represent dimensionless quantities that appear once the equation has been scaled appropriately.

For equation (7) the flux is $f(\rho) = \bar{v}\rho$, and the appropriate flux condition at the inflow boundary is given in terms of the density variable.

Also note that we have relabelled the notation returning to our original variable so that $\rho(x,t)$ represents the dimensionless density variable, and x and t are dimensionless in what follows. The model in (7) - (9) is convenient for expressing an analytical solution, testing simulation methods and analysing the impact of small perturbations on key measures. Each of these is carried out in the following section.

3.2. Analytical Solutions, Parameter Dependence and Perturbations

Equation (7) is a scalar, linear, constant-coefficient PDE that is of hyperbolic type. The general solution of this equation has characteristics in the form of $x(t) = \bar{v}t + x_0$. For the equation (7), we see that along $x(t)$ the time derivative of $\rho(x(t),t)$ is:

$$\frac{d}{dt}\rho(x(t),t) = \rho_t(x(t),t) + \rho_x(x(t),t)\frac{dx}{dt} \tag{10}$$
$$= \rho_t(x(t),t) + \bar{v}\rho_x(x(t),t) \tag{11}$$
$$= 0 \tag{12}$$

Thus $\rho(x,t)$ is constant along the characteristics. If we denote the initial and boundary conditions as $\rho(x,0) = \rho_0$ and $\rho(0,t) = \rho_I(t) = \rho_\ell$ respectively, where ρ_ℓ is a fixed parameter, then the analytical solution is given by,

$$\rho(x,t) = \begin{cases} \rho_\ell, & 0 < x < \bar{v}t \\ \rho_0, & \bar{v}t < x < 1. \end{cases} \tag{13}$$

We investigate perturbations in velocity and initial and boundary conditions, their effects on the solution and the accuracy of our numerical schemes in Section 6. Given a small perturbation in the advection velocity, say $\bar{v} = 1 + \epsilon$ where $0 \le \epsilon \le 0.5$, the analytical solution is constant along the characteristic lines of $x(t) = (1+\epsilon)t + x_0 = \bar{v}t + x_0$ and is described by (13).

We study perturbation in the inflow condition separately; that is, we fix the advection velocity at $\bar{v} = 1$ when the inflow parameter is perturbed. Let $0 \le \epsilon \le 0.1$ be a perturbation in the inflow ρ_ℓ, and let $\hat{\rho}$ denote the solution to the *perturbed* equation given by,

$$\hat{\rho}_t + \hat{\rho}_x = 0 \tag{14}$$
$$\hat{\rho}(x,0) = \rho_0 \tag{15}$$
$$\hat{\rho}(0,t)) = \rho_\ell \pm \epsilon \tag{16}$$

The characteristics lines are in the form of $x(t) = t + x_0$ and the analytical solution is formulated as,

$$\hat{\rho}(x,t) = \begin{cases} \rho_\ell \pm \epsilon, & 0 < x < t \\ \rho_0, & t < x < 1. \end{cases} \tag{17}$$

4. Numerical Schemes and Time Filter for the Model

Two methods for numerical simulation of the linear model are considered here. The first is the basic upwind scheme, and we introduce a time filter that is amended to the upwind scheme in order to improve accuracy. A description of the discretization is given, and both schemes are outlined. A stability condition for the filtered scheme is derived, and it is shown that the filtered scheme is guaranteed to converge when the appropriate CFL condition is satisfied.

Discretize the domain $[0,1] \times [0,T]$ using a mesh width $h = \Delta x$ and a time step $k = \Delta t$. This will define a collection of grid points by

$$x_j = jh, \;\; j = 0, 1, 2, \;\; \ldots \; M \tag{18}$$
$$t_n = nk, \;\; n = 0, 1, 2, \;\; \ldots \; N \tag{19}$$

For simplicity, we use a uniform mesh with h and k constants. We denote the true solution ρ evaluated at the mesh points as follows:

$$\rho_j^n = \rho(x_j, t_n),$$

and two algorithms are proposed for approximating these values at the grid points. The approximations are dentoed as P_j^n. That is,

$$P_j^n \approx \rho_j^n = \rho(x_j, t_n)$$

4.1. Upwind Scheme

The upwind scheme is a classical method that provides a simple choice for solving the linear advection equation. It is first order accurate, and the method is guaranteed to be stable and convergent provided that the CFL condition is satisfied. In terms of the notation presented here, the method is expressed as

$$P_j^{n+1} = P_j^n - \nu \left(P_j^n - P_{j-1}^n \right), \tag{20}$$

for $j = 1, 2, \ldots M$ and $n = 1, 2, \ldots N$. The parameter ν can be defined to simplify the statement of the method

$$\nu = \bar{v}\frac{\Delta t}{\Delta x},$$

and for a fixed choice of the parameters, the value of ν is constant. The upwind method is known to be stable when

$$0 \leq \nu \leq 1. \tag{21}$$

This is the well-known CFL condition, and it is necessary for this inequality to be satisfied in order for the numerical method to be convergent. [28, 29, 30, 31].

4.2. Upwind with Time Filter Scheme

This section introduces a scheme that combines a time filter with the fully explicit upwind method introduced in Section 4.1. This approach is inspired from the article [32] as well as early work applying a time filter within a PDE framework [33]. The idea as introduced in [32] is that the filter increases the accuracy of the scheme and reduces the discrete curvature of the solution simply by adding a single line of code. In addition, the scheme can be easily extended to a variable time step. In our application, the algorithm can be viewed as a two-step method where the unfiltered approximation at each time step is computed using the upwind method as step 1, and the approximation is then filtered in step 2. The filtered approximation becomes the next iterate in the scheme and is constructed as follows.

Step 1:

$$\frac{\tilde{P}_j^{n+1} - P_j^n}{\Delta t} + \bar{v}\frac{P_j^n - P_{j-1}^n}{\Delta x} = 0 \tag{22}$$

Step 2:

$$P_j^{n+1} = \tilde{P}_j^{n+1} - \frac{\gamma}{2}(\tilde{P}_j^{n+1} - 2P_j^n + P_j^{n-1}) \tag{23}$$

where $\tilde{P}^{n+1}$ and P^{n+1} stand for the unfiltered and filtered density at any x_j respectively. As shown in [32], the combination of backward Euler with the time filter in step 2 is 0-stable for $-2 \leq \gamma \leq 2$ and A-stable for $\frac{-2}{3} \leq \gamma \leq \frac{2}{3}$. Furthermore it is noted that the filtering process can be applied twice. In that case, the second step will consist of two filtering equations resulting in an increase value of the filter parameter γ to $\gamma(2 - \frac{\gamma}{2})$. The two filter equations can be combined and written as one single equation given below,

$$P_j^{n+1} = \tilde{P}_j^{n+1} - \frac{\gamma}{2}(2 - \frac{\gamma}{2})(\tilde{P}_j^{n+1} - 2P_j^n + P_j^{n-1}) \tag{24}$$

For the case of linear differential equations the two-step method for time filtering can be simplified into a single step method. Solving equation (23) for the filtered variable and substituting into the upwind scheme in equation (22) yields a one-step scheme. Below we show the obtained one-step scheme for the choice of filter parameter $\gamma = \frac{2}{3}$.

Algorithm 4.1. Upwind method with the time filter assuming a constant time step with $\gamma = \frac{2}{3}$. Given P_{j-1}^n, P_j^n and P_j^{n-1}, find P_j^{n+1} satisfying the following

$$P_j^{n+1} = \frac{2}{3}(2 - \nu)P_j^n - \frac{1}{3}P_j^{n-1} + \frac{2}{3}\nu P_{j-1}^n \tag{25}$$

Similarly, the two filtering equation in (24) provides the following single step method

$$P_j^{n+1} = \frac{4}{9}(\frac{7}{2} - \nu)P_j^n - \frac{5}{9}P_j^{n-1} + \frac{4}{9}\nu P_{j-1}^n. \tag{26}$$

Proof The equation in (25) is obtained by solving (23) for $\tilde{P}^{n+1}$ in step 2 and replacing it into the upwind scheme in step 1. Equation (26) is obtained in a similar fashion using (24).

4.3. Stability Condition for the Upwind with Time Filter

The upwind scheme in (20) is first order accurate, and the method is guaranteed to be stable and convergent provided that the CFL condition is satisfied [28, 29]. One can derive similar consistency and stability results for the upwind method that uses the time filter defined in Algorithm 4.1.

Proposition 4.1. The scheme (25) in Algorithm 4.1 is consistent of first order in both $\mathcal{O}(\Delta t)$ and $\mathcal{O}(\Delta x)$. When $0 < \nu \leq \frac{1}{2}$, the method is stable and convergent.

Proof

The local truncation error for Algorithm 4.1 is defined by

$$\tau^n = \frac{1}{\Delta t}\left(\mathcal{N}(\rho^n) - \rho^{n+1} \right)$$

where $\mathcal{N}(\rho)$ represents the right side of equation (25) evaluated at true solution ρ of (7).

$$\mathcal{N}(\rho_j^n) = \frac{2}{3}(2 - \nu)\rho_j^n - \frac{1}{3}\rho_j^{n-1} + \frac{2}{3}\nu\rho_{j-1}^n$$

with $\rho_j^n = \rho(x_j, t^n)$.

