MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **12**, Number **5**, October **2015**

doi:10.3934/mbe.2015.12.965

pp. **965–981**

STOCHASTIC MODELLING OF PTEN REGULATION IN BRAIN TUMORS: A MODEL FOR GLIOBLASTOMA MULTIFORME

Margherita Carletti and Matteo Montani

Department DISBEF, University of Urbino "Carlo Bo", Italy

Valentina Meschini

Department DISBEF, University of Urbino "Carlo Bo"
and Gran Sasso Science Institute, Italy

Marzia Bianchi and Lucia Radici

Department DISB, University of Urbino "Carlo Bo", Italy

(Communicated by Yang Kuang)

Abstract. This work is the outcome of the partnership between the mathematical group of Department DISBEF and the biochemical group of Department DISB of the University of Urbino "Carlo Bo"in order to better understand some crucial aspects of brain cancer oncogenesis. Throughout our collaboration we discovered that biochemists are mainly attracted to the instantaneous behaviour of the whole cell, while mathematicians are mostly interested in the evolution along time of small and different parts of it. This collaboration has thus been very challenging. Starting from [23, 24, 25], we introduce a competitive stochastic model for post-transcriptional regulation of PTEN, including interactions with the miRNA and concurrent genes. Our model also covers protein formation and the backward mechanism going from the protein back to the miRNA. The numerical simulations show that the model reproduces the expected dynamics of normal glial cells. Moreover, the introduction of translational and transcriptional delays offers some interesting insights for the PTEN low expression as observed in brain tumor cells.

1. **Introduction and motivation.** Glioblastoma multiforme (GBM) is the most common and most aggressive malignant primary brain tumor in humans. Experimental evidences show that patients affected by GBM have a lower level of protein PTEN than usual and the issue of PTEN downregulation in cancer has been addressed by researchers in different fields such as biological, biochemical and, recently, mathematical.

PTEN (phosphatase and tensin homolog) is a tumor suppressor that acts as a phosphatase for the lipid signaling intermediate phosphatidylinositol-3,4,5 trisphosphate (PIP3), producing phosphatidylinositol-4,5 bisphosphate. PIP3 anchors AKT to the membrane, where AKT is activated through its phosphorylation by phosphoinositide-dependent kinase-1 (PDK1) and mammalian target of rapamycin complex 2 (mTORC2). Upon activation, AKT phosphorylates numerous targets to transduce signals for growth, proliferation, and survival. PTEN, through PIP3

2010 *Mathematics Subject Classification.* Primary: 92D10, 65C99, 92-08, 92C45.

Key words and phrases. Brain tumors, stochastic modelling, gene regulation, PTEN, glioblastoma multiforme, numerical methods for genetic regulatory networks.

dephosporylation, inhibits the AKT/mTOR pathway. In addition to its effect on PIP3/AKT pathway, PTEN also regulates p53 function [28]. In the central nervous system, inhibition of PTEN leads to increased stem cell proliferation, somatic, dendritic and axonal growth, accelerated spine maturation, diminished synaptic plasticity, and altered intrinsic excitability [13]. PTEN gene, located on chromosome 10q23, is one of the most commonly mutated and deleted tumor suppressors in human cancer [10]. Loss of 10q, including PTEN gene, is the most common alteration associated with glioblastoma (70% incidence) [28]. PTEN levels are frequently found down-regulated in cancer also in the absence of genetic loss or mutation. PTEN is heavily regulated by transcription factors, microRNAs, ceRNAs (*competitive endogenous* RNAs) and methylation, while the tumor suppressive activity of the PTEN protein can be altered at multiple levels through aberrant phosphorylation, ubiquitination and acetylation [10]. These regulatory cues are presumed to play a key role in tumorigenesis through the alteration of the appropriate levels, localization and activity of PTEN.

PTEN protein levels are mainly regulated by microRNAs (miRNAs) that are 21-23 nucleotides (nt) long, endogenous, noncoding RNA molecules, able to perform post-transcriptional regulation by specifically binding target messenger RNAs (mRNAs), typically leading to a reduction in the levels of the corresponding proteins. miRNAs are transcribed from independent miRNA genes or from introns of protein-coding transcripts (in this last case called *mirtrons*) [6]. After being processed into maturity, a miRNA is loaded onto a specialized class of proteins to form the RNA-induced silencing complex (RISC), which specifically binds *miRNA response elements* (MREs) located in target mRNAs through a base-pairing recognition mechanism which requires at least 6-nt complementarity. The whole process, known as *RNA interference*, results in gene silencing through translation inhibition and/or mRNA destabilization [7]. Each mRNA can typically interact with several miRNAs, and each miRNA can target many different mRNAs. Moreover, miRNA-based regulation is strongly affected by global properties such as the total concentration of available targets (a feature known as dilution effect). The combination of the repressive effects of miRNAs on their targets and of the weakening of such repression due to *dilution effects* leads to effective, positive interactions between joint targets of a given miRNA (*crosstalk interactions*). ceRNAs contain the same MREs and are able to compete for the same miRNAs, thus acting as miRNA sponges and alleviating the repressive effects of the miRNA on its target mRNAs [11]. Transcripts acting as ceRNAs include both protein coding mRNAs and long noncoding RNAs (lncRNAs), and among the last ones pseudogene derived RNAs. The majority of validated ceRNAs are mRNAs, and their ability to sequester miRNAs from alternative targets can confer a protein coding-independent role on mRNAs. However, different classes of non-coding RNAs, such as long non-coding RNAs, have been shown to display ceRNA activity, the only prerequisite being the presence of MREs capable of binding miRNAs to inhibit their repression of protein coding targets [11]. A remarkable example is represented by pseudogenes where the preservation of regulatory sequences makes them good candidates as endogenous miRNA sponges, if co-expressed with the gene of origin. In order to be biologically effective, ceRNA crosstalk must occur under the right conditions: one of the most important is the relative abundance of miRNAs and the pool of ceRNAs that they target [25]. A titration mechanism has been described, whereby ceRNA crosstalk is optimal when the abundance of miRNA and ceRNA transcripts within a network

are near equimolarity. ceRNA crosstalk is minimal when ceRNA transcript level vastly exceeds that of miRNAs and vice versa: in the first case competition is minimized as all the miRNAs are bound and thus many targets remain unrepressed, while in the second one most targets will be fully repressed and ceRNA binding to miRNA will be ineffective to repress other targets [25]. This emerging scenario suggests that a broad class of ceRNAs may participate in a microRNA-dependent crosstalk, producing robust networks that, if perturbed, may lead to cancer.

