Derbal, Y., "State Machine Modeling of MAPK Signaling Pathways", *2014 IEEE Engineering in Medicine and Biology Conference*, Chicago, Michigan, 2014.

State Machine Modeling of MAPK Signaling Pathways

Youcef Derbal

Ryerson University
yderbal@ryerson.ca

***Abstract*— Mitogen activated protein kinase (MAPK) signaling pathways are frequently deregulated in human cancers with potential involvement in most if not all cellular processes leading to tumorigenesis. Mathematical/computational models of MAPK signaling are indispensable to the study of pathway deregulation dynamics and their nonlinear effects on cell fate and carcinogenesis. A finite state machine model of MAPK cellular signaling is explored as an alternative to differential equations-based models of kinetics. The proposed approach is applied to the Ras-Extracellular signal-regulated kinase (Ras-ERK) pathway which includes the frequently mutated Ras and RAF proteins in many types of carcinomas.**

I. INTRODUCTION

The mechanisms underlying the functions of multicellular organisms rely on the transduction and processing of extra-cellular stimuli by a cellular signaling circuitry controlling the regulation of gene expression and ultimately cell fate. Deregulated cellular signaling pathways have been implicated in a variety of pathologies including cancer, cardiovascular diseases and asthma. Mathematical/computational models of cellular signaling are instrumental to the study of pathway deregulation dynamics and their effects on cell fate and carcinogenesis. In theory, they can be used for a variety of ends, including the rational design of therapies, the assessment of pharmaceutical drugs' efficacy in altering a disease course and the estimation of their side effects. Ordinary differential equations (ODEs) are often used to represent the kinetics of the relevant biochemical processes, including the ubiquitous phosphorylation-dephosphorylation reactions underlying the mitogen activated protein kinase (MAPK) signal transduction [1], [2], [3-5] [6]. These kinetics models can yield insights about the dynamic behaviour of signaling pathways such as the role of feedback in the amplification of signals [7], the emergence of bistable and oscillatory behaviours in signaling networks [8], and the realization of ultrasensitivity and amplification of mitogen signals under enzyme saturation [9]. ODE based models can also shed light on the relationship between signal amplification and speed of signal propagation as well as enable insight as to how both transient and sustained signaling can be achieved through the same pathways under different stimuli [1]. On the other hand, the reaction rates associated with these ODE-based models are often estimated from sparse experimental data, if at all available, or curated from various literature sources which undoubtedly made use of different measurement protocols. In this paper, we explore an alternative modeling approach built around an information processing abstraction of cellular mechanisms, which has been recently discussed vis-à-vis the computational modeling of living organisms [10]. As such the resulting models are intrinsically free from the need to replicate the mechanistic details of reaction kinetics, avoiding hence the shortcoming highlighted above for ODE based signaling models.

II. MAPK SIGNALING AND CANCER

MAPK pathways include a three-tier cascade of protein kinases activated by growth and stress signals as well as cytokines via cell receptors such as tyrosine kinase receptors (RTKs) and G-protein coupled receptors (GCPR). These pathways mediate the regulation of various cellular processes, such as cell growth, proliferation, differentiation and cell death. The known major MAPK cascades are distinguished by their respective end effector kinases, namely: ERK, p38, JNK, and ERK 5. These pathways are frequently deregulated in human cancers [11, 12], with potential involvement in most if not all cellular processes leading to tumorigenesis, including independence from growth stimuli, unresponsiveness to anti-growth signals, circumvention of apoptosis, boundless replicative potential, maintained angiogenesis, and capacity for tissue invasion and metastasis [13]. In this paper we focus on the Ras-ERK pathway as an application example for the proposed modeling approach (see figure 1). This pathway has a particular importance to cancer research because of the frequently mutated Ras and RAF proteins in many types of carcinomas that affect the skin, breasts, pancreas, liver and the thyroid respectively. The deregulated activation of the Ras-ERK pathway, underlying these pathologies, is driven by multiple sources acting independently or in coalescence. These includes: overexpression of cell membrane receptors and ligands, mutations of receptors and signaling proteins, and the sustained ligand-receptor stimuli mediated by autocrine or paracrine means. This understanding of the Ras-ERK signaling pathway inspired various therapeutic approaches focused on drug mediated inhibition of the mutant Ras-RAF axis, considered to be the key driver of the deregulated activation of ERK. Unfortunately, this target inhibition is circumvented by cancerous cells through the activation of the PI3K-mTOR pathway whose deregulated control of cell growth and survival co-exist with that of the Ras-ERK signaling pathway for many of the above listed cancers [11] [14]. In addition to the cross-talk between Ras-ERK and PI3K-mTOR signaling pathways and their co-regulation of differentiation, cell growth and proliferation, both pathways are under the inhibitory control of the energy sensing AMPK (see figure 1). This later modulates the Ras-ERK pathway through the inhibition of wild type BRAF, a member of the RAF family of proteins [15], and controls mTORC1 through the phosphorylation of TSC2 and Raptor [16]. The interaction between the Ras-ERK and PI3K-mTOR pathways as co-regulators of survival, growth proliferation and their modulation by AMPK lead to complex signaling dynamics. This solidifies the importance of mathematical/computational modeling as an indispensable tool in any effort to tease out an actionable insight about these coupled nonlinear dynamics and to prognosticate about cancer response to therapeutic interventions.