It follows that

$$\begin{aligned}\tau^n &= -\frac{2}{3}\left(\rho_t + \bar{v}\rho_x + \Delta t\rho_{tt} - \frac{1}{2\bar{v}}\rho_{xx}\Delta x + \mathcal{O}(\Delta t^2) \right) \\ &= -\frac{2}{3}\left(\Delta t\rho_{tt} - \frac{1}{2\bar{v}}\Delta x\rho_{xx} + \mathcal{O}(\Delta t^2) \right)\end{aligned}$$

Therefore, the local truncation error for the method is $\mathcal{O}(\Delta t)$, and the method is first order.

For stability of the method, we follow the Modified

Equation analysis in order to derive a CFL condition for the method, see Chapter 8 of [28]. Define a function $v(x,t)$ that agrees with P_j^n in (25) at the grid points in the mesh, where we identify $x = x_j$ and $t = t_n$. Then $u(x,t)$ satisfies

$$u(x,t+\Delta t) = \frac{2}{3}(2-\nu)u(x,t) - \frac{1}{3}u(x,t-\Delta t) + \frac{2}{3}\nu u(x-\Delta x,t)$$

Using Taylor expansions about the point (x,t), one arrives at

$$u_t + \bar{v}u_x = -\Delta t\ u_{tt} + \frac{1}{2}\bar{v}\Delta x\ u_{xx} + \mathcal{O}(\Delta x^2)$$

or

$$u_t + \bar{v}u_x = \frac{1}{2}\bar{v}\Delta x\ u_{xx} - \Delta t u_{tt} \tag{27}$$

if we neglect higher order terms. Differentiating (27) with respect to t, one can show that

$$u_{tt} = \bar{v}^2 u_{xx} + \mathcal{O}(\Delta t).$$

Neglecting the terms of order $\mathcal{O}(\Delta t)$, this expression can replace the term involving u_{tt} in (27) to obtain

$$\begin{aligned} u_t + \bar{v}u_x &= \frac{1}{2}\bar{v}\Delta x\ u_{xx} - \Delta t\ \bar{v}^2 u_{xx} \\ &= \bar{v}\Delta x\left(\frac{1}{2} - \nu\right)\ u_{xx} \end{aligned}$$

Hence the numerical solution for the scheme propagates with the correct advection speed $\bar{v}$, but it also exhibits some diffusion and smearing of the wavefront. If one chooses the parameter $\nu = \frac{1}{2}$, then the diffusion coefficient vanishes. Moreover, the diffusion coefficient is positive only when

$$0 < \nu < \frac{1}{2},$$

and this yields a stability condition for the scheme in Algorithm 4.1. Since the method is both consistent and stable when the time step Δt is chosen so that the CFL condition is satisfied, then by the Lax Equivalence Theorem, the method is convergent.

For the sample problems considered in this paper, the CFL condition reduces to imposing the inequality

$$\Delta t \leq \frac{1}{2\bar{v}}\Delta x.$$

For the results presented in the next section, the two cases arise where $\bar{v} = 1$ or $\bar{v} = 1 + \epsilon$. A similar stability and convergence conclusion holds for when the double filtering process is applied. The following statement provides the CFL condition for this particular case.

Proposition 4.2. The scheme (26) in Algorithm 4.1 is consistent of first order in both $\mathcal{O}(\Delta t)$ and $\mathcal{O}(\Delta x)$. When $0 < \nu \leq \frac{2}{7}$, the method is stable and convergent.

Proof

This result can be obtained easily following the same steps shown in Proposition 4.1.

5. Numerical Results

This section presents a few numerical illustrations including the stability and convergence properties for the numerical schemes introduced above. First we examine the CFL condition for both filtered upwind schemes when the filter is applied to the original linear model in equations (7)-(9) once and twice. Second an error analysis and convergence is shown for upwind, filtered upwind and double filter upwind schemes leading us to a proper time step for our calculations in the following sections. All the computations in this paper are carried out using the programming language Python. The results in this section are given for the case when $\rho_0 = 0$, and the inflow condition below. These conditions are chosen in order to compare the performance of the time filtered scheme with that of the standard upwind on a model problem that contains a jump discontinuity.

$$\rho_I(t) = \begin{cases} \rho_\ell, & 0 < t < 3 \\ 0, & t \geq 3 \end{cases} \tag{28}$$

5.1. Numerical Stability

The numerical results in this section convey our choice of time step considering a balance between stability and generating a smooth curve free of oscillations. As proved in Proposition 4.1, the CFL condition for filtered upwind in (25) is $0 < \nu \leq \frac{1}{2}$. Figure 2 shows that with the choice of $\Delta t = \frac{1}{2}\Delta x$, the solution surface in phase one exhibits oscillations on the top surface. As numerically tested decreasing the time step by about 20% to $\Delta t = \frac{2}{5}\Delta x$, the oscillations disappear resulting in a smooth surface. Similarly, when using a double time filter with the time step $\Delta t = \frac{2}{7}\Delta x$ based on the CFL condition shown in Proposition 4.2, we experience the presence of small oscillations on the top surface area. Reducing the time step to $\Delta t = 0.22\Delta x$ by about 23%, the surface becomes smooth, see Figure 3.

5.2. Numerical Error and Convergence

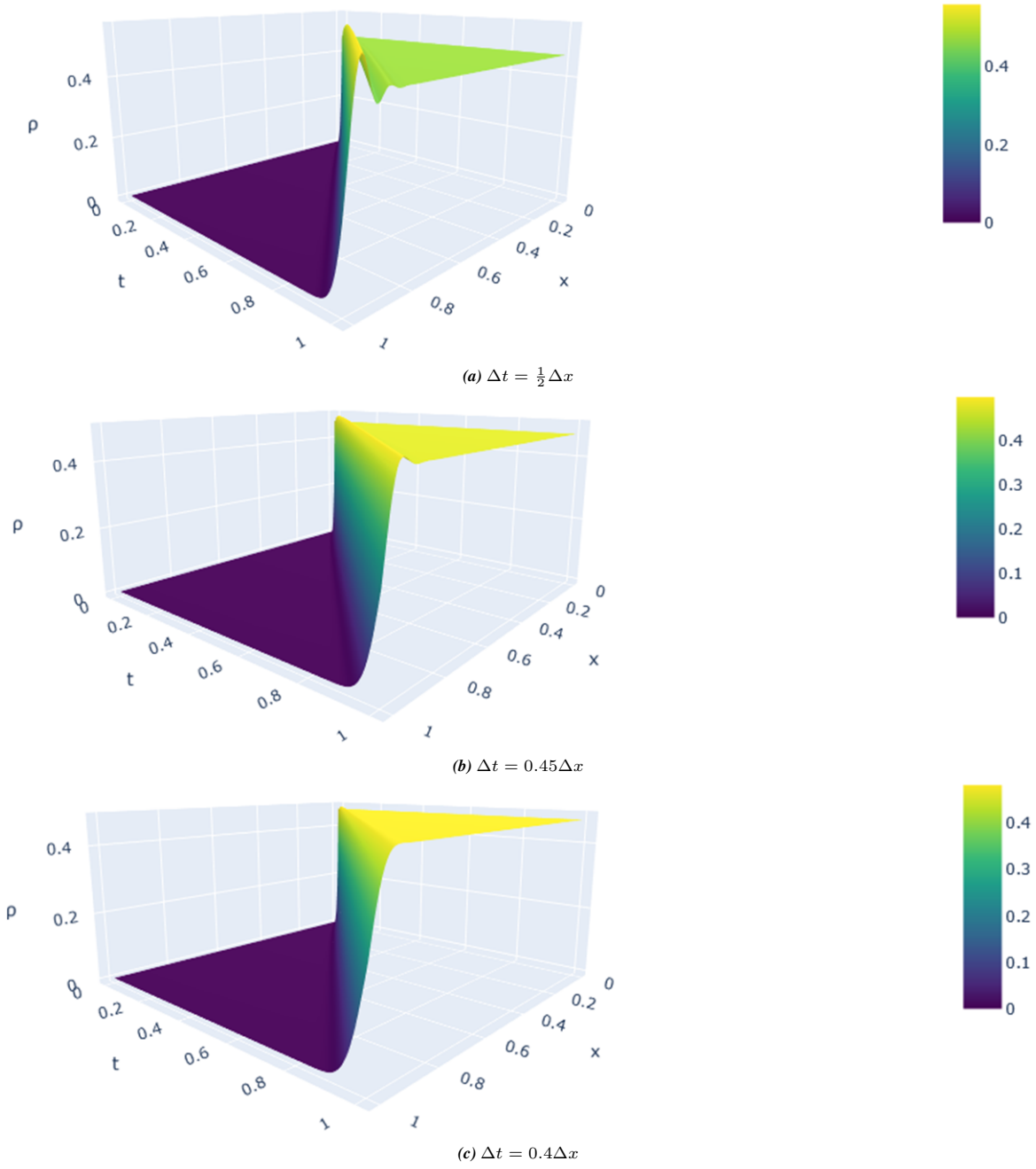
In order to estimate accuracy and demonstrate convergence in this section, we examine the computational error for approximated density using upwind, upwind with one time filter and double time filter over the finite time interval $[0,1]$ or final time $T = 1$. Denote the error at the grid point (x_j, t_n) as $E_j^n = \rho_j^n - P_j^n$, and our choice for the error norm is the L^1-norm defined as $||E^n|| = \Delta x \sum_{j=0}^{M} |E_j^n|$ for $n = 0, 1, 2, \ \ldots\ , N$. In the error calculations for (20) and (25), we assume a fixed ratio $\nu = \frac{\Delta t}{\Delta x} = \frac{2}{5}$ for the case of $\bar{v} = 1$ as discussed in the previous section for a numerical solution free of oscillations. This ratio is taken equal to $\frac{1}{5}$ for computations using a double time filter introduced in equation (26). Letting $\Delta t \to 0$ by refining the spatial grid, i.e. Δx, the error calculations at different times within the interval $[0,1]$ are given in Tables 1 and 2. As one observes at each fixed time , i.e. $t = 0.2, 0.4, 0.6, 0.8$ and 1, the errors using all three schemes become smaller converging to zero as $\Delta x \to 0$. Comparing