The gene for the tumor suppressor PTEN is one with the most extensively characterised ceRNA network. Aside from the non-coding pseudogene PTENP1 [21], the network includes multiple protein-coding transcripts. Three papers in Cell 2011 remarked the importance of PTEN ceRNA-mediated regulation in cancer cells [24, 23, 18]. These studies allowed the identification and validation, by different computational and experimental approaches, of a list of protein-coding competing endogenous RNAs that regulate PTEN mRNA and protein in a miRNA-dependent fashion. Given the aberrant expression of PTEN in many cancers, with an intact PTEN locus, and the finding that PTEN regulation mostly takes place at post-transcriptional level as part of a miRNA-mediated crosstalk, a better dissection of the ceRNA crosstalk and the identification of other potential ceRNA interactors may have important implications for human diseases with PTEN involvement. Despite the body of experimental evidences, a clear quantitative understanding of miRNA-mediated regulation is still lacking and this represents an unmet goal [12].

Recently Ala et al. in [3] introduced a mathematical mass-action deterministic model to determine the optimal conditions for ceRNA activity *in silico*. From a computational point of view the ceRNA networks revealed that transcription factor and ceRNA networks are closely connected as already anticipated in [20]. They found that ceRNA networks are responsible for PTEN up-regulation and its aberrant expression in cancer. This suggests that optimal molecular conditions and alterations of one ceRNA can have dramatic effects on integrated ceRNA and transcriptional networks. Independently of Ala's work we developed a stochastic mathematical model for post-transcriptional PTEN regulation relying on the paper by Sumazin [23]. Our model extends Ala's results about the interaction between PTEN, ceRNAs and miRNAs including protein formation and backward mechanisms from proteins to the related miRNA both for PTEN and ceRNA.

2. **A competitive stochastic model for the regulation of PTEN.**

It is known that the biochemical kinetics involving small numbers of molecules can be very different to kinetics described by the law of mass action and differential equations [27]. This effect is a consequence of the intrinsic noise of the system and is associated with the uncertainty of knowing when a reaction occurs and what that reaction is. When considering a collection of molecules, the intrinsic noise is accentuated when some chemical species have small numbers, as is often the case in genetic regulatory models where there are small numbers of key transcription factors that can bind to a limited number of operator regions on DNA.

We introduce a chemical reactions model that describes the regulation of PTEN acting as a tumor suppressor. In modelling we consider the 'sponge effect 'due to ceRNAs, that is one of the main regulatory factors [23]. The main topic is that the translation of PTEN is blocked by miRNAs, so the concentration of this latter one is a first regulator. Moreover the miRNAs that regulates PTEN can be regulated in turn by other concurrent genes. Thus the concentration of the miRNAs of these

concurrent genes is a regulator of PTEN concentration: if the miRNAs join to concurrent genes, then they let PTEN free to produce the PTEN protein.

Looking at the dynamics of the model shown in Figure 1 we derive the following kinetic reactions which regulate PTEN expression, translating the biological meaning into the mathematical one:

$$\begin{aligned}
DNA_p &\xrightarrow{c_1} RNA_p + DNA_p \\
RNA_p &\xrightarrow{c_2} 0 \\
RNA_p &\xrightarrow{c_3} PTEN + RNA_p \\
PTEN &\xrightarrow{c_4} 0 \\
RNA_p + miRNA &\xrightarrow{c_5} RNA_{p*u} \\
RNA_{p*u} &\xrightarrow{c_6} RNA_p + miRNA \\
RNA_{p*u} &\xrightarrow{c_7} 0 \\
DNA_c &\xrightarrow{c_8} RNA_c + DNA_c \\
RNA_c &\xrightarrow{c_9} 0 \\
RNA_c &\xrightarrow{c_{10}} C_{prot} + RNA_c \\
C_{prot} &\xrightarrow{c_{11}} 0 \\
RNA_c + miRNA &\xrightarrow{c_{12}} RNA_{c*u} \\
RNA_{c*u} &\xrightarrow{c_{13}} RNA_c + miRNA \\
RNA_{c*u} &\xrightarrow{c_{14}} 0 \\
miDNA + C_{prot} &\xrightarrow{c_{15}} miRNA + miDNA + C_{prot} \\
miRNA &\xrightarrow{c_{16}} 0.
\end{aligned} \tag{1}$$

In (1) $PTEN$ is the protein, DNA_p is the PTEN gene, RNA_p is the mRNA of PTEN, RNA_{p*u} is the blocked miRNA/mRNA complex, DNA_c are the concurrent genes, RNA_c are the mRNAs of concurrent genes, C_{prot} are the proteins, RNA_{c*u} is the mRNA of blocked concurrent genes, $miDNA$ is the microRNA genes and finally $miRNA$ is the microRNA. Moreover each of the 16 reactions is characterised by a specific coefficient c_i $(i = 1, \ldots, 16)$, providing the probability that such a reaction will occur. Below we will explain in detail their biological meaning, how we define and derive them. This set of chemical reactions is uniquely characterised by two sets of quantities:

- the stoichiometric vectors $\nu_1, \ldots, \nu_{16}$, which can be collected in a stoichiometric matrix ν. This is a 10×16 matrix, where 10 is the number of variables and 16 is the number of reactions;
- the propensity functions $a_1, \ldots, a_{16}$, that describe the relative probabilities with which each reaction occurs.

In this way we can translate all the chemical reactions into a discrete ODE system using the Chemical Master Equation and the Law of Mass Action. Including all the state variables listed above in a vector called X, the ODE system becomes

$$X'(t) = \sum_{j=1}^{16} \nu_j a_j(X(t)), \qquad X(0) = X_0.$$

For our specific set of chemical reactions we obtain the following stoichiometric matrix ν. In each of its places we report the coefficient with which reactants (with minus sign) and products (with positive sign) appear.

$$\nu = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 1 & -1 \end{pmatrix}.$$

We still have to define the propensity functions and the coefficients of the chemical reactions. For these definitions we refer to the paper [16] by Goutsias. Given a system in a biological state, let us call it $X(t) = x$ at time t, the probability that the i-th reaction will occur during the time interval $[t, t + dt)$ is given by the propensity function $a_i(X(t))$ of the i-th reaction channel. This function is expressed as

$$a_i(x) = c_i \; h_i(x), \quad i \in S$$

where $S = 1, 2, \ldots, 16$, $c_i > 0$ is the specific probability rate constant of the i-th reaction and $h_i(x)$ is the number of all possible distinct combinations of the reactant molecules associated with the i-th reaction channel when the system is at state x, given by

$$h_i = \begin{cases} x_i & \text{for monomolecular reactions} \\ x_i(x_i - 1)/2 & \text{for bimolecular reactions with identical reactants} \\ x_i \; x_j & \text{for bimolecular reactions with distinct reactants} \end{cases}$$

for some $1 \leq i, \; j \leq N, \; i \neq j$. The specific probability rate constant c_i is the probability per unit time that a randomly chosen combination of reactant molecules of the i-th reaction will react. If the reaction rate constant K_i of the i-th reaction is known, then

$$c_i = \begin{cases} K_i & \text{for monomolecular reactions} \\ 2K_i/AV & \text{for bimolecular reactions with identical reactants} \\ K_i/AV & \text{for bimolecular reactions with distinct reactants} \end{cases}$$

where A is the Avogadro's number and V is the volume that, for simplicity, we assume constant and equal to 10^{-5}.