Derbal, Y., "State Machine Modeling of MAPK Signaling Pathways", *2014 IEEE Engineering in Medicine and Biology Conference*, Chicago, Michigan, 2014.

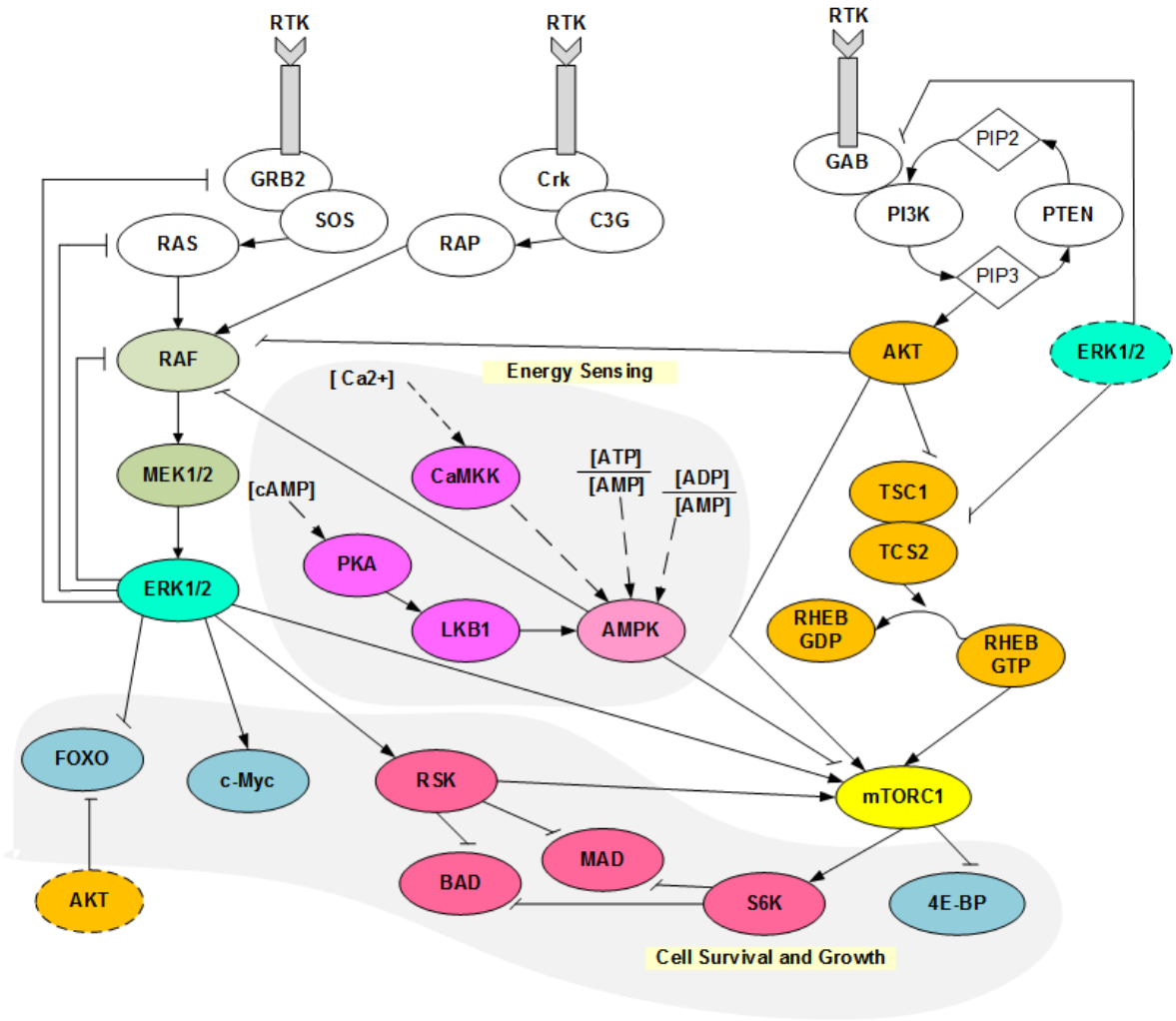

Figure 1. Cell growth and survival co-regulation by the Ras-ERK and PI3K-mTOR signaling pathways

Derbal, Y., "State Machine Modeling of MAPK Signaling Pathways", *2014 IEEE Engineering in Medicine and Biology Conference*, Chicago, Michigan, 2014.

III. FINITE STATE MACHINE MODEL OF THE RAS-ERK PATHWAY

Each stage of the Ras-ERK cascade corresponds to a phosphorylation-dephosphorylation cycle (PdPC) governed by the following reactions (see figure 2):

$$X + E_1 \underset{d_1}{\overset{a_1}{\rightleftarrows}} XE_1 \xrightarrow{k_1} X^* + E_1 \qquad (1.1)$$

$$X^* + E_2 \underset{d_2}{\overset{a_2}{\rightleftarrows}} X^*E_2 \xrightarrow{k_2} X + E_2 \qquad (1.2)$$

X and X* represent a signaling protein and its phosphorylated form respectively. X* acts as the phosphorylating enzyme of the next stage in the MAPK cascade. Taking the second tier as an example, X and X* represent MEK and MEK* respectively. In this case E_1 represents RAF*, while E2 is a phosphatase that dephosphorylates MEK*. In the proposed model, each tier of the MAPK cascade includes a finite state machine (FSM) used to model the temporal dynamics of signal propagation resulting from the physicochemical transformation applied to the involved signaling proteins. The sates of the FSM are (see figures 3 and 4): DP (dephosphorylating), TP (transitioning towards phosphorylation), P (phosphorylating), TDP (transitioning towards dephosphorylation). In this model the rates of enzyme association, catalysis and dissociation are indirectly accounted for by the average time T_S spent by the cascade stage in the corresponding