A Decision Making Model Where the Cell Exhibits Maximum Detection Probability: Statistical Signal Detection Theory and Molecular Experimental Data

Ali Emadi
Dept. of Electrical and Computer Engineering
New Jersey Institute of Technology
Newark, NJ, USA
ae378@njit.edu

Tomasz Lipniacki
Institute of Fundamental Technological Research
Polish Academy of Sciences
Warsaw, Poland
tlipnia@ippt.pan.pl

Andre Levchenko
Yale Systems Biology Institute and Dept. of Biomedical Eng.
Yale University
New Haven, CT, USA
andre.levchenko@yale.edu

Ali Abdi
Dept. of Elect. Comput. Eng. and Dept. of Biological Sciences
New Jersey Institute of Technology
Newark, NJ, USA
ali.abdi@njit.edu

***Abstract*—Molecular noise and signaling abnormalities in biochemical signaling systems in cells affect signaling events and consequently may alter cellular decision making results. Since unexpected and altered cellular decisions may contribute to the development of many pathological conditions and diseases, it is of interest to develop proper models to characterize and measure molecular signal detection parameters and cellular decisions. In this paper and using the Neyman-Pearson signal detection theorem, we propose a signal detection model in which the cell maximizes its signal detection probability in the presence of noise. To evaluate the usefulness of the proposed model, we use measured molecular experimental data of the important TNF–NF-κB cell signaling system. Our results demonstrate that the proposed model provides biologically relevant findings. The introduced Neyman-Pearson-based molecular signal detection framework allows to systematically model and quantify the signal detection behavior and failure of molecular signaling systems, and compute their key decision making parameters such as detection and false alarm probabilities. With regard to the specific TNF–NF-κB system case study in this paper and given the high involvement of the transcription factor NF-κB in cell survival, programmed cell death, immune signaling and stress response, the developed signal detection framework can serve as a useful tool to model the associated cell decision making processes.**

Keywords—cell decision making, biochemical signals, detection theory, Neyman-Pearson detector, A20 deficiency, cancer

I. INTRODUCTION

Characterization of cellular decision making processes is an important chapter in molecular and cell biology. Due to the signal transduction noise or signaling abnormalities or both, a cell may not correctly sense the presence of a signal and therefore may not respond as expected [1]-[3]. This can result in unexpected signaling outcomes, anomalous cellular functions, altered cell fates, and abnormal intercellular processes [4][5], that in turn can contribute to the development of various types of pathological conditions and diseases [6]. It is therefore of interest to model and measure what the signal detection probability is and how it varies with the signal strength.

In a prior study [2], we developed a signal detection theoretical framework to model and measure cell decision making probabilities. In there we assumed the presence and absence of the signal are equally probable, i.e., 50% each, and computed decision making probabilities using a Bayesian approach. However, if the likelihood of the presence of the signal is not known in advance, one needs to consider a different modeling approach. In this study, we introduce a signal detection model where the cell maximizes its signal detection probability, for any given false alarm probability, which is the probability of mistakenly declaring a signal when it is not present. The model relies on the Neyman-Pearson signal detection theorem [7].

To examine the proposed model, we use molecular experimental data [3] of a signaling system through which the tumor necrosis factor, TNF, regulates the transcription factor Nuclear Factor κB (NF-κB), as depicted in Fig. 1. This transcription factor controls many genes involved in cell death and survival, as well as different types of cancer and autoimmune diseases. The molecule A20 in Fig. 1 acts as a negative - inhibitory - feedback, and its deficiency such as mutation or loss can lead to chronic inflammation and cancer [8][9]. For example, experimental data show that mice with A20 deficiency develop inflammation and ultimately die [10].