In our model we obtain the following formulation for the propensity functions a_i, with the corresponding coefficients c_i(molecules/seconds):

$$\begin{array}{ll} a_1 = c_1 \; x_1 & c_1 = 0.3769 \\ a_2 = c_2 \; x_2 & c_2 = 0.0004 \\ a_3 = c_3 \; x_2 & c_3 = 0.0007 \\ a_4 = c_4 \; x_3 & c_4 = 0.0002 \\ a_5 = c_5 \; x_2 \; x_{10} & c_5 = 0.0006 \\ a_6 = c_6 \; x_4 & c_6 = 0.0003 \\ a_7 = c_7 \; x_4 & c_7 = 0.0003 \\ a_8 = c_8 \; x_5 & c_8 = 0.07 \\ a_9 = c_9 \; x_6 & c_9 = 0.0004 \end{array} \tag{2}$$

$$
\begin{array}{ll}
a_{10} = c_{10}\ x_6 & c_{10} = 0.00002 \\
a_{11} = c_{11}\ x_7 & c_{11} = 0.0002 \\
a_{12} = c_{12}\ x_6\ x_{10} & c_{12} = 0.0006 \\
a_{13} = c_{13}\ x_8 & c_{13} = 0.0003 \\
a_{14} = c_{14}\ x_8 & c_{14} = 0.0003 \\
a_{15} = c_{15}\ x_9\ x_7 & c_{15} = 0.0001 \\
a_{16} = c_{16}\ x_{10} & c_{16} = 0.0002
\end{array}
$$

where some values of the coefficients c_i $(i = 1 \dots 16)$ are found in the literature [3] and some others are obtained with considerations on the stability of the system at equilibrium.

3. **Modelling regimes for Genetic Regulatory Networks.** A *Genetic Regulatory Network* (also called Gene Regulatory Network or GRN) is a collection of DNA segments in a cell that interact with each other (indirectly through their RNA and protein expression products) and with other substances in the cell, thereby governing the rates at which genes in the network are transcribed into mRNA. Regulation of gene expression, or gene regulation, refers to the cellular control of the amount and timing of changes to the appearance of the functional product of a gene. Although a functional gene product may be an RNA or a protein, the majority of the known mechanisms regulate the expression of protein coding genes. Any step of the gene expression may be modulated, from DNA-RNA transcription to the post-translational modification of a protein. Gene regulation gives the cell control over its structure and function, and it is the basis for cellular differentiation, morphogenesis and the versatility and adaptability of any organism. The controls that act on gene expression, i.e. the ability of a gene to produce a biologically active protein, are based on biochemical processes that are inherently stochastic.

In modelling GRN three different modelling regimes can be used. These include the discrete and stochastic, the continuous and stochastic and the continuous and deterministic.

The first regime is characterised by **Stochastic Simulation Algorithms** (SSA) in which there are low molecular numbers. The first SSA is due to Gillespie and it is an essentially exact procedure for numerically simulating the evolution of a set of chemical reactions in a well-stirred, homogeneous chemical reacting system Ω, by taking into account the randomness inherent in such a system. If $S_1, \dots, S_N$ represent N interacting molecular species through M reaction channels, then the evolution of such a system is characterised by a discrete nonlinear Markov process in which a vector $X(t)$ of dimension N, representing numbers (integer values) of the N molecular species at time t, is evolved through time. The state vector $X(t)$ is a discrete jump Markov process, whose time evolution equation describing the probability that, given $X(t_0) = x_0$, then $X(t) = x$, i.e. $P(x, t|x_0, t_0)$, is called the **Chemical Master Equation** (CME) and can be written as

$$\frac{\partial}{\partial t} P(x, t|x_0, t_0) = \sum_{j=1}^{M} (a_j(x - \nu_j) P(x - \nu_j, t|x_0, t_0) - a_j(x) P(x, t|x_0, t_0)).$$

In general this discrete parabolic partial differential equation is too difficult to solve (either analytically or numerically) and other techniques are needed to simulate $X(t)$.

Now, any set of chemical reactions is uniquely characterised by two sets of quantities: the stoichiometric vectors $\nu_1, \dots, \nu_M$ that represent, in turn, the update

of the numbers of molecules in the system if reactions $R_1, \ldots, R_M$ occur, respectively, and a set of M propensity functions $a_1(X(t)), \ldots, a_M(X(t))$, that describe the relative probabilities of reactions $R_1, \ldots, R_M$ occurring respectively. The SSA is rigorously based on the same microphysical premise that underlies the Chemical Master Equation and gives a more realistic representation of a system evolution than the deterministic Reaction Rate Equation (RRE). In particular, the RRE is completely inappropriate if the molecular population of some critical reactant species is so small that microscopic fluctuations can produce macroscopic effects. As for the CME, the SSA converges, in the limit of large numbers of reactants, to the same solution as the Law of Mass Action.

The algorithm for the SSA, developed in [14], is:

Algorithm 1 SSA

Data: stoichiometric vectors, reaction rates, initial state, simulation time
while $t < T$ **do**
  generate two random variables r_1, r_2 from $U(0,1)$
  compute $a_0(X) = \sum_{j=1}^{M} a_j(X)$
  compute $dt = \frac{1}{a_0(X)} \ln\left(\frac{1}{r_1}\right)$
  select j such that $\sum_{k=1}^{j-1} a_k(X) < r_2 a_0(X) \leq \sum_{k=1}^{j} a_k(X)$
  upgrade state vector as $X(t+dt) = X(t) + \nu_j$
  upgrade time as $t = t + dt$
end while
Result: state dynamics

It takes time steps of different length, based on the rate constants and population size of each chemical species. The probability of one reaction occurring relative to another is obtained by multiplying the rate constant of each reaction with the numbers of its substrate molecules. According to the correct probability distribution derived from the statistical thermodynamic theory, two random variables are simulated to choose which reaction will occur and how long the step will last. The chemical populations are altered according to the stoichiometry of the reaction and the process is repeated. The cost of this detailed algorithm is the large amount of computational time due to the fact that the time step for the next reaction to take place can be very small if we want to guarantee that only one reaction will fire in that time interval. Thus, SSA, also called slow regime, is a computationally demanding approach limiting its applicability especially for large reaction networks required for modelling realistic gene networks [8].