state $s \in \{0,1,2,3\}$. The average is understood to be over the ensemble of the relevant proteins being covalently modified in the cascade tier under consideration. The corresponding average time T_C needed for a complete PdPC can therefore be written as the sum of the average times spent in each of the four states of the cycle. Normalizing the sum over T_C leads to $\tau_0 + \tau_1 + \tau_2 + \tau_3 = 1$, where $\tau_s = T_S/T_c$ is the proportion of the average time spent in state s relative to the average phosphorylation-dephosphorylation cycle time. These time proportions depend on the kinetic rates associated with the transitions to and from the states in question. In particular, based on the kinase and phosphate reactions given in (1), we can write the following estimates: $\tau_0 = \frac{k_2+d_1}{a_1}$, $\tau_1 \approx \frac{a_1}{d_1+k_1}$, $\tau_2 \approx \frac{k_1+d_2}{a_2}$, $\tau_3 \approx \frac{a_2}{k_2+d_2}$.

Assuming the dissociation rates d_1 and d_2 to be comparable, these approximations can be rewritten as: $\tau_0 \approx \alpha k_{m2}$, $\tau_1 \approx \frac{1}{k_{m1}}$, $\tau_2 \approx \frac{1}{\alpha} k_{m1}$, $\tau_3 \approx \frac{1}{k_{m2}}$, $k_{m1} = \frac{k_1+d_1}{a_1}$, and $k_{m2} = \frac{k_2+d_2}{a_2}$, where k_{m1} and k_{m2} are the Michealis constants. The signaling proteins spend the majority of the PdPC time in the unmodified or phosphorylated states respectively. We can therefore assume that $\tau_1 + \tau_3 \ll \tau_0 + \tau_2$. With the assumption that the Michealis constants $k_{m1}$and k_{m2} are comparable, we have: $\tau_o = \frac{\alpha^2}{1+\alpha^2}$, and $\tau_2 = \frac{1}{1+\alpha^2}$. The parameter $\alpha = a_2/a_1$ represents the phosphatase to kinase relative ratio of affinity vis-à-vis their respective substrates.