the errors in Table 1 for the approximated density, P_j^n by upwind and upwind with time filter, the latter scheme shows a better accuracy at all computed times. The errors in Table 2 for the density approximations via the double time filter in equation (26) presents a close accuracy to the upwind with one time filter; however, it requires a finer time step. Therefore it is not viewed to be as computationally efficient as upwind with one time filter.

Calculations to estimate the rate of convergence at $t = 1$ for all the schemes are listed in Table 3. Given a grid size defined as $h_i = 2^{-(i+5)}$ for $i = 1, 2, 3, \ldots, 6$, the convergence rate of the error is determined by calculating the exponent a in the following expression,

$$\left(\frac{h_i}{h_{i+1}}\right)^a = \frac{||E(h_i)||}{||E(h_{i+1})||} \tag{29}$$

The convergence rate values indicate that the 1-norm of the errors decay at the rate of about $\frac{1}{2}$ in all cases. This means that error decreases like $(\Delta x)^{\frac{1}{2}}$. Although the methods are all first order based on the local truncation error analysis in Section 5.3, these results are valid for smooth solutions only. One might expect a lower rate of convergence for discontinuous solutions as indicated in Section 8.7 of [28].

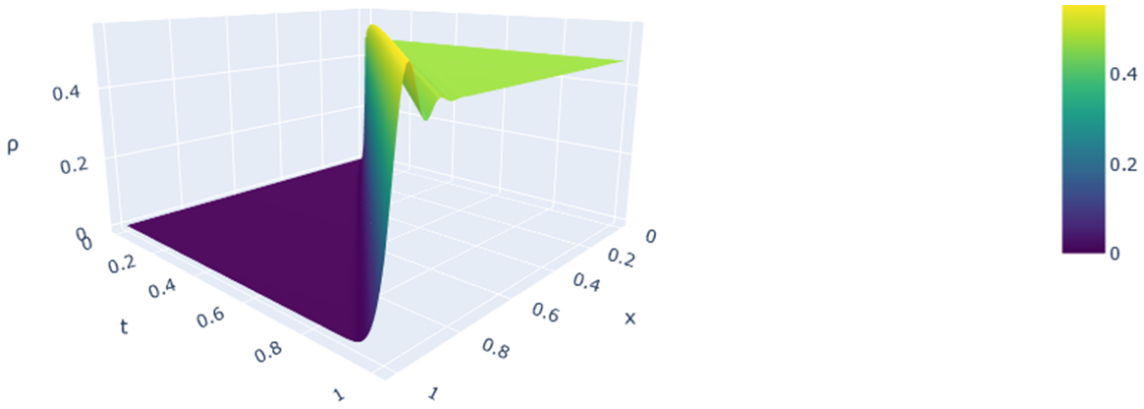


(a) $\Delta t = \frac{1}{2}\Delta x$

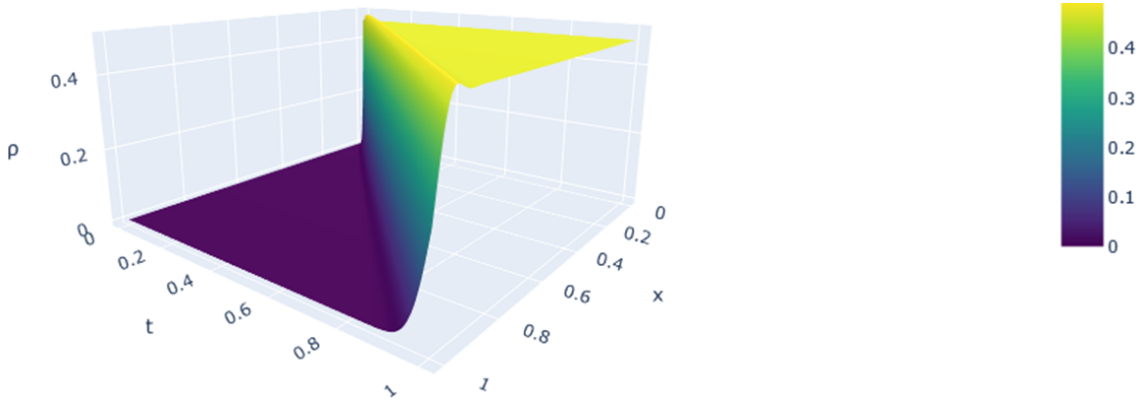
(b) $\Delta t = 0.45\Delta x$

(c) $\Delta t = 0.4\Delta x$

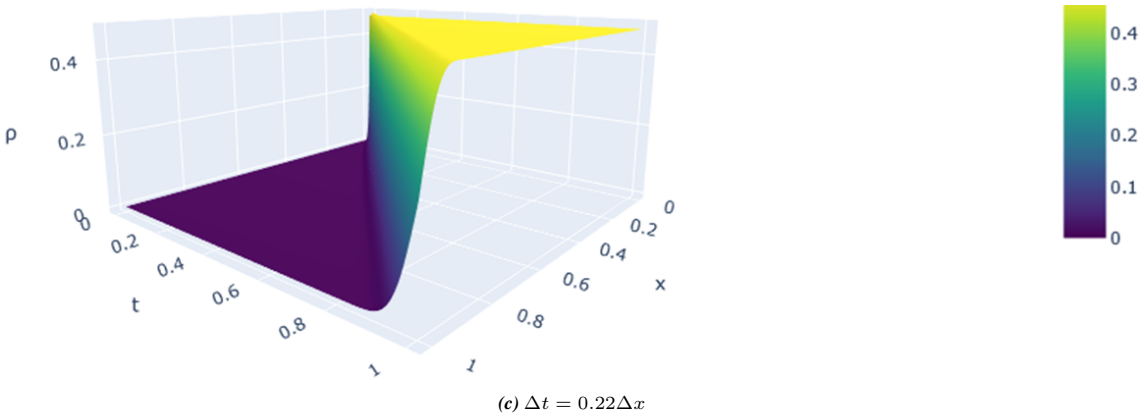
Figure 2. *The density surface plot for first phase using upwind with one time filter in equation (25) of Algorithm 4.1 and different choices of* Δt.



(a) $\Delta t = \frac{2}{7}\Delta x$



(b) $\Delta t = 0.25\Delta x$



(c) $\Delta t = 0.22\Delta x$

Figure 3. *The density surface plot for first phase using upwind with double time filter in equation (26) of Algorithm 4.1 and different choices of* Δt.

Table 1. *The density approximation for upwind using* 1*-norm with* $\Delta t = \frac{2}{5}\Delta x$.

			L^1 **Errors using Upwind**		
Δ x	**t=0.2**	**t=0.4**	**t=0.6**	**t=0.8**	**t=1**
2^{-6}	$1.67227e^{-2}$	$2.31087e^{-2}$	$2.45148e^{-2}$	$3.22691e^{-2}$	$1.46734e^{-2}$
2^{-7}	$1.15544e^{-2}$	$1.62118e^{-2}$	$1.98373e^{-2}$	$2.29609e^{-2}$	$1.10689e^{-2}$
2^{-8}	$8.10591e^{-3}$	$1.14828e^{-2}$	$1.40368e^{-2}$	$1.62686e^{-2}$	$8.1833e^{-3}$
2^{-9}	$5.74139e^{-3}$	$8.13431e^{-3}$	$9.94978e^{-3}$	$1.14867e^{-2}$	$5.96816e^{-3}$
2^{-10}	$4.06715e^{-3}$	$5.74334e^{-3}$	$6.80282e^{-3}$	$8.11829e^{-3}$	$4.31221e^{-3}$
2^{-11}	$2.87167e^{-3}$	$4.05914e^{-3}$	$4.97113e^{-3}$	$5.74109e^{-3}$	$3.09567e^{-3}$

Table 2. *The density approximation for upwind with filter using* 1*-norm with* $\Delta t = \frac{2}{5}\Delta x$.