The key quantity is the step size, or waiting time, τ, whose value inversely depends on the size of the propensities of the different channel reactions and which needs to be re-evaluated after every firing event. Such a discrete event simulation may be extremely expensive, in particular for stiff systems where τ can be very short due to the fast kinetics of some of the channel reactions. Several alternative methods have been introduced to increase the integration step size. The so-called τ-leap approach takes a larger step size by allowing all the reactions to fire, from a Poisson or Binomial distribution, within that step.

Although τ-leap methods can, in some cases, substantially improve computational efficiency compared with the SSA, when there is moderate stiffness in the system the efficiencies can be quite poor.

The fast regime is the deterministic regime that describes the behaviour when there are large numbers of molecules, so that we can talk about concentrations. This averaged behaviour is described by an ODE of initial value type of the form

$$X'(t) = \sum_{j=1}^{M} \nu_j a_j(X(t)), \qquad X(0) = X_0.$$

This regime is the standard regime of chemical kinetics where the Law of Mass Action applies.

The intermediate regime is the regime where noise effects are still important but continuity arguments hold. By the application of the Central Limit Theorem and by matching the first two moments of the CME, we can write down an SODE that describes the evolution of X. This equation is sometimes called the **Chemical Langevin Equation** (CLE) and takes the following form

$$dX = \sum_{j=1}^{M} \nu_j a_j(X)dt + B(X)dW, \qquad X(0) = X_0.$$

Here $W(t) = (W_1(t), \ldots, W_N(t))'$ is an N dimensional vector whose individual elements are independent Wiener processes and $B(X)$ is a $N \times N$ matrix satisfying

$$B^2(X) = C(X) = \nu \mathrm{Diag}(a_1(X), \ldots, a_M(X))\nu^T$$

where $\nu = [\nu_1, \ldots, \nu_M]$ is the $N \times M$ stoichiometric matrix. The CLE is an Itô SODE (Stochastic Ordinary Differential Equation) and represents the continuous stochastic evolution of the equation describing internal noise [26].

4. **Modelling delays in GRNs.** Time delay is an important aspect in modelling genetic regulation due to slow biochemical reactions such as gene transcription and translation and protein diffusion between the cytosol and the nucleus [9]. The delay chemical master equation and the delay reaction rate equation can be developed for describing biological reactions with time delay, thus leading to stochastic delay differential equations derived from the Langevin approach and modelling intrinsic noise. In GRNs delays are associated with transcription and translation that do not occur instantaneously but take time due to other processes being involved, such as nuclear import and export, RNA polymerase activation, splicing, protein synthesis and folding.

In this section we describe how we can introduce delays into the SSA to investigate the dynamics of discrete models with delay. Unlike the SSA, there is not necessarily a unique implementation of delay SSA (DSSA). DSSA implementations can differ in the way they handle the waiting time for delayed reactions, the time steps in the presence of delayed reaction updates and the delayed consuming reactions. The DSSA version we use to produce the results presented in this paper works as follows: initially we specify which non-consuming reactions are delayed and the delay size (constant or variable) associated with each reaction. Inserting translational delay in our model it turns out that in the corresponding reaction reactants are still present as products so there are no delayed consuming reactions. Simulation proceeds by drawing reactions and their waiting times (for delayed and non-delayed reactions). If a non-delayed reaction is selected then the state is updated in the standard way (SSA), but if it is a delayed reaction that is selected, then it is not updated until the appropriate time point would be passed by another simulation step. In this case, the last drawn reaction is ignored and instead the

state is updated according to the delayed reaction. Simulation continues at the corresponding time point.

Algorithm 2 provides a pseudo-code description of our DSSA implementation, [5]. It is crucial to observe that the time step used by the DSSA is self-selecting,

Algorithm 2 DSSA

Data: stoichiometric vectors, reaction rates, initial state, simulation time, delay τ
while $t < T$ **do**
  generate two random variables r_1, r_2 from $U(0,1)$
  compute $a_0(X) = \sum_{j=1}^{M} a_j(X)$
  compute $dt = \frac{1}{a_0(X)} \ln\left(\frac{1}{r_1}\right)$
  select j such that $\sum_{k=1}^{j-1} a_k(X) < r_2 a_0(X) \leq \sum_{k=1}^{j} a_k(X)$
  if delayed reactions are scheduled within $(t, t+dt]$ **then**
    let k be the delayed reaction scheduled next at time $t+\tau$
    upgrade state vector as $X(t+\tau) = X(t) + \nu_k$
    upgrade t as $t = t + \tau$
  else
    if j is not a delayed reaction **then**
      upgrade state vector as $X(t+dt) = X(t) + \nu_j$
    else
      record time $t + dt + \tau$ for delayed reaction j
    endif
  endif
end while
Result: state dynamics

based on the assumption of exponential waiting times, as for SSA. The more stiff the kinetics system becomes, due to large rate constants and/or large numbers of molecules, the smaller the time step. Thus, the algorithm intrinsically controls the stability of the evolution. In the case of the continuous DDE representation an important issue is stepsize selection for any numerical method in order to avoid instabilities in the computed solutions.

5. **Numerical methods for SODEs and SDDEs.** Given the stochastic ordinary differential equation

$$\begin{array}{rcl} dy(t) & = & f(t,y(t))dt + \sum_{j=1}^{d} g_j(t,y(t))dW_j(t), \quad t \in [0,T] \\ y(0) & = & y_0 \end{array} \tag{3}$$

and a discretization of the time interval $[0,T]$

$$t_0 = 0 < t_1 < t_2 < \cdots < t_n < \cdots < t_N = T,$$

the simplest stochastic numerical scheme for the Itô SODE is the multidimensional Euler-Maruyama (EM) method

$$y_{n+1} = y_n + f(y_n)\Delta t_n + \sum_{j=1}^{d} g_j(y_n)\Delta W_{n_j}$$

where y_n is the numerical solution of (3) at time t_n, $\Delta t_n = t_{n+1} - t_n$ and $\Delta W_n = W_{n+1} - W_n = W_{t_{n+1}} - W_{t_n} = W(t_{n+1}) - W(t_n),\ n = 0, 1, 2, \ldots, N$.