Derbal, Y., "State Machine Modeling of MAPK Signaling Pathways", *2014 IEEE Engineering in Medicine and Biology Conference*, Chicago, Michigan, 2014.

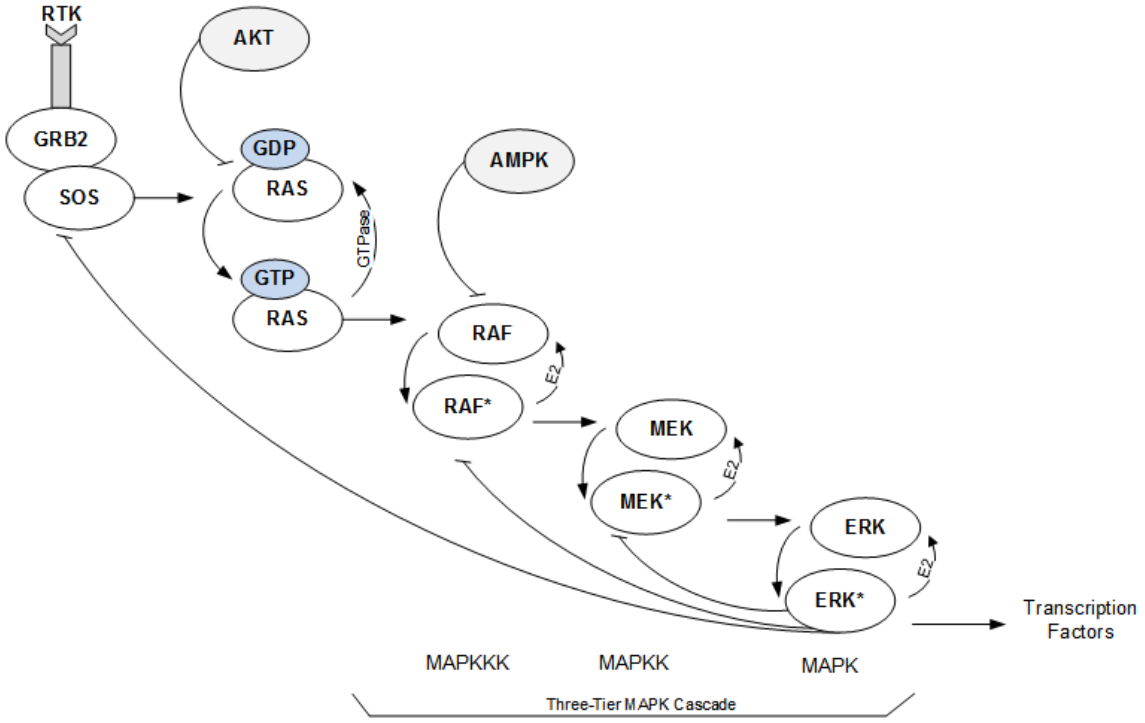

Figure 2. The Ras-ERK cascade.

The Ras-ERK simulation model consists of three FSMs, each representing one tier of the MAPK cascade. The phosphorylated state of one stage triggers the life cycle of the FSM in the next stage, which can be affected by inhibitory signals from other protein kinases, such as AMPK. The proposed model of phosphorylation and dephosphorylation dynamics is given for the j-th MAPK cascade stage as follows:

$$u_j(t_n) = \sum_{i=1}^{n} N(s(t_i), u_{j-1}(t_i)), \quad j = 1,2,3 \qquad (2.1)$$

$$N\big(s(t_i), u(t_i)\big) = \begin{cases} -\beta & if \;\; s(t_i) = 0 \\ \dfrac{\lambda * u(t_i)}{1 + \gamma * y(t_i)} & if \;\; s(t_i) = 2 \\ 0 \quad if \; s(t_i) = 1 \; or \;\; s(t_i) = 3 \end{cases} \qquad (2.2)$$

Derbal, Y., "State Machine Modeling of MAPK Signaling Pathways", *2014 IEEE Engineering in Medicine and Biology Conference*, Chicago, Michigan, 2014.