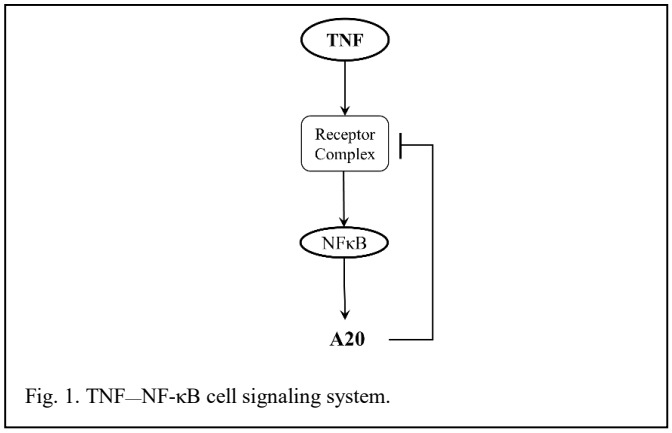

Fig. 1. TNF—NF-κB cell signaling system.

The rest of the paper is organized as follows. The proposed Neyman-Pearson cell signal detection framework and the molecular experimental data of the TNF–NF-κB molecular signaling system are presented in Section II. Molecular data analysis results of the proposed modeling framework are discussed in Section III, and some concluding remarks are given in Section IV.

II. Cell Signal Detection Theory and Molecular Experimental Data: The TNF–NF-κB Signaling System

A. Neyman-Pearson Signal Detection Theorem

In order to model cell signal detection theoretically, one can consider the following hypothesis testing problem:

$$\begin{cases} \mathrm{H}_0 : \text{Cell is receiving noise only,} \\ \mathrm{H}_1 : \text{Cell is receiving a noisy signal.} \end{cases} \quad (1)$$

In the Bayesian approach [7], one needs to know the prior probabilities of the hypotheses, i.e., $p(\mathrm{H}_0)$ and $p(\mathrm{H}_1)$, to be able to determine the decision threshold and then compute probabilities of making incorrect and correct decisions. Since such information is typically not available in cell signaling studies, we previously assumed that the absence and presence of signal are equally probable, i.e., $p(\mathrm{H}_0) = p(\mathrm{H}_1) = 0.5$ [2], to study the cell signal detection performance using the Bayesian framework.

As a different alternative, in this paper we consider the Neyman-Pearson framework [7][11] which does not need the knowledge of the prior probabilities of the two hypotheses, and therefore can serve as a suitable model for cell signal detection studies.

Suppose the vector $\mathbf{x}$ contains N data samples to be used to decide H_0, i.e., no signal, versus H_1, which indicates that there is a signal. Also let $p(\mathbf{x};\mathrm{H}_0)$ and $p(\mathbf{x};\mathrm{H}_1)$ represent probability density functions (PDFs) under H_0 and H_1, respectively. Additionally, let P_{FA} and P_{D} be false alarm and detection probabilities, respectively, where these two events refer to mistakenly declaring a signal when it is not present and correctly declaring a signal when indeed it is present, respectively. The Neyman-Pearson signal detection theorem [7] states that for a given false alarm probability α, a decision making system maximizes P_{D} if it decides H_1 when:

$$L(\mathbf{x}) = \frac{p(\mathbf{x};\mathrm{H}_1)}{p(\mathbf{x};\mathrm{H}_0)} > \gamma . \quad (2)$$

In the above equation, $L(\mathbf{x})$ is the likelihood ratio and γ is the decision threshold, calculated by solving the following equation:

$$P_{\mathrm{FA}} = \int_{\{\mathbf{x}:\, L(\mathbf{x})>\gamma\}} p(\mathbf{x};\mathrm{H}_0)\, d\mathbf{x} = \alpha . \quad (3)$$

Furthermore, the maximized P_{D} can be computed using the following formula:

$$P_{\mathrm{D}} = \int_{\{\mathbf{x}:\, L(\mathbf{x})>\gamma\}} p(\mathbf{x};\mathrm{H}_1)\, d\mathbf{x} . \quad (4)$$

Proof of the Neyman-Pearson signal detection theorem relies on using Lagrange multipliers to find the maximum P_{D} for a given P_{FA} value.