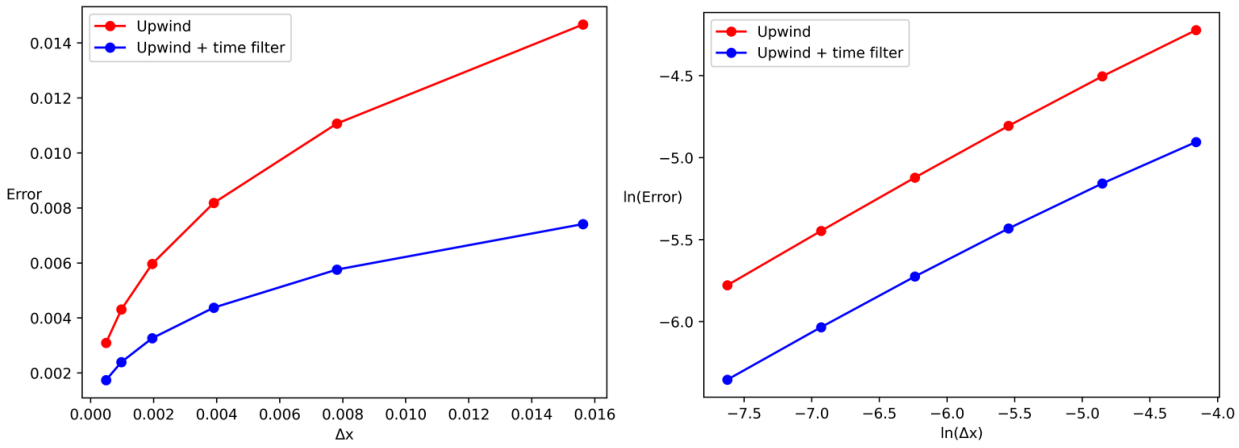
L^1 Errors using Upwind with Filter					
Δ x	**t=0.2**	**t=0.4**	**t=0.6**	**t=0.8**	**t=1**
2^{-6}	$1.08676e^{-2}$	$1.42594e^{-2}$	$1.35085e^{-2}$	$1.93542e^{-2}$	$7.41581e^{-3}$
2^{-7}	$7.12969e^{-3}$	$9.67841e^{-3}$	$1.17014e^{-2}$	$1.346341e^{-2}$	$5.75938e^{-3}$
2^{-8}	$4.83920e^{-3}$	$6.73171e^{-3}$	$8.18264e^{-3}$	$9.46543e^{-3}$	$4.37694e^{-3}$
2^{-9}	$3.36585e^{-3}$	$4.73276e^{-3}$	$5.77234e^{-3}$	$6.65238e^{-3}$	$3.26458e^{-3}$
2^{-10}	$2.36636e^{-3}$	$3.32619e^{-3}$	$3.84139e^{-3}$	$4.69451e^{-3}$	$2.39717e^{-3}$
2^{-11}	$1.66309e^{-3}$	$2.34726e^{-3}$	$2.87294e^{-3}$	$3.31711e^{-3}$	$1.74039e^{-3}$

Table 3. *The density approximation error for upwind with twice time filter using 1-norm and* $\Delta t = \frac{1}{5}\Delta x$.

L^1 Errors using Upwind with Double Filter					
Δ x	**t=0.2**	**t=0.4**	**t=0.6**	**t=0.8**	**t=1**
2^{-6}	$1.23805e^{-2}$	$1.66519e^{-2}$	$2.00012e^{-2}$	$2.30880e^{-2}$	$9.45654e^{-3}$
2^{-7}	$8.32595e^{-3}$	$1.15513e^{-2}$	$1.41535e^{-2}$	$1.62979e^{-2}$	$7.34205e^{-3}$
2^{-8}	$5.77565e^{-3}$	$8.14898e^{-3}$	$9.96176e^{-3}$	$1.15347e^{-2}$	$5.5362e^{-3}$
2^{-9}	$4.07449e^{-3}$	$5.76734e^{-3}$	$6.59036e^{-3}$	$8.13089e^{-3}$	$4.09231e^{-3}$
2^{-10}	$2.88367e^{-3}$	$4.06544e^{-3}$	$4.97358e^{-3}$	$5.74322e^{-3}$	$2.98432e^{-3}$
2^{-11}	$2.03272e^{-3}$	$2.87161e^{-3}$	$3.5171e^{-3}$	$4.06051e^{-3}$	$2.15619e^{-3}$

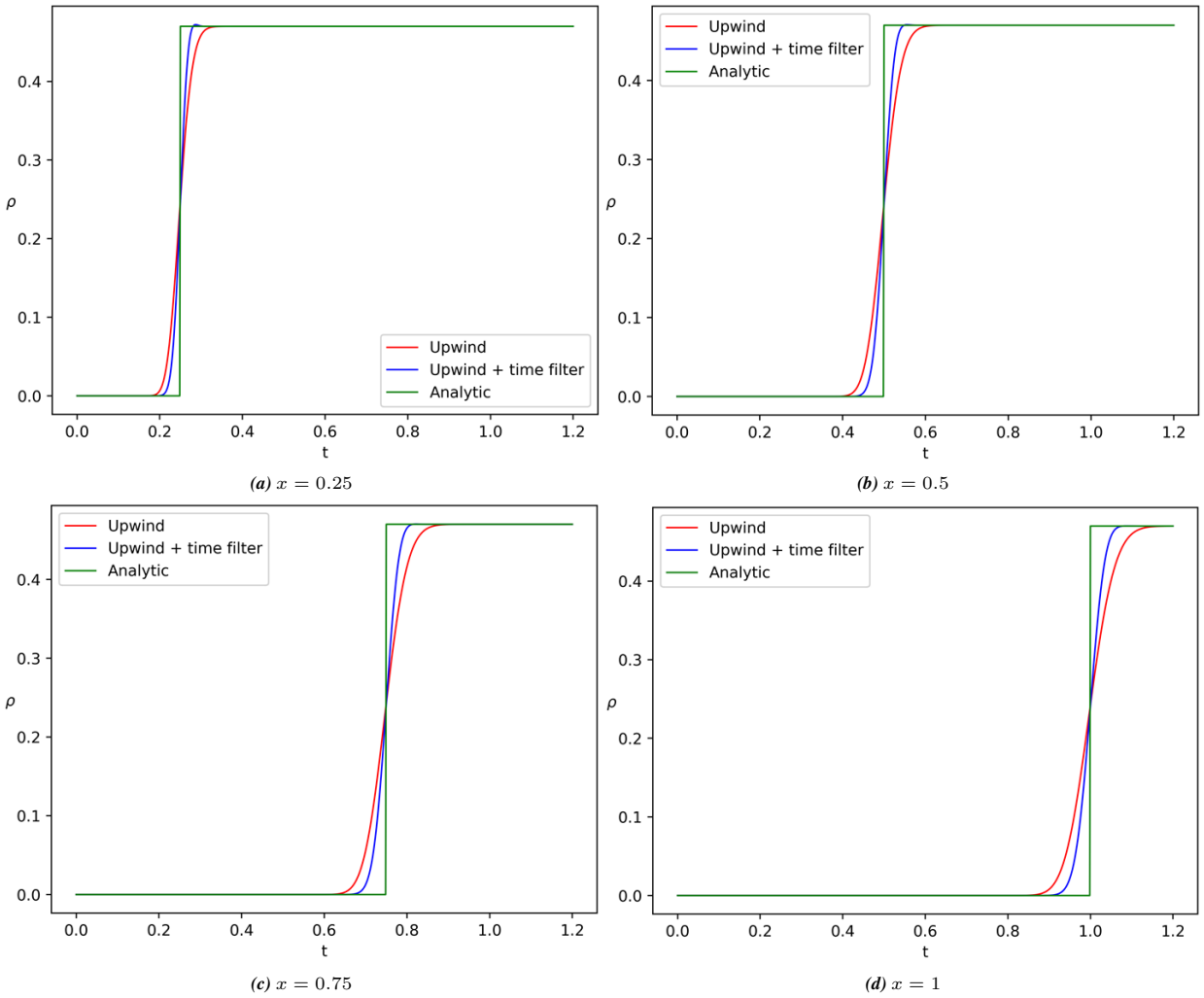
Table 4. *The convergence rate estimate for upwind with double time filter, filtered once and upwind at* $t = 1$ *according to the error values in Tables 1 and 2 using the values of* i *to estimate the value of* α *in equation (29).*

Rate of Convergence			
i	**Double Filter**	**Filtered**	**Upwind**
1	0.365129996	0.364690772	0.406698343
2	0.407270352	0.395991797	0.435754044
3	0.435998053	0.423025599	0.455396211
4	0.455513924	0.445564016	0.468859993
5	0.468917266	0.461922391	0.478176194



(a) *Error plots at* $t = 1$ ***(b)*** *Log-log plots at* $t = 1$

Figure 4. *The* L^1*-error and log-log plots for the filtered and upwind schemes at* $t = 1$ *with* $\Delta t = \frac{2}{5}\Delta x$.