t_n and t_i are the n-th and i-th discrete simulation times respectively. $u_j(t_n)$ is the number of phospohorylated molecules at time t_n. This signal constitutes the output of the j-th MAPK cascade and the driving input of the cascade stage j+1 (see figures 3 and 4). The proposed model of phosphorylation dynamics uses a linear approximation of the cooperative ultrasensitivity that ensues when the kinase and phosphatase enzymes operate under saturating conditions [9]. The parameter λ quantifies the signal amplification gain associated with such ultrasensitivity. The model includes the factor $\frac{1}{1+\gamma * y(t_i)}$ to integrate the contribution of an inhibitory signal $y(t_i)$ such as that of the feedback from ERK to RAF in the MAPK example at hand. The parameter γ represents the inhibition strength. The dephosphorylation process is on the other hand assumed to take place with the same constant rate β for all MAPK tiers. For M phosphorylating signals and Q inhibitory signals, the proposed model can be generalized as follows:

$$N(s(t_i)) = \begin{cases} -\beta & if \;\; s(t_i) = 0 \\ \dfrac{\sum_{q=1}^{Q} \lambda_q x_q}{1 + \sum_{m=1}^{M} \gamma_m * y_m(t_i)} & if \;\; s(t_i) = 2 \\ 0 & if \; s(t_i) = 1 \; or \; s(t_i) = 3 \end{cases} \qquad (3)$$

Where x_q and λ_q, $q = 1, .., Q$ are the phosphorylating signals and their associated cooperative amplification gains. y_m and $\gamma_m, m = 1, .., M$ are the inhibitory signals and their respective inhibition strengths.

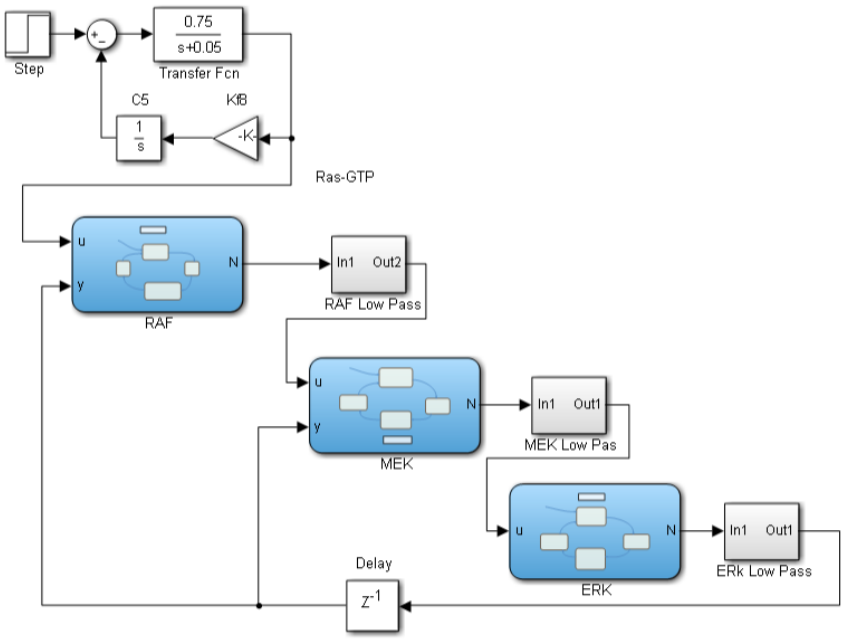

Figure 3. MATLAB ® model of the Ras-ERK pathway.

Derbal, Y., "State Machine Modeling of MAPK Signaling Pathways", *2014 IEEE Engineering in Medicine and Biology Conference*, Chicago, Michigan, 2014.

The times spent in the states of the FSM model were estimated using averages over the population of involved signaling proteins. Consequently, such in-state times are simulated as normally distributed random variables $T(s)$, $s = 0,1,2,3$ with means equal to τ_s and standard deviations of $0.1\ \tau_s$ respectively. The simulation of the proposed model is illustrated in figures 5 and 6 for $\gamma = 0$ (no inhibitory signals), and $\gamma = 1$ respectively. The simulation results are obtained using $T_C = 15$ seconds, $\alpha = 1.5$, and $k_{m1} = k_{m2} = 0.01W_T$, where $W_T = 2.5 \times 10^3$ is the total number of RAF molecules involved in the corresponding phosphorylation-dephosphorylation cycle of the first MAPK cascade stage. Furthermore, λ is set to 0.0025, 0.0035, and 0.0007 for the cascade stages 1, 2, and 3 respectively. The dephosphorylation rate β is set to 100 for all the cascade tiers.