To model the decisions regarding the absence or presence of the TNF signal in Fig. 1 using the Neyman-Pearson signal detection theorem, we use measured nuclear concentrations of NF-κB in N individual cells, for different TNF doses. Details of the molecular measured data are provided in the next subsection.

B. The Molecular Experimental Data

The data set is composed of wild-type and A20$^{-/-}$ (A20-deficient) 3T3-immortalized mouse embryonic fibroblasts [3], representing normal and abnormal cells, respectively. Several TNF doses were examined simultaneously using common reagents, to be able to quantitatively compare responses under different conditions. Nuclear concentrations of NF-κB were

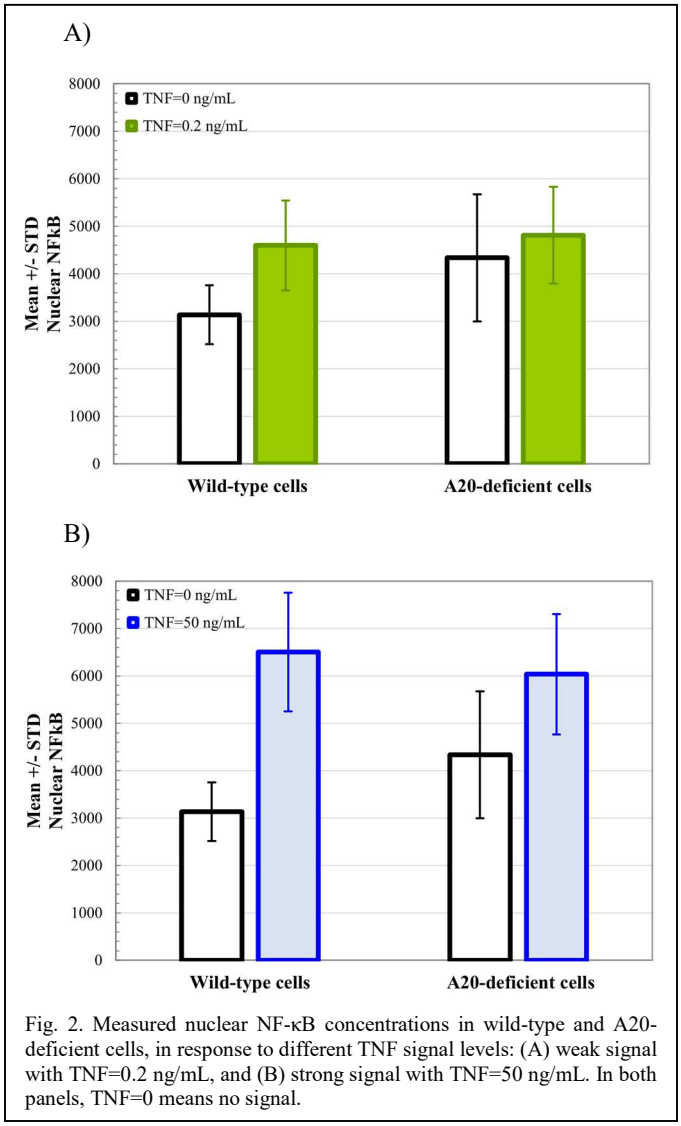

Fig. 2. Measured nuclear NF-κB concentrations in wild-type and A20-deficient cells, in response to different TNF signal levels: (A) weak signal with TNF=0.2 ng/mL, and (B) strong signal with TNF=50 ng/mL. In both panels, TNF=0 means no signal.

assayed using immunocytochemistry in thousands of mouse fibroblasts exposed to different TNF doses [3]. Methods used for quantifying the nuclear NF-κB concentrations can be found in [12].

III. Molecular Data Analysis Results of the TNF–NF-κB Signaling System in the Light of the Neyman-Pearson Signal Detection Theory

A. Average Responses of the Molecular System

To model the H_0 scenario in which there is no signal, we use the NF-κB nuclear concentrations measured in $N = 350$ cells, where the TNF concentration is 0 ng/mL. For modeling the H_1 scenario when a weak signal is present, we use the measured nuclear NF-κB levels of N cells where the TNF concentration is 0.2 ng/mL, whereas to model H_1 to study a strong signal scenario, we use the nuclear NF-κB levels of N cells exposed to the TNF dose of 50 ng/mL.