(a) $x = 0.25$ ***(b)*** $x = 0.5$

(c) $x = 0.75$ ***(d)*** $x = 1$

Figure 5. *The density plot for* $0 < t < 1$ *at* $x = 0.25, 0.5, 0.75$ *and* 1 *respectively with the choice of* $\Delta x = \frac{1}{128}$ *and* $\Delta t = \frac{2}{5}\Delta x$.

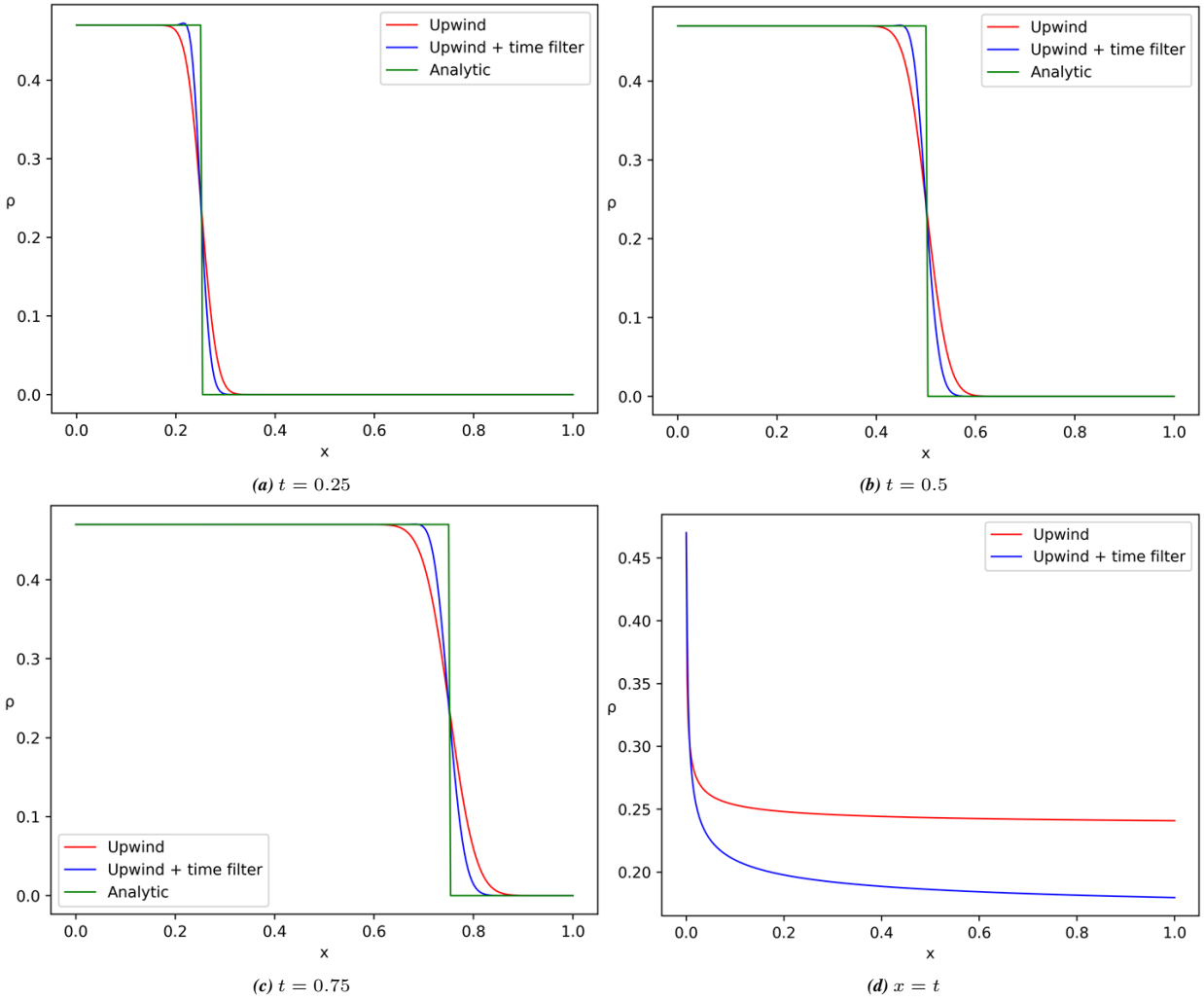
Figure 4 shows a graphical comparison for the error between the density approximation using the upwind scheme and the filtered upwind scheme at time $t = 1$ as well as a graphical comparison for the rate of convergence using a log-log plot. Although both schemes have the same order of convergence as shown in the log-log plot, the error plot in 4a shows that the filtered upwind scheme exhibits a significant reduction in error over that of the upwind scheme for a coarse mesh size. Hence the filtered scheme has advantages when used in a broader ensemble model simulation framework.

The numerical solutions obtained by upwind and filtered upwind versus the analytical solution over the time interval $[0, 1]$ are demonstrated in figures 5 and 6. These figures demonstrate a comparison of the numerical solutions as functions of one variable at a time. In Figure 5, the approximated density is calculated at different locations as a function of time and in Figure 6 (a)-(c), it is shown at different times as a function of space. Figure 6 (d) illustrates the numerical solutions along the characteristic line $x = t$. As one can observe, the upwind with time filter drops down to zero faster than upwind thus capturing the jump discontinuity quicker. All the plots present a better accuracy of the numerical solution via filtered upwind. Overall our numerical results is in favor of using the time filter. Although all the numerical schemes that we have tested share the same rate of convergence, the calculated errors speak of a smaller error and therefore a better estimation when a time filter is applied.

6. Parameter Analysis for the Biological Model

In this section, parameters related to the physical properties of the *rrn operon* are used to connect the output of the PDE model simulations given in Section 5 with the original transcription model. We focus on the estimation of the average transcription time (ATT) for an RNA polymerase, an important biological measure of cell fitness. An accurate estimate of the ATT is crucial to understanding a full ribosome assembly model as it describes the amount of time required to produce a single strand of rRNA, a fundamental

building block of the ribosome. Results in this section illustrate that the filtered algorithm is superior to upwind when approximating the ATT, and the effects of perturbations on various parameters are explored in this context. Perturbations of the elongation velocity are considered first as the velocity is key to determining the average transcription time of individual RNAPs. The effects of perturbations in the initiation rate of RNAPs at the promoter site are also discussed.



(a) $t = 0.25$ *(b)* $t = 0.5$

(c) $t = 0.75$ *(d)* $x = t$

Figure 6. *Figures (a)-(c) are the density plots for* $0 < x < 1$ *at* $t = 0.25, 0.5$ *and* 0.75 *respectively and figure (c) is the density plots along the characteristic line* $x = t$ *with the choice of* $\Delta x = \frac{1}{128}$ *and* $\Delta t = \frac{2}{5}\Delta x$.

The results contained in this section are based on gene information gathered experimentally and reported in the paper [34]. In particular, the length of the *rrn* operon is $L = 5500$nts. The average density of RNAPs on the gene is reported as $\rho_\ell = 51.5$ RNAPs per strand, and we refer to this average value as the *equilibrium density* for the purposes of the continuum model considered in this paper. The maximum density of RNAPs on the strand is given by $\rho_{max} = 110$ RNAPs per strand. The average initiation rate, or the rate at which RNAP molecules initiate onto the strand, is assumed to be 0.86 RNAP molecules per second. The parameter $v_{max} = 91$nts per second is the average elongation rate of the RNAP, and when it is converted to the units of strand per second, then it is given as below.

$$v_{max} = \frac{91}{5500} \approx 0.016545 \text{ strand per second}$$

While v_{max} describes the average elongation rate, it is known that the elongation rate of individual RNAPs can vary, and a range for the elongation rate is reported as 70 - 200nts per second, see page 3744 of [34]. We consider perturbations in the velocity of the linear model problem once it has been nondimensionalized. For the simplest case, we do not consider a time dependent velocity but simply a constant velocity where the parameter is perturbed. That is, $\bar{v} = 1 + \epsilon$ in equation (7). Taking into account the rescaling in Section 7, the perturbations correspond to values of average elongation rate

that range from 90nts per second to 136nts per second. In the results that follow, we report on only the positive perturbations of the elongation velocity. Similar results were obtained with negative perturbations of the velocity, but we omit those for brevity.

The initial and boundary conditions for the linear models (unperturbed and perturbed) are imposed so that at $t = 0$, the strand is empty, that is, $\rho_0 = 0$ in (7). The total time that the model is simulated is partitioned into three phases. These correspond to a *ramp-up (to equilibrium density)* phase and a *ramp-down* phase separated by an *equilibrium* phase. We refer to these in the following way. Phase I is the time required for the density across the entire strand to ramp up to the equilibrium density, ρ_ℓ, after the simulation begins. Phase II corresponds to the amount of time the density, $\rho(x,t)$ remains at the equilibrium density value. During the simulation at some time after Phase II is established, the inflow condition is set to 0 so that RNAP initiation is no longer permitted. For the results presented here, we chose this time to be at $t = 3$ units of dimensionless time. In terms of the rescaling outlined in Section 7, this corresponds to an inflow boundary condition of the form in equation (28). Phase III corresponds to the amount of time required for the entire strand to be emptied of RNAPs after the inflow condition is set to 0. For all calculations in this work, the parameter $\rho_\ell = 0.47$. Figure 7 shows the solution surface for all the three phases using the analytical solution.