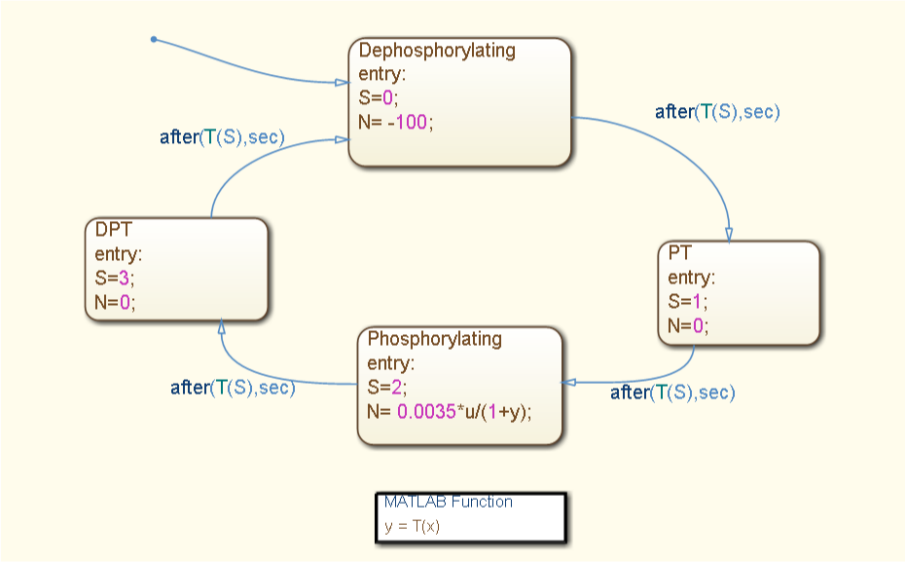

Figure 4. The FSM for the second stage of the Ras-ERK cascade.

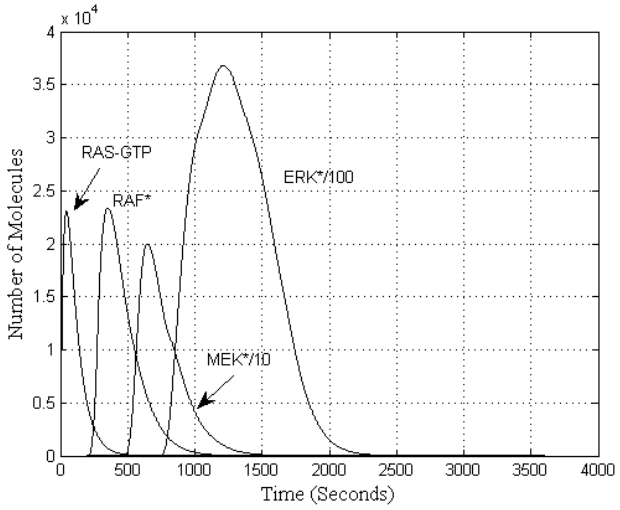

Figure 5. The temporal profiles of signaling protein numbers for $\gamma = 0$.

Derbal, Y., "State Machine Modeling of MAPK Signaling Pathways", *2014 IEEE Engineering in Medicine and Biology Conference*, Chicago, Michigan, 2014.