The noise increments ΔW_{n_j} are $N(0, \Delta t_n)$-distributed independent random variables that can be generated numerically by pseudo-random number generators. The EM method has strong order of accuracy $\gamma = 0.5$ and converges to the Itô solution of the Itô system (3). If we apply the EM method to find an approximation to the weak solution of the Itô SODE we are free to select any $\sqrt{\Delta t}N(0,1)$ sample for the Wiener increment $\Delta W(t)$ on any step (since weak convergence concerns only the mean of the solution). Thus, the order of convergence is maintained if the increment is replaced by an independent two-point random variable $\Delta u_j = \sqrt{\Delta t}V_j$, where V_j takes the values $+1$ and -1 with equal probability. In this way, Δu_j has the same mean, variance and third moment as $\sqrt{\Delta t}N(0,1)$. The weak EM method obtained by replacing the $n-$th Wiener increment ΔW_{n_j} by $\sqrt{\Delta t_n}V_{n_j}$ has weak order 1. Note that a method may have a certain order of accuracy in general, but this order may be increased for SODEs of a particular type. For example, the EM method behaves as strong order 1 in the case of systems with additive noise. The concepts of strong and weak convergence concern the accuracy of a numerical method over a finite interval $[0, T]$ for small stepsizes Δt. *Numerical stability* of a one-step method concerns the property of errors to remain bounded with respect to an initial error for any SODE of the kind (3) having drift coefficients f satisfying global Lipschitz conditions. Namely, a one-step numerical method is said to be *numerically stable* if, for each time interval $[t_0, T]$ and stochastic ordinary differential equation (3) with $f(t, y)$ satisfying a Lipschitz condition, there exist positive constants Δ_0 and M such that

$$|y_n - \tilde{y_n}| \leq M|y_0 - \tilde{y_0}|,$$

where $n = 0, \ldots, N$ and $y_n, \tilde{y_n}$ are any two solutions obtained by the numerical method using time discretizations Δ_n such that $\max_n \Delta_n < \Delta_0$.

As for general SDDEs of the form

$$\begin{array}{rcl} dy(t) & = & f(y(t), y(t-\tau))dt + \sum_{j=1}^{d} g_j(y(t), y(t-\tau))dW(t), \quad t \in [0, T], T > 0 \\ y(\theta) & = & \psi(\theta), \qquad \theta \in J := [-\tau, 0], \quad \tau > 0, \end{array} \tag{4}$$

we define a mesh with uniform stepsize on the interval $[0, T]$, i.e. $h = T/N$ and $t_n = nh$ with $n = 0, \ldots, N$. We also assume that there exists an integer N_τ such that the delay can be expressed as a multiple of the stepsize, $\tau = N_\tau h$. We consider strong approximations $\tilde{y}_n$ of the solution y of (4) by means of the explicit stochastic one-step method

$$\tilde{y}_{n+1} = \tilde{y}_n + \Phi(h, \tilde{y}_n, \tilde{y}_{n-N_\tau}, I_\Phi), \qquad n = 0, \ldots, N-1, \tag{5}$$

where the initial values are given by $\tilde{y}_{n-N_\tau} := \psi(t_n - \tau)$ for $n - N_\tau \leq 0$. The increment function $\Phi(h, \cdot, \cdot, I_\Phi)$ incorporates a finite number of multiple Itô integrals, I_Φ, that is

$$I_{j_1,\ldots,j_l}(h) = \int_{t_k}^{t_{k+1}} \ldots \int_{t_k}^{s_2} dW^{j_1}(s_1) \ldots dW^{j_l}(s_l),$$

where $j_i \in \{0, 1\}$ and $dW^0(t) = dt$. The EM method takes the form

$$\tilde{y}_{n+1} = \tilde{y}_n + hf(\tilde{y}_n, \tilde{y}_{n-N_\tau}) + \sum_{j=1}^{d} g_j(\tilde{y}_n, \tilde{y}_{n-N_\tau})\Delta W_{n_j},$$

where $\Delta W_n = W_{(n+1)h} - W_{nh}$, $n = 0, \ldots, N-1$ are independent $N(0,h)$-distributed Gaussian random variables.

Baker and Buckwar in [4] proved that under suitable conditions on the increment function Φ, if f and g are such that (4) has a unique strong solution the one-step method (5) is convergent, as $h \to 0$ with $\frac{\tau}{h} \in \mathbb{N}$, in the mean-square sense with order $p = \frac{1}{2}$ and

$$\max_{1 \le n \le N} \left(\mathbb{E} \left| y(t_n) - \tilde{y}(t_n) \right|^2 \right)^{1/2} \le C h^p \qquad \text{as } h \to 0,$$

where $\epsilon_n = y(t_n) - \tilde{y}(t_n)$ is the global error. If the SDDE has additive noise, then the EM method is convergent with order $p = 1$ in the mean-square sense. Under the same assumptions on the increment function to achieve convergence, in the same paper, Baker and Buckwar also proved stability in the quadratic mean-square sense of the numerical method [4].

6. **Numerical simulations.** In this section we present the outcomes of the numerical simulations related to the chemical reaction system (1) in the discrete stochastic regime. Figure 2 depicts the concentrations of all reactants and products involved in the stochastic model (1), with propensity functions a_j and coefficients c_j as in formulation (2). We used the SSA described in Algorithm 1 in Section 3. This represents the slow regime. All the concentrations go to the steady state, characterising the normal behaviour of the interaction between PTEN and ceRNA within the cell. In Genetic Regulatory Networks, both the transcriptional and translational processes take some time due to other processes being involved in the cell. From the mathematical point of view these time lags are appropriately modelled by the inclusion of delays, called transcriptional and translational delay. If we first consider only transcriptional delay, the numerical simulations show very little effects on the behaviour of the system. Translational delay, i.e. the time lag between the mRNA production and the protein formation, has instead a much bigger impact on the dynamics of the system, due to the inclusion of the feedback processes that reflect interactions between the proteins and miRNAs, as we can see in Figure 1.

Figure 3 shows the dynamical behaviour of model (1), considering translational delay both for PTEN and concurrent genes. In order to observe the typical damped oscillations of the concentrations going to the steady state, the translational delay must be very large, of the order $10^4 s$ (approximately 3 hours). This value is larger than expected for normal biological time ranges.

Figure 4 shows sustained oscillations of system (1), all of them having the same period for all the concentrations. In this case the translational delay is of the order $10^5 s$ (approximately one day). For both Figures 3 and 4 the numerical simulations are obtained by the DSSA as in Algorithm 2.