Fig. 2 presents average NF-κB responses in wild-type and A20-deficient cells for different levels of TNF. Comparison of average levels of the nuclear NF-κB for TNF doses of 0.2 (Fig. 2A) versus 50 ng/mL (Fig. 2B) indicates that a stronger signal generates a stronger response, in both types of normal and abnormal cells.

To understand the fatal role of A20 deficiency, one can compare the normal and abnormal responses for each non-zero TNF level. Inspection of each panel of Fig. 2 reveals that abnormal cells experience less increase in the nuclear NF-κB response, with respect to the associated nuclear NF-κB level for zero TNF (no signal). As we will see in the following subsections, utilizing the Neyman-Pearson signal detection theory allows to systematically model and study the signal detection failure of the TNF–NF-κB and perhaps other molecular systems, and compute key decision making parameters such as detection and false alarm probabilities.

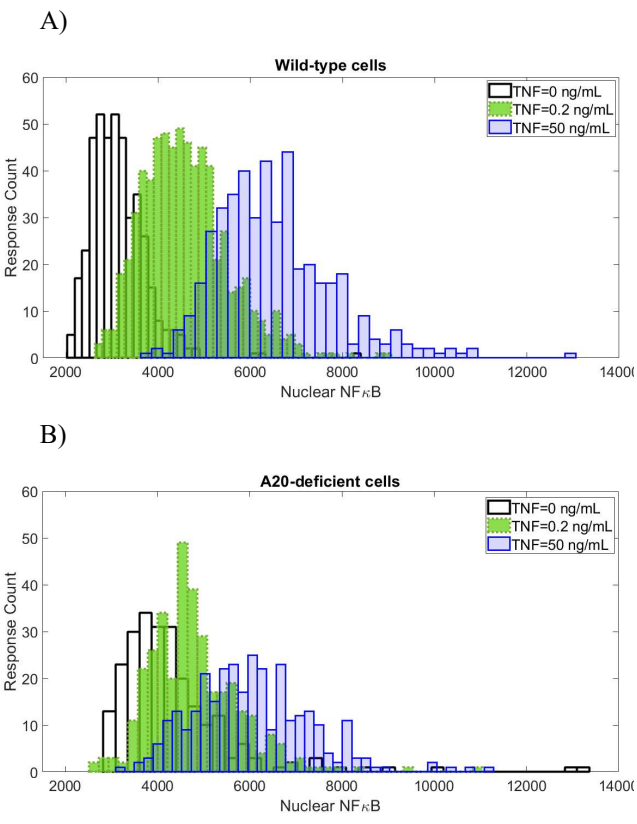

Fig. 3. Histograms of measured nuclear NF-κB concentrations for zero, weak and strong TNF signals: (A) wild-type cells, and (B) A20-deficient cells.

B. Probabilistic Responses of the Molecular System

Fig. 3A shows histograms of the measured nuclear NF-κB concentrations of wild-type cells for no TNF signal, as well as weak and strong TNF signals, that correspond to the TNF doses of 0, 0.2 and 50 ng/mL, respectively. To study how the cellular signal detection behavior is affected in the presence of an abnormality, i.e., A20 deficiency, histograms of the measured nuclear NF-κB concentrations of A20-deficient cells are shown in Fig. 3B for zero, weak and strong TNF signals as well.

We note that in normal cells, Fig. 3A, as the TNF signal becomes stronger, the overlap between the no signal response histogram and the TNF-stimulated response histogram decreases. This indicates smaller decision making errors that can be quantified using the Neyman-Pearson theorem, as demonstrated in the next subsection. The overlap between the response histograms and the associated decision making errors in abnormal cells, Fig. 3B, exhibit more complex patterns that are less likely to be visually understood. As shown in the next subsection, the Neyman-Pearson theorem allows us to compute and discover these complex response overlap patterns in abnormal cells, and compare them with what we see in normal cells. This can assist with understanding the altered decision making outcomes in abnormal cells, with respect to the normal cells.