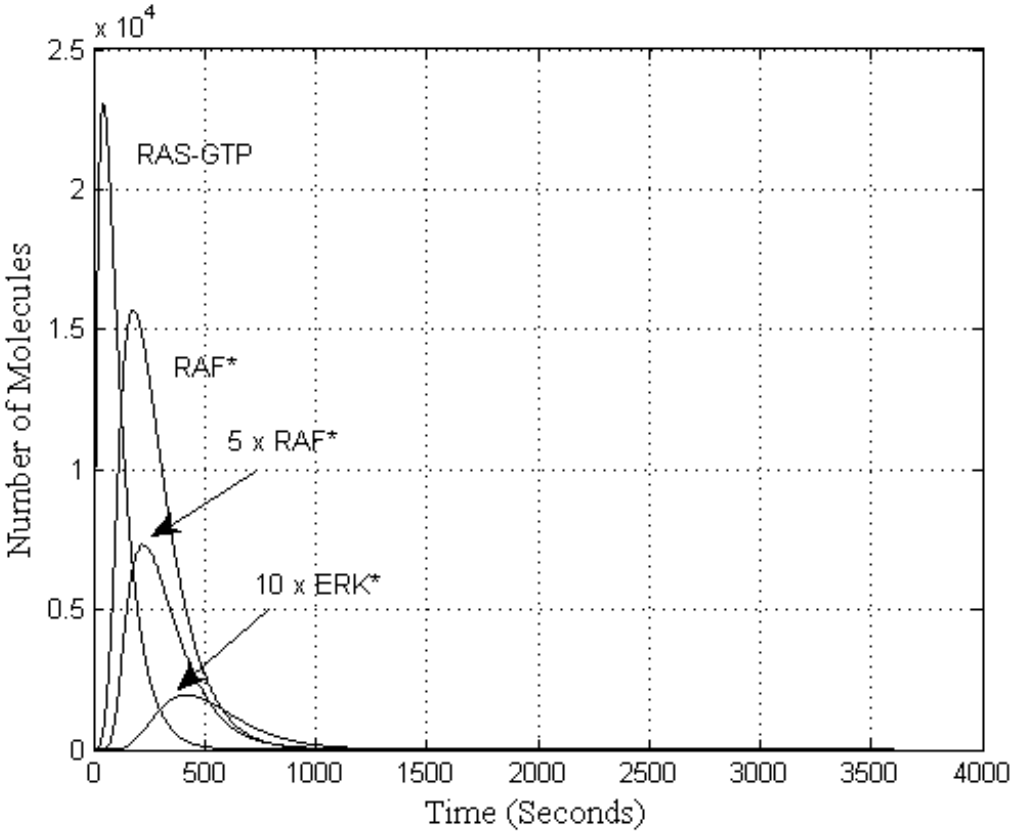

Figure 6. The temporal profiles of the signaling protein numbers for $\gamma = 1$.

A unity gain low pass filter with a cut-off frequency of 0.0016 Hz is applied to the discrete summation in (2.1), which is computed with a sampling period of 0.01 seconds. This smoothing operation may be needed for the integration of the model with continuous-time simulation of processes such as the cell cycle, the transcription regulatory network, and metabolic reactions. The simulation results illustrate the model's ability to recapitulate some key signal transduction properties of MAPK signaling cascades, including the amplification of extracellular stimuli and the effect of inhibitory feedback on the relayed signal strength (figure 6). However, since the proposed approach is not reliant on detailed reaction kinetics, the model is not intended to yield high resolution predictions of physiochemical variables such as protein concentrations. Instead, the objective is to achieve an approximation of the temporal complexity of signal transduction dynamics that can yield a physiologically plausible model of the cell communication circuitry. In the next section, a short discussion is provided about the potential contribution of the proposed model to the representation of the cause-effect dynamics linking deregulated signaling pathways and cancer.

IV. Modeling Cellular Signaling and Cancer

The cellular signaling circuitry plays a central role in mediating the complex coupling between the various cellular processes determining cell fate, which is a decisive factor in tumorigenesis. A model of this circuitry must therefore be amenable to integration with the physiological processes that are key determinants of the cellular context, including the cell cycle and the transcription regulatory network. It has been argued that information and energy make up the fundamental forces underlying space-time biological complexity [10]. On this account, the proposed information processing abstraction of cellular signaling may lead to a physiologically plausible model of the relevant cellular context. As a catalyst of this desirable objective, the proposed model focuses on: (1) recapitulating the logic integrating various extracellular and cytoplasmic signals to control the gene expression program behind the determination of cell fate, (2) reflect the structural motifs of the cellular signaling circuitry so as to provide physiologically plausible effector points in the cell communication system. These effector points would serve as the means to induce simulated genetic/epigenetic perturbations and analyse their impact on cell physiology. They may also be used to study the impact of drug inhibitory interventions intended to abrogate the pathological implications of deregulated signaling pathways. In this respect, the proposed approach may be used to study the signaling circuitry underlying cell fate co-regulation by the Ras-ERK and PI3K-mTOR pathways and the modulation of their dynamics by the energy sensing AMPK (figure 1). Despite the high level coupling between the signaling pathways illustrated in figure 1, the proposed FSM-based modeling approach can be applied in a straightforward fashion to yield an integrated model of the entire signaling circuitry at hand. The drastically reduced number of parameters compared to an ODE based kinetics model is a clear advantage in enabling the consideration of a wider scope of interactions with the involvement of multiple signaling pathways. This may provide a deeper insight about the dynamics underlying the regulation of various aspects of cell physiology. Another attractive aspect of the modeling approach being proposed is the model's faithful representation of the signaling protein interaction network and its structural motifs. This may be instrumental in translating model generated insights into actionable therapeutic strategies targeting specific pathway proteins. Further research is however needed to achieve such goal. In particular, a systematic method is needed to estimate the model parameters using experimental data. This would provide a solid anchor for the biological plausibility of the proposed modeling approach.