Note that the time durations of all phases are analytically computable for the linear model, and the values are controlled by specification of initial and boundary conditions, along with the perturbation in the velocity in (7). The time duration of Phases I and II depend on the elongation velocity, specifically whether it is perturbed by ϵ or not. The time duration of Phase III is exactly the same value as that of Phase I. In the following sections, we give the analytical expressions of those values, relate them to the original biological model outlined above and report on the numerical simulations' ability to accurately estimate these time durations.

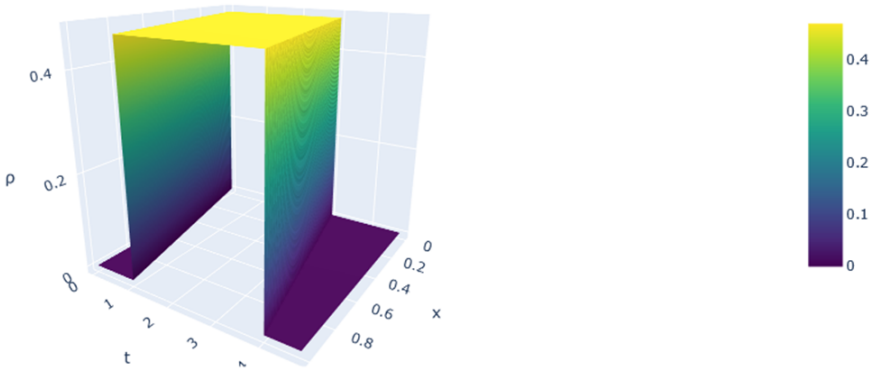


Figure 7. *The density surface plot for* $0 \leq t \leq 4$ *and* $0 \leq x \leq 1$.

6.1. Perturbations in the Average Elongation Velocity

For the linear PDE model, the ATT for each RNAP is constant, and it is determined by the flow velocity parameter. Another consequence of the linear PDE regime is that the length of Phase I is equal to that of the ATT. The characteristics of the governing PDE are all one needs to consider in order to analytically compute the ATT. Perturbations in the average elongation velocity directly influence the value, and here we give the analytical value based on the nature of the perturbation in the velocity term. With regards to Phase I, the ATT is also the time required for the density across the strand to ramp up to the equilibrium density after the simulation begins at $t = 0$. The ATT is easily computed analytically using a sketch of characteristic lines.

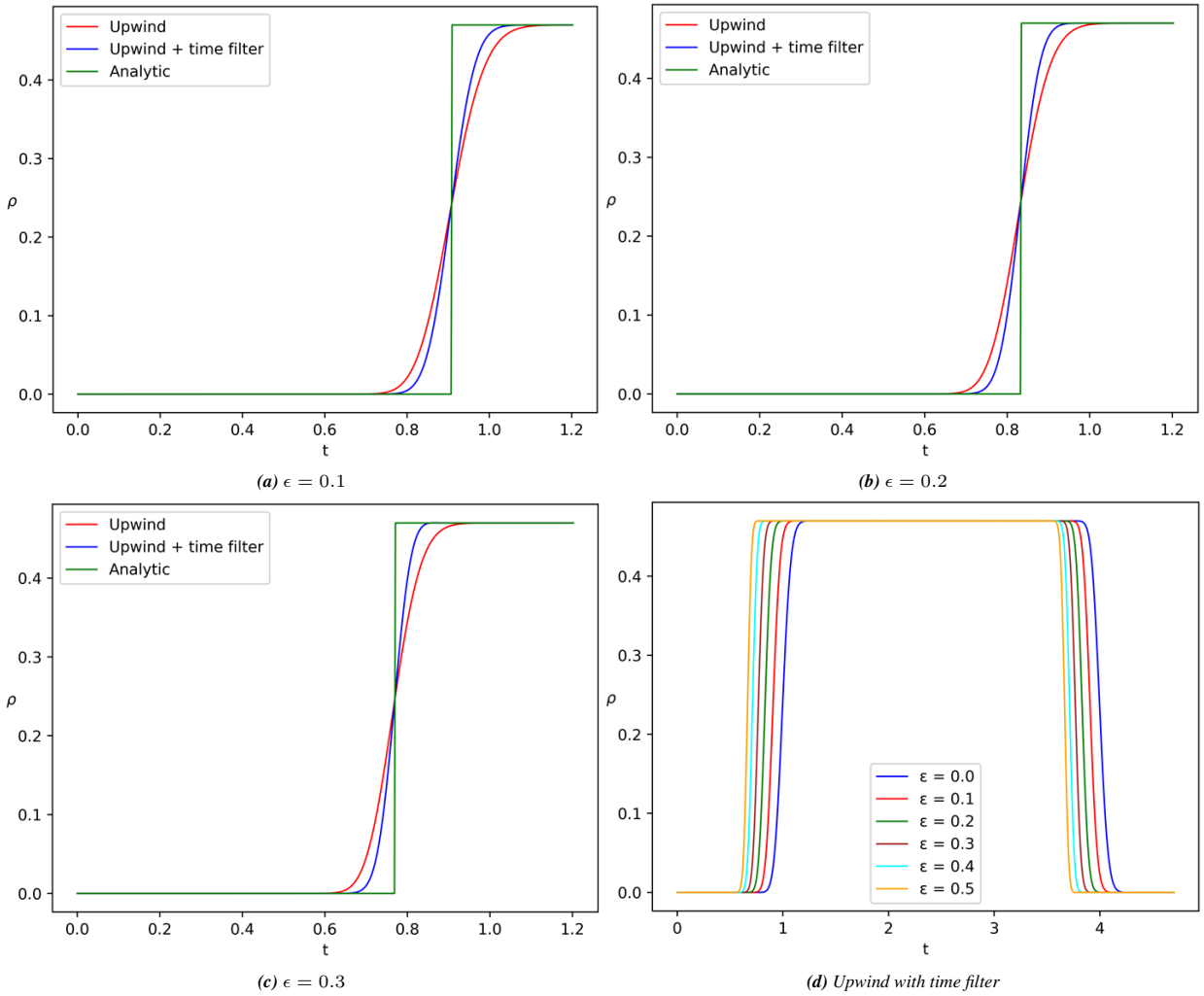
With $\epsilon = 0$ and $\bar{v} = 1$ in the governing PDE (7) - (9), the time to reach equilibrium density from the beginning of the simulation is one non-dimensional unit of time. This value is labelled $t^\star = 1$ in the discussion that follows. If the velocity is perturbed as $\bar{v} = 1 + \epsilon$ in (7) - (9), then we denote the time to reach equilibrium density as t^+, which is given by

$$t^+ = \frac{1}{1+\epsilon}$$

One can explore the accuracy with which each numerical scheme is able to estimate the ATT. In order to estimate this value, one finds the first discrete time value t_K when $P_M^K - 0.46 \geq 0$, where $x_M = 1.0$. In order to assess the scheme's ability to estimate the ATT in the presence of velocity perturbations, this is computed for several values of ϵ. Table 5 gives the analytical values of the time duration of Phase I along with the errors in the approximations calculated using both the filtered scheme in Algorithm 4.1 and the upwind scheme described in equation (20). The filtered scheme is consistently more accurate than the upwind scheme, reducing the error by a factor of about $1/3$. It is notable that the errors for both schemes are smaller for larger values of the perturbation parameter ϵ, and this is expected since the results in Figure 8 illustrate that both numerical schemes more accurately estimate the jump in the function $\rho(1,t)$ for larger perturbations. The wave travels more quickly across the domain, and the approximation schemes have less time for the discontinuity to smear.

Table 5. *Analytical and numerical estimates for the duration time of Phases I and III. The second column gives the dimensionless value of the time duration based on the analytical expressions above. The columns labelled Filtered report the error in the numerical estimate of the duration time using the filtered upwind algorithm from Algorithm 4.1. Those columns labelled Upwind give the error in numerical estimates of the duration time using only the upwind scheme in (20) without filtering. The grid size for all calculations presented in the table is* $\Delta x = \frac{1}{128}$, *and* $\Delta t = 0.3\Delta x$. *Note that this choice for* Δt *satisfies the CFL condition in Proposition 4.1 for all choices of* ϵ *in Column 1.*

		Estimation Errors		Estimation Errors	
	Phase I, III	**Phase I**	**Phase I**	**Phase III**	**Phase III**
0	$t^\star = 1.0$	0.114	0.157	0.044	0.056
0.1	$t^+ = 0.909$	0.093	0.138	0.037	0.051
0.2	$t^+ = 0.833$	0.077	0.124	0.030	0.044
0.3	$t^+ = 0.769$	0.058	0.11	0.023	0.039
0.4	$t^+ = 0.714$	0.044	0.100	0.021	0.037
0.5	$t^+ = 0.667$	0.030	0.091	0.016	0.032



(a) $\epsilon = 0.1$ ***(b)*** $\epsilon = 0.2$

(c) $\epsilon = 0.3$ ***(d)*** *Upwind with time filter*

Figure 8. *The density plot along* $x = 1$ *for perturbed velocity by* ϵ. *Plots (a), (b), and (c) show the comparison between upwind and upwind with time filter with the perturbed velocities. Plot (d) presents the approximated density using upwind with time filter for all velocity perturbations over the entire process including all the phases. The grid size for all calculations presented in the table is* $\Delta x = \frac{1}{128}$, *and* $\Delta t = 0.3\Delta x$. *Note that this choice for* Δt *satisfies the CFL condition in Proposition 4.1 for all choices of* ϵ *in Column 1.*

One can relate the range of perturbations in ϵ to that of the original dimensioned model in (4)-(6) to assess their effects on the corresponding perturbations in average transcription time for the *rrn* gene. We explore this for Phase I in Table 6 below. The numerical algorithms show similar behavior in their predictions of the time duration for Phase III.