Figure 5 is the enlargement of a small portion of Figure 3, where only miRNA and mRNAs are plotted. The picture shows the expression of mRNA with respect to miRNA concentrations. For high miRNA values, mRNA (both for PTEN and concurrents) is very low concentrated. On the contrary, if miRNAs are low concentrated, they have very little effect on mRNAs, thus allowing mRNA to remain at higher concentration. The graph shows, although for just a short period of time, the dramatic decrease of PTEN that is characteristic of tumor cells.

Figures 6 depicts the simulations in the fast regime, mathematically described by a systems of ODEs. An implicit ODE solver for moderately stiff systems is used. In these simulations the concentrations of each reactant are plotted separately, thus

enhancing the behaviour of PTEN and concurent genes. Graphs are obtained using the same parameters and delays as in the discrete stochastic regime and they are in agreement with the more realistic stochastic dynamics.

7. **Discussion.** With this work we meant to propose a mathematical stochastic model for PTEN post-transcriptional regulation, that is a crucial aspect in Glioblastoma oncogenesis. We considered the inward and backward interaction between the transcription factor and the miRNA, including the sponge effect due to concurrent genes [23]. We performed simulations for the discrete stochastic regime modelling intrinsic noise through the SSA. Varying the model coefficients in deterministic setting it is possible to observe the pathological downregulation of PTEN in cancer [3]. Instead, we tried to offer some insights for the low expression of PTEN in brain cancer without considering extrinsic variations of the coefficients. In stochastic setting and introducing (translational) delays, the system is able to capture the correct oscillatory dynamics around the steady state, although for large values of the delays and this is somehow difficult to justify biologically. In turn, the decrease of PTEN levels is visible only for such large delay values. This can be due to the (stiff) nature of the discrete system. New insights could be offered by simulating the kinetics in the intermediate modelling regime, that is using SODEs and, if delays are considered, SDDEs. However, in the case of moderately stiff systems explicit Runge-Kutta methods for SODEs with extended stability regions along the negative real axis have proven to be especially effective [1, 2]. Runge-Kutta methods are a class of one step methods which gain their efficacy by computing intermediate approximations to the solution within a step. Explicit Runge-Kutta methods with extended stability regions are based on explicit Runge- Kutta methods whose stability function is a shifted and scaled Chebyshev polynomial or some variant thereof. In the stochastic setting, there are difficul ties designing fully implicit methods due to possible unboundedness of the solution as the Wiener increment can take positive or negative values with equal likelihood [17]. Thus most methods are semi-implicit, that is implicit in the deterministic component. Abdulle and Cirilli [2] have extended the ideas of explicit Chebyshev methods with extended stability regions to the SODE setting, through the class of S-ROCK methods.

Our future research will be addressed to simulate the discrete model using Poisson or Binomial $\tau-$ leap methods, that are known to be more efficient for stiff systems [15, 19], and explore the simulation in the intermediate (stochastic and continuous) regime using appropriate methods. We plan to obtain more appropriate coefficients c_i by means of specific biological experiments and finally, we hope to find more efficient strategies to introduce delay, in order to observe downregulation of PTEN due to intrinsic properties of the model itself.

Acknowledgments. The first author wishes to thank Professor Kevin Burrage for preliminary discussions on competitive models for PTEN regulation and following chats on the final version of the model. All the authors are grateful to the Referees for their careful and detailed work in the reviewing process, allowing them to improve the quality of the article very much.

REFERENCES

[1] A. Abdulle and A. Medivikov, Second order Chebyshev methods based on orthogonal polynomials, *Numerische Mathematik*, **90** (2001), 1–18.

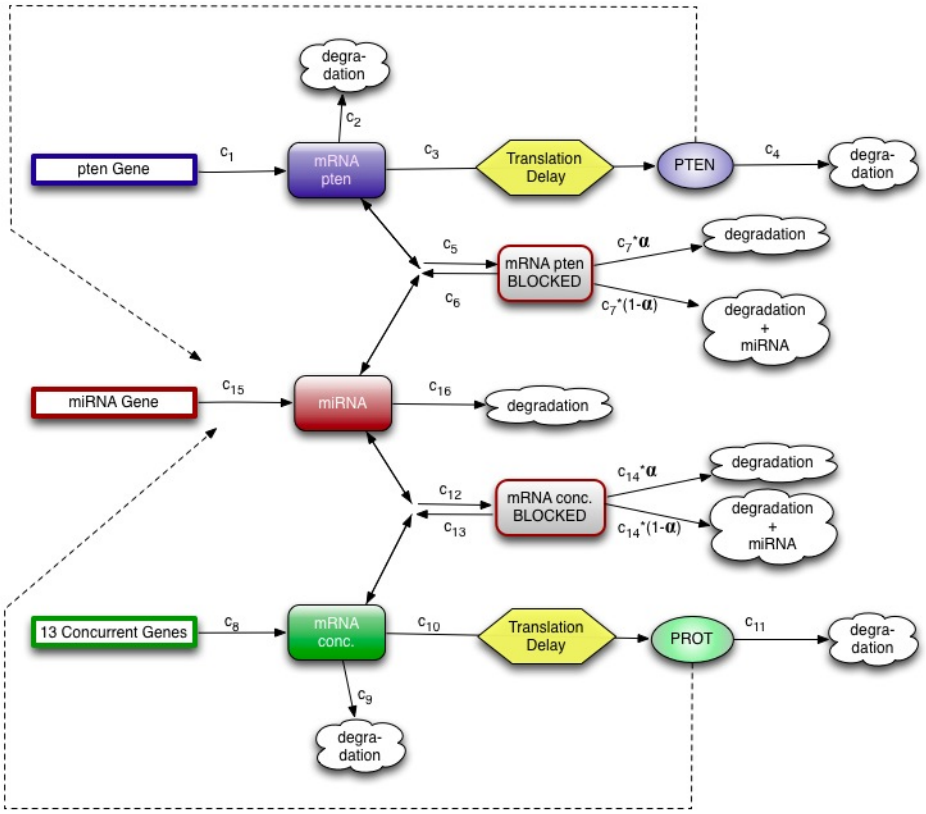

FIGURE 1. Dynamics of the model for the regulation of PTEN.