V. Conclusion

The proposed state machine model of Ras-ERK signaling pathway has been illustrated to reflect some key signal transduction properties of MAPK signaling pathways. These include the amplification of extracellular stimuli and sensitivity to inhibitory feedback. The reliance on a relatively small number of model parameters may enable the analysis of large systems of coupled signaling pathways that otherwise may not be possible using detailed kinetics-based models.

Derbal, Y., "State Machine Modeling of MAPK Signaling Pathways", *2014 IEEE Engineering in Medicine and Biology Conference*, Chicago, Michigan, 2014.

REFERENCES

[1] R. Heinrich, *et al.*, "Mathematical models of protein kinase signal transduction," *Molecular cell*, vol. 9, pp. 957-970, 2002.
[2] J. Vera, *et al.*, "Power-law models of signal transduction pathways," *Cellular Signalling*, vol. 19, pp. 1531-1541, 2007.
[3] R. J. Orton, *et al.*, "Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway," *Biochem J*, vol. 392, pp. 249-61, Dec 1 2005.
[4] B. N. Kholodenko, *et al.*, "Quantification of short term signaling by the epidermal growth factor receptor," *J Biol Chem*, vol. 274, pp. 30169-81, Oct 15 1999.
[5] F. A. Brightman and D. A. Fell, "Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12 cells," *FEBS Lett*, vol. 482, pp. 169-74, Oct 6 2000.
[6] B. Schoeberl, *et al.*, "Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors," *Nat Biotechnol*, vol. 20, pp. 370-5, Apr 2002.
[7] A. R. Asthagiri and D. A. Lauffenburger, "A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model," *Biotechnology progress*, vol. 17, pp. 227-239, 2001.
[8] U. S. Bhalla and R. Iyengar, "Emergent Properties of Networks of Biological Signaling Pathways," *Science*, vol. 283, pp. 381-387, 1999.
[9] A. Goldbeter and D. E. Koshland, Jr., "An amplified sensitivity arising from covalent modification in biological systems," *Proc Natl Acad Sci U S A*, vol. 78, pp. 6840-4, Nov 1981.
[10] Y. Derbal, "On modeling of living organisms using hierrarchical coarse-graining abstractions of knowledge," *Journal of Biological Systems*, vol. 21, p. 1350008, 2013.
[11] A. De Luca, *et al.*, "The RAS/RAF/MEK/ERK and the PI3K/AKT signalling pathways: role in cancer pathogenesis and implications for therapeutic approaches," *Expert opinion on therapeutic targets*, vol. 16 Suppl 2, pp. S17-S27, 2012.
[12] A. S. Dhillon, *et al.*, "MAP kinase signalling pathways in cancer," *Oncogene*, vol. 26, pp. 3279-3290, 2007.
[13] D. Hanahan and R. A. Weinberg, "The Hallmarks of Cancer," *Cell*, vol. 100, pp. 57-70, 2000.
[14] M. C. Mendoza, *et al.*, "The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation," *Trends in Biochemical Sciences*, vol. 36, pp. 320-328, 2011.
[15] C. Shen, "AMPK Phosphorylates and Negatively Regulates BRAF," *Cancer Discovery*, October 10, 2013 2013.
[16] J. L. Jewell and K. L. Guan, "Nutrient signaling to mTOR and cell growth," *Trends Biochem Sci*, vol. 38, pp. 233-42, May 2013.