Table 6 shows the average elongation rate and the corresponding ATT in terms of the units associated with the original dimensioned biological model for the range of perturbations in ϵ. Both the filtered upwind and the original upwind scheme consistently overestimate the ATT; however, the filtered scheme arrives at an estimate which is consistently overestimating by about 10% while the upwind

scheme overestimates by about 15%. This is also consistent with the graphs shown in Figure 8 where we see the density approximations using the filtered scheme exhibit more accuracy close to the jump in the density at the point $x = 1$ as a function of the time variable. Hence the filtered upwind scheme allows one to better estimate the ATT in the linear regime.

Table 6. *Analytical and numerically estimated values for duration time of Phase I in terms of the dimensioned time for the original biological model. The second column gives the dimensioned value of the average transcription time using the time scaling in (36).*

	Dimensioned Values		Avg. TT Estimation	
ϵ	**Elong. Rate**	**Phase I Avg.TT**	**Filtered Upwind**	**Upwind Estimate**
0	**91** (nt/sec)	60.44**s**	67.39**s**	69.93**s**
0.1	100.1	49.95s	55.11s	57.53s
0.2	109.2	42.03s	45.88s	48.25s
0.3	118.3	35.85s	38.54s	40.96s
0.4	127.4	30.82s	32.77s	35.18s
0.5	136.5	26.92s	28.12s	30.58s

6.2. Perturbations in the Initiation Rate

In this section, we consider small perturbations in the inflow rate for (7) - (9), and the elongation velocity is fixed at $\bar{v} = 1$. The parameter $\rho_\ell = 0.47 \pm \epsilon$ is perturbed, and this corresponds to small perturbations in the initiation rate of the biological model. Note that the length of Phase I is then $t^\star = 1$, and the density along the right boundary at $x = 1$ should attain the value $\rho_\ell = 0.47 \pm \epsilon$ at that time. The beginning of Phase II occurs at time $t^\star = 1$ and continues until $t = 3$ for the simulations. Here we report how closely numerical schemes reach the analytical value of the density $\rho_\ell = 0.47 \pm \epsilon$ for various values of ϵ. Table 7 illustrates that both methods perform equally well for density approximations under perturbations in ρ_ℓ. The differences between the analytic value of the density and the approximations are on the order of the grid size. As noted, the ATT is independent of any perturbation in the parameter ρ_ℓ, and we note that the results in Table 7 indicates that the calculations of ATT done in this case with the filtered scheme are independent of the size of ϵ for this type of perturbation. The upwind scheme shows some dependence of the ATT for perturbations in the inflow parameter. In addition, the ATT estimations by the filtered scheme presents about 50% better accuracy compared to that of the traditional upwind for all the ϵ values using the same grid size. It is notable that the filtered scheme delivers increased accuracy over the range of perturbations and with the same computational overhead.

Table 7. *The density values and average transcription time approximated by filtered upwind and upwind along line $x = 1$ with perturbed inflow for $\epsilon = \pm 0.01, \pm 0.025, \pm 0.05, \pm 0.075, \pm 0.1$ and the grid size of $\Delta x = \frac{1}{128}$ and $\Delta t = \frac{2}{5}\Delta x$. The analytic values are given for comparison.*

	Density	Density Estimation		Avg TT Estimation	
ϵ	**Analytic**	**Filtered**	**Upwind**	**Filtered**	**Upwind**
-0.1	0.37	0.3620	0.3605	1.0719	1.1375
-0.075	0.395	0.3865	0.3858	1.0719	1.1406
-0.05	0.42	0.4109	0.4102	1.0719	1.1406
-0.025	0.445	0.4354	0.4356	1.0719	1.1438
-0.01	0.46	0.4501	0.4502	1.0719	1.1438
0.01	0.48	0.4721	0.4708	1.0750	1.1469
0.025	0.495	0.4868	0.4855	1.0750	1.1469
0.05	0.52	0.5114	0.5109	1.0750	1.1500
0.075	0.545	0.5360	0.5355	1.0750	1.1500
0.1	0.57	0.5606	0.5601	1.0750	1.1500

7. Conclusions and Future Work

Motivated by the LWR traffic flow model, this work gives the derivation of a simple continuum model for RNAPs transcribing a gene on a DNA strand. After nondimensionalizing, the linear model describes the density of RNAPs evolving as a function of space and time within a velocity field. Our initial study considers the simple case of a constant velocity. The paper illustrates the advantage of combining a time filter with the standard upwind in all the numerical computations, and a CFL condition is derived for the cases of filtering once and twice. The filtering technique shows an increased accuracy and a reduced numerical dissipation along the jump discontinuity compared

to the upwind. The improved estimation using a filtered scheme is consistent throughout the parameter analysis studies where we perturb the average elongation velocity and the initiation rate and illustrate their impact on the average transcription time for the model.

Future work will include consideration of a spatially dependent, piecewise defined elongation velocity function that is parametrized to allow for local variation in RNAP elongation rates. A long term research goal is to combine a version of this compartment model with a nonlinear system of ordinary differential equations as well as other compartments to construct a comprehensive model of ribosome assembly. The results contained here provide a positive initial step towards that goal.

Acknowledgements

The authors would like to acknowledge the support of the National Science Foundation under Award DMS-1951510. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The authors offer special thanks to Prof. Zafer Hatahet, Penn State University-Abington, for his continued support providing lectures, consultations and many useful references on *rrn* operon polymerization process. The authors also acknowledge Prof.T.Gedeon, Montana State University, for the many helpful conversations about ribosomal assembly mechanisms and for the ongoing collaboration.

Appendix: Nondimensionalization for the Linear Problem

The simplest PDE model is the linear model. The velocity is constant $v(\rho) = v_{max}$. Using this expression for velocity in the general conservation law form in equation (1) leads to a flux function that is defined to be

$$f(\rho) = v_{max}\ \rho,$$

and the linear PDE is given by

$$\frac{\partial \rho}{\partial T} + \frac{\partial}{\partial x}\left(v_{max}\rho\right) = 0, \quad 0 < x < 1, T > 0 \tag{30}$$

with a boundary condition and an initial condition given by

$$\begin{aligned} \rho(x,0) &= \rho_0 \\ f(\rho(0,T)) &= v_{max}\rho(0,T) = f_\ell \end{aligned} \tag{31}$$

The spatial variable is already scaled so that it corresponds to the dimensionless length of the gene on the DNA strand; $x = 1$ corresponds to a length of 5500 nucleotides for the gene under consideration in this work. The time variable T is measured in seconds. The units associated with ρ are in the number of RNAPs per strand. Hence, we first scale the density variable so that it is unitless. Define,

$$\hat{\rho}(x,T) = \frac{1}{\rho_{max}}\rho(x,T)$$

Multiplying (30) by $\frac{1}{\rho_{max}}$, we have the PDE

$$\frac{\partial}{\partial T}\left(\frac{\rho}{\rho_{max}}\right) + \frac{\partial}{\partial x}\left(v_{max}\left(\frac{\rho}{\rho_{max}}\right)\right) = 0 \tag{32}$$

$$\frac{\partial}{\partial T}\left(\hat{\rho}\right) + \frac{\partial}{\partial x}\left(v_{max}\hat{\rho}\right) = 0 \tag{33}$$

with a boundary condition and an initial condition given by

$$\hat{\rho}(x,0) = \frac{\rho(x,0)}{\rho_{max}} = \frac{\rho_0}{\rho_{max}} \tag{34}$$

$$\hat{f}(\hat{\rho}(0,T)) = v_{max}\hat{\rho}(0,T) = v_{max}\frac{\rho(0,T)}{\rho_{max}} = \frac{f_\ell}{\rho_{max}} \tag{35}$$

Finally, we rescale time so that the dimensionless time variable is

$$t = v_{max}T \ \text{ so that } [t] = [v_{max}T] = \frac{strand}{sec}\cdot sec = 1 \tag{36}$$

Based on the dimensionless independent variables, we define a new variable to represent the concentration $z(x,t) = \hat{\rho}(x,T(t)) = \hat{\rho}(x,T)$. Given an arbitrary function $g(x,t) = \hat{g}(x,T(t))$, we note that the partial derivatives are related in the following way.