[2] A. Abdulle and S. Cirilli, S-ROCK: Chebyshev methods for stiff stochastic differential equations, *SIAM J. Sci. Comput.,* **30** (2008), 997–1014.
[3] U. Ala, F. A. Karreth, C. Bosia, A. Pagnani, R. Taulli, V. Léopold, Y. Tay, P. Provero, R. Zecchina and P. P. Pandolfi, Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments, *PNAS,* **110** (2013), 7154–7159.
[4] C. T. H. Baker and E. Buckwar, Numerical analysis of explicit one-step methods for stochastic delay differential equations, *LMS J. Comput. Math.,* **3** (2000), 315–335.
[5] M. Barrio, K. Burrage, A. Leier and T. Tian, Oscillatory Regulation of Hes1: Discrete stochastic delay modelling and simulation, PLoS Comput Biol, 2006.
[6] D. P. Bartel, MicroRNAs: Genomics, biogenesis, mechanism, and function, *Cell,* **116** (2004), 281–297.
[7] D. P. Bartel, MicroRNAs: Target recognition and regulatory functions, *Cell,* **136** (2009), 215–233.
[8] K. Burrage, T. Tian and P. M. Burrage, A multi-scaled approach for simulating chemical reaction systems, *Progress in Biophysics and Molecular Biology,* **85** (2004), 217–234.
[9] M. Carletti, *Stochastic Modelling of Biological Processes*, PhD Thesis, The University of Queensland, Brisbane, Australia, 2008.
[10] A. Carracedo, A. Alimonti and P. P. Pandolfi, PTEN level in tumor suppression: How much is too little?, *Cancer Res.,* **71** (2011), 629–633.
[11] A. de Giorgio, J. Krell, V. Harding, J. Stebbing and L. Castellano, Emerging roles of competing endogenous RNAs in cancer: Insights from the regulation of PTEN, *Mol Cell Biol.*, **33** (2013), 3976–3982.

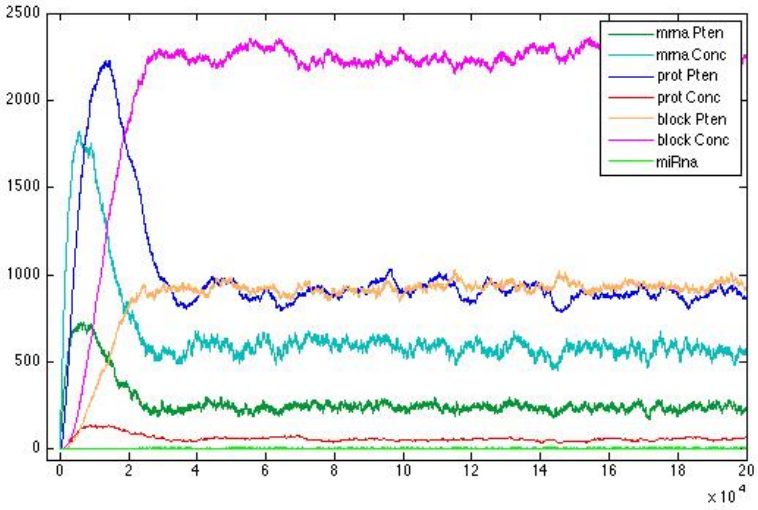

FIGURE 2. Simulation of the discrete system (1) using SSA as in Algorithm 1. The Figure shows the concentrations of all the reactants and products vs time as indicated in the legend.

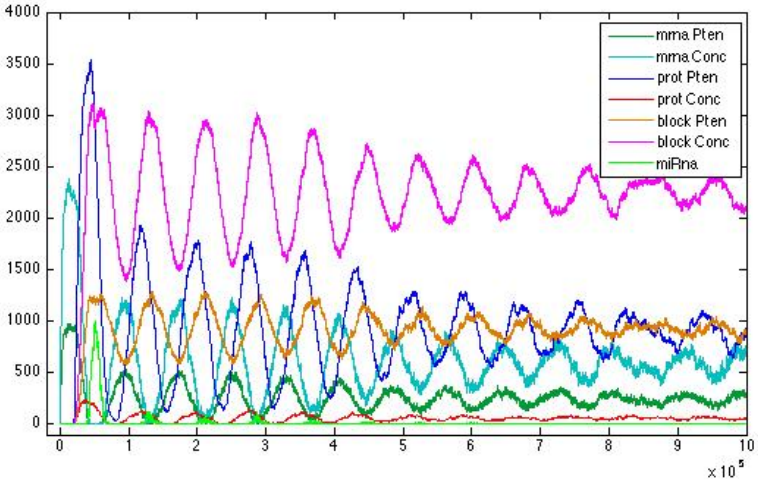

FIGURE 3. Simulation of the discrete system (1) including translational delay, using DSSA as in Algorithm 2. The Figure shows the concentrations of all the reactants and products vs time as indicated in the legend. Damped oscillations can be seen only for large delays.

[12] M. Figliuzzi, E. Marinari and A. De Martino, MicroRNAs as a selective channel of communication between competing RNAs: A steady-state theory, *Biophys J.,* **104** (2013), 1203–1213.
[13] P. Garcia-Junco-Clemente and P. Golshani, PTEN: A master regulator of neuronal structure, function, and plasticity, Commun Integr Biol., 2014.

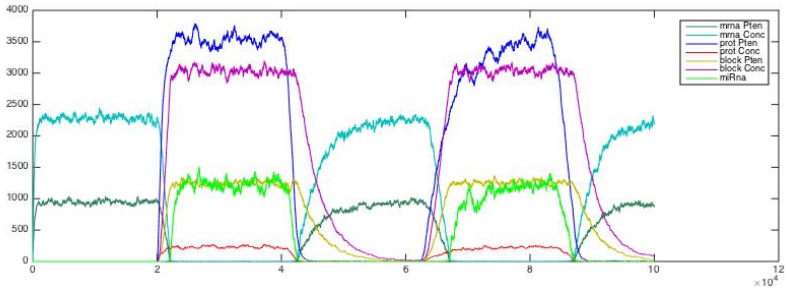

FIGURE 4. Simulation of the discrete system (1) including translational delay, using DSSA as in Algorithm 2. The Figure shows the concentrations of all the reactants and products vs time as indicated in the legend. Sustained oscillations can be seen for larger delays than in Figure 3.

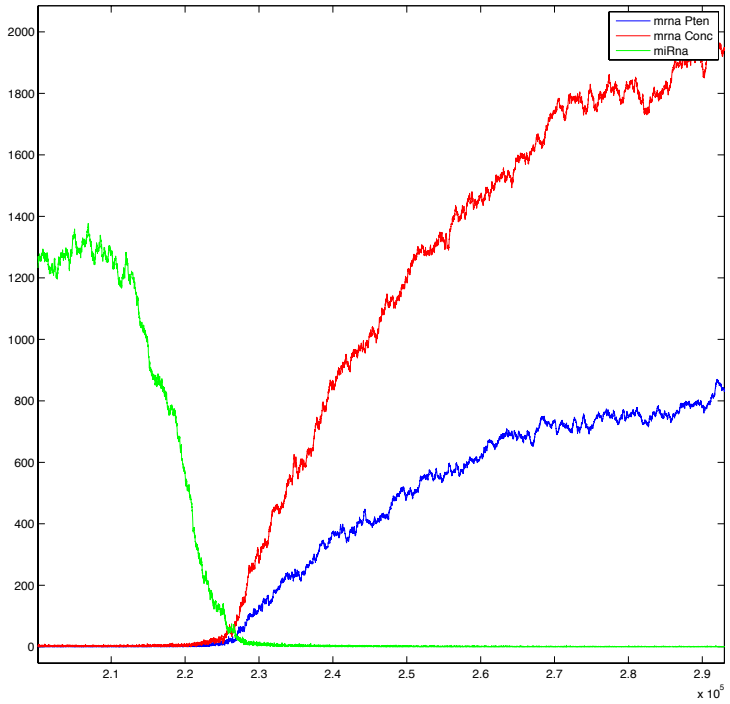

FIGURE 5. miRNA (of PTEN and concurrent genes) and the mRNA concentration are plotted vs time. High miRNA values imply low PTEN expression.

[14] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.*, **81** (1977), 2340–2361.

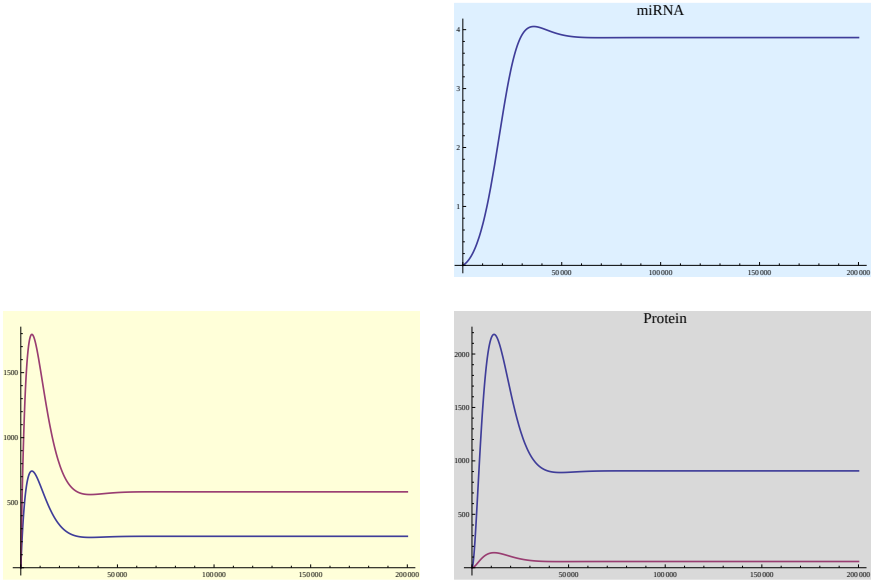

FIGURE 6. Simulations of miRNA, protein, blocked and mRNA of both PTEN and concurrent genes vs time using the ODE system. In each picture the blue line refers to PTEN and the red one to the concurrent genes.

[15] D. T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems, *J. Chem. Phys.,* **115** (2001), 1716–1733.
[16] J. Goutsias, Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems, *J. Chem. Phys.,* **122** (2005), 184102.
[17] D. Hernandez and R. Spigler, Convergence and stability of implicit Runge-Kutta methods for systems with multiplcative noise, *BIT Num. Math.,* **33** (1993), 654–669.
[18] F. A. Karreth, Y. Tay, D. Perna, U. Ala, S. Mynn Tan, A. G. Rust, G. De Nicola, K. A. Webster, D. Weiss, P. A. P. Mancera, M. Krauthammer, R. Halaban, P. Provero, D. J. Adams, D. A. Tuveson and P. P. Pandolfi, In vivo identification of Tumor-suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma, *Cell,* **147** (2011), 382–395.
[19] A. Leier, T. T. Marquez-Lago and K. Burrage Generalized binomial tau-leap method for biochemical kinetics incorporating both delay and intrinsic noise, *J. Chem. Phys.,* **128** (2008), 205107.
[20] S. Mukherji, M. S. Ebert , G. X. Zheng, J. S. Tsang, P. A. Sharz and A. van Oudenaarden, MicroRNAs can generate thresholds in target gene expression, *Nat. Genet,* **43** (2011), 854–859.
[21] L. Poliseno, L. Salmena, J. Zhang, B. Carver, W. Haveman and P. P. Pandolfi, A coding-independent function of gene and pseudogene mRNAs regulates tumour biology, *Nature,* **465** (2010), 1033–1038.
[22] P. Rue, J. Villa-Freixa and K. Burrage, Simulation methods with extended stability for stiff biochemical kinetics, *BMC Systems Biology,* (2010), p110.
[23] P. Sumazin, X. Yang, H. S. Chiu, W. J. Chung, A. Iyer, D. Llobet-Navas, P. Rajbhandari, M. Bansal, P. Guarnieri, J. Silva and A. Califano, An extensive MicroRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in Glioblastoma, *Cell,* **147** (2011), 370–381.
[24] Y. Tay, L. Kats, L. Salmena, D. Weiss, S. M. Tan, U. Ala, F. Karreth, L. Poliseno, P. Provero, F. Di Cunto , J. Lieberman, I. Rigoutsos and P. P. Pandolfi, Coding Independent Regulation of the Tumor Suppressor PTEN by Competing Endogenous mRNAs, *Cell,* **147** (2011), 344–357.

[25] Y. Tay, J. Rinn and P. P. Pandolfi, The multilayered complexity of ceRNA crosstalk and competition, *Nature,* **505** (2014), 344–352.
[26] T. Tian, K. Burrage, P. M. Burrage and M. Carletti, Stochastic delay differential equations for genetic regulatory network, *J. Comp App. Math.,* **205** (2007), 696–707.
[27] T. E. Turner, S. Schnell and K. Burrage, Stochastic approaches for modelling in vivo reactions *Comput. Biol. and Chem.,* **28** (2004), 165–178.
[28] J. Xu, Z. Li, J. Wang, H. Chen and J. Y. Fang, Combined PTEN Mutation and Protein Expression Associate with Overall and Disease-Free Survival of Glioblastoma Patients, *Transl Oncol.,* **7** (2014), 196–205.

Received October 28, 2014; Accepted April 13, 2015.

E-mail address: margherita.carletti@uniurb.it
E-mail address: matteo.montani@uniurb.it
E-mail address: meschini.valentina@gmail.com
E-mail address: radici.lucia@gmail.com
E-mail address: marzia.bianchi@uniurb.it