$$g_t(x,t) = \frac{\partial \hat{g}(x,T)}{\partial T}\ \frac{dT}{dt}$$

This leads to

$$z_t = \frac{1}{v_{max}}\frac{\partial \hat{\rho}}{\partial T} \quad \Rightarrow \quad v_{max}z_t = \frac{\partial \hat{\rho}}{\partial T} \quad \text{and} \quad z_x = \hat{\rho}_x$$

Substituting this into equation (33) above, we have

$$\begin{aligned} \frac{\partial}{\partial t}\left(v_{max}z\right) + \frac{\partial}{\partial x}\left(v_{max}z\right) &= 0 \\ z_t + z_x &= 0 \end{aligned}$$

with boundary and initial condition as follows:

$$z(x,0) = \frac{\rho_0}{\rho_{max}} \tag{37}$$

$$f((0,t)) = z(0,t) = \frac{f_\ell}{v_{max}\rho_{max}} \tag{38}$$

Note that the inflow boundary is computed carefully here. Since $f(z,t) = z$, the appropriate flux condition at the inflow boundary must use the last equation in (35) to obtain

$$f(0,t) = z(0,t) = \hat{\rho}(0,t) = \frac{\rho(0,t)}{\rho_{max}} = \frac{1}{v_{max}}\frac{f_\ell}{\rho_{max}}$$

In summary, the dimensionless form of the linear PDE

problem is given below.

$$z_t + z_x = 0 \tag{39}$$

$$z(x,0) = \frac{\rho_0}{\rho_{max}} \tag{40}$$

$$f((0,t)) = z(0,t) = \frac{f_\ell}{v_{max}\rho_{max}} \tag{41}$$

Once the equation is derived, we relabel the z variable so that we return to our original variable and $\rho(x,t)$ represents the dimensionless density variable that is treated in the main body of this paper.

References

[1] T. W. Turowski and E. Petfalski, B. D. Goddard, S. L. French, A. Helwak and D. Tollervey (2020) "Nascent Transcript Folding Plays a Major Role in Determining RNA Polymerase Elongation Rates," Molecular Cell, 79: 3, pp. 488-503.

[2] A. Klindziuk and A. Kolomeisky (2021) "Long-Range Supercoiling-Mediated RNA Polymerase Cooperation in Transcription," J. Phys. Chem. B 2021, 125: 18, 4692-4700.

[3] J. Saba, X. Y. Chua, T. V. Mishanina, D. Nayak, T. A. Windgassen, R. A. Mooney and R. Landick (2019) The elemental mechanism of transcriptional pausing. Elife. 2019 Jan 8; 8: e40981. doi: 10.7554/eLife.40981.

[4] J. Y. Kang, T. V. Mishanina, R. Landick, S. A. Darst (2019) Mechanisms of Transcriptional Pausing in Bacteria. J Mol Biol. 2019; 431 (20): 4007-4029. doi: 10.1016/j.jmb.2019.07.017.

[5] C. T. MacDonald, J. H. Gibbs, J. H. and A. C. Pipkin, A. C., (1968) "Kinetics of biopolymerization on nucleic acid templates", Biopolymers, Vol. 6, p. 1-25.

[6] C. T. MacDonald and J. H. Gibbs (1969) "Concerning the kinetics of polypeptide synthesis on polyribosomes", Biopolymers, Vol. 7, p. 707-725.

[7] F. Spitzer, Frank (1970) "Interaction of Markov processes", Advances in Math., Vol. 5, pp. 246 - 290.

[8] R. Heinrich and T. A. Rapoport (1980) "Mathematical modelling of translation of mRNA in Eucaryotes: Steady states, time-dependent processes and application to reticulocytes", J. Theor. Biol., Vol. 86, pp. 279-313.

[9] B. Derrida, M. Evans, V. Hakim and V. Pasquier (1993) "Exact solution of a 1D asymmetric exclusion model using a matrix formalism," J. Phys. A: Math. Gen., Vol. 26, pp. 1493-1517.

[10] L. B. Shaw, R. K. Z. Zia and K. H. Lee, K. H. (2003) "Totally asymmetric exclusion process with extended objects: A model for protein synthesis," Phys. Rev. E, Vol. 68, 021910.

[11] A. Kolomeisky (1998) "Asymetric simple exclusion model with local inhomogeneity", J. Phys. A, Vol. 31, pp. 1153-1164.

[12] S. Klumpp and T. Hwa (2008) "Stochasticity and traffic jams in the transcription of ribosomal RNA: Intriguing role of termination and antitermination," PNAS, Vol. 105, No. 47, pp. 18159-164.

[13] S. Klumpp (2011) "Pausing and backtracking in transcription under dense traffic conditions," J. Stat. Phys., Vol. 142, pp. 1252 - 1267.

[14] C. A. Brackley, M. C. Romano, M. C., C. Grebogi and M. Thiel, M. (2010) "Limited resources in a driven diffusion process,", Phys. Rev. Let., Vol. 105.

[15] L. Ciandrini, L., I. Stansfield and M. C. Romano, M. C. (2010) "Role of the particle's stepping cycle in an asymmetric exclusion process: model of mRNA translation," Phys. Rev. E, Vol. 81, No. 2.

[16] J. Wang, J., B. Pfeuty, Q. Thommen, M. C. Romano, M. C. and M. Lefranc, M. (2014) "Minimal model of transcriptional elongation processes with pauses," Phys. Rev. E, Vol. 90.

[17] L. Mier-y-Teran-R, M. Silber and V. Hatzimanakatis (2010) "The origins of time-delay in template biopolymerization processes," PloS Comput Biology, Vol. 6, No. 4, pp. e1000726.

[18] R. K. P. Zia, J. J. Dong, and B Schmittmann, B. (2011) "Modeling translation in protein synthesis with TASEP: a tutorial and recent developments," J. Stat. Phys., Vol. 144, No. 2, pp. 405 - 428.

[19] L. Davis, T. Gedeon, J. Gedeon, and J. Thorenson (2013) "A Traffic Flow Model for Bio-polymerization Processes," J. Math. Bio., pp. 1416-1432.

[20] L. Davis, T. Gedeon and J. Thorenson (2014) "Discontinuous Galerkin Calculations for a Nonlinear PDE Model of DNA Transcription with Short, Transient and Frequent Pausing," J. Comp. Math. Vol. 32, No. 6, pp. 601-629.

[21] L. Davis, T. Gedeon and J. Thorenson (2014) "Parameter Studies for a Nonlinear Continuum Model of Transcription," Proceedings of 21st International Symposium on the Mathematical Theory of Networks and Systems.

[22] C. F. Daganzo (1995) "Requiem for second-order fluid approximations of traffic flow," Transportation Research Part B: Methodological, Vol. 29, No. 4, pp. 277-286.

[23] A. Aw and M. Rascle (2000) "Resurrection of "second order" models of traffic flow," SIAM J. Appl. Math., Vol. 60, No. 3, pp. 916–938.

[24] H. M. Zhang (2002) "A non-equilibrium traffic model devoid of gas-like behavior," Transportation Research Part B: Methodological, 36: 3 275-290.

[25] R. Mohan and G. Ramadurai (2015) "Submission to the DTA2012 Special Issue: A Case for Higher-Order Traffic Flow Models in DTA," Networks and spatial economics, 15: 3 765-790.

[26] R. Haberman (1998) Mathematical models: mechanical vibrations, population dynamics, and traffic flow: an introduction to applied mathematics, in series Classics in Applied Mathematics, SIAM, Philadelphia.

[27] M. D. Rosini (2013) Macroscopic Models for Vehicular Flows and Crowd Dynamics: Theory and Applications Classical and Non-Classical Advanced Mathematics for Real Life Applications in series Understanding Complex Systems, Springer.

[28] R. J. LeVeque (2002) "Finite volume methods for hyperbolic problems," in series Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge.

[29] K. W. Morton and D. F. Mayers (2005) "Numerical Solution of Partial Differential Equations: An Introduction," Ed. 2, Cambridge University Press, Cambridge.

[30] J. W. Thomas (2013) Numerical Partial Differential Equations: Finite Difference Methods in series Texts in Applied Mathematics, Vol. 22, Springer.

[31] J. W. Thomas (2013) Numerical Partial Differential Equations: Conservation Laws and Elliptic Equations, in series Texts in Applied Mathematics, Vol. 33, Springer, New York.

[32] A. Guzel and W. Layton (2018) "Time filters increase accuracy of the fully implicit method," BIT Numerical Mathematics, 58: 2 301-315.

[33] R. E. Schlesinger, L. W. Uccellini, Louis W and D. R. Johnson (1983) "The Effects of the Asselin Time Filter on Numerical Solutions to the Linearized Shallow-Water Wave Equations," Monthly Weather Review, 111: 3 455-467.

[34] P. P. Dennis, M. Ehrenberg, D. Fange and H. Bremer (2009) "Varying Rate of RNA Chain Elongation during rrn Transcription in Escherichia coli," Journal of Bacteriology, 191: 11 3740-3746.