C. Neyman-Pearson Modeling Results for the Signal Detection Performance of the Molecular System

To understand the signal detection behavior of the TNF–NF-κB molecular system in normal cells, we have computed and graphed its receiver operating characteristic (ROC) curve (Fig. 4, curves with the * marker). In a ROC figure, a useful

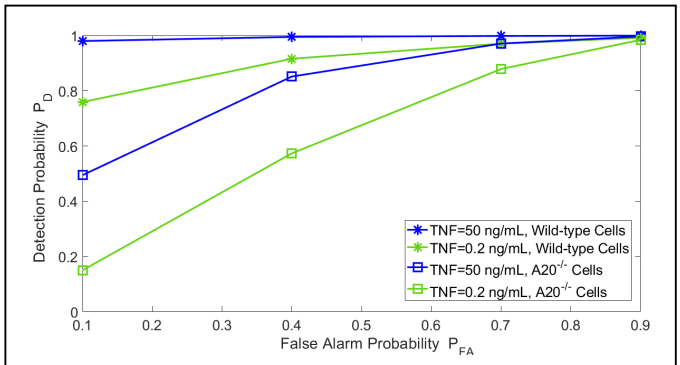

Fig. 4. Receiver operating characteristic (ROC) curves of the TNF–NF-κB system, computed based on the proposed cell decsion making model using the Neyman-Pearson signal detection theorem and molecular experimental data. Curves with the * and □ markers correspond to normal and abnormal cells, respectively.

graphical tool [1][7], the detection probability P_D of a decision making system is graphed versus the false alarm probability P_{FA}, for various decision thresholds. More specifically and for any given P_{FA} in Fig. 4, we have solved Equation (3) and substituted the resulting decision threshold γ in Equation (4), to compute the corresponding maximum P_D. For the PDFs under H_0 and H_1 in (3) and (4), respectively, we have similarly [2][3] used Gaussian distributions that reasonably represent the data. This allows to obtain closed-form and simple-to-use mathematical expressions for the likelihood ratio $L(\mathbf{x})$, the false alarm probability P_{FA} and the detection probability P_D defined in Equations (2), (3) and (4), respectively.

Analysis of Normal Cells: Comparison of the wild-type cell ROC curves (Fig. 4, curves with the * marker), reveal that when the signal is strong, 50 ng/mL TNF, the signal detection probability is higher, compared to the weak signal, and is very close to 1. This is a biologically relevant result, i.e., there is almost no missed signal and no decision mistake, when the signal is strong. This indicates the suitability of the proposed Neyman-Pearson-based model for the cellular signal detection.

Analysis of Abnormal Cells: For abnormal cells (Fig. 4, curves with the □ marker), we note a drop in the ROC curves, compared to the normal cell curves. For example, for the false alarm probability of 10% and when the TNF signal is strong, the detection probability drops from 98% to 49%. For the same false alarm probability of 10% and the weak TNF signal, the detection probability decreases further from 76% to 15%. These detection probability decreases for both strong and weak TNF signals in A20-deficient cells are noteworthy, and demonstrate that the introduced model generates biologically appealing predictions, and quantifies the signal detection capability degradation, when there is a mutation or abnormality in the signaling network. Such signal detection and cellular decision making quantifications are not feasible via inspection of the average responses (Fig. 2), and allow for dissecting out the complex overlap patterns observed among the response histograms (Fig. 3) of the TNF–NF-κB molecular system.

IV. Conclusion

A normal wild-type cell is supposed to properly respond to different received signals and produce correct signaling outcomes and decisions. However, in the presence of noise or mutated signaling components, it may exhibit unexpected and incorrect responses which may eventually develop some pathological conditions. Here and based on the Neyman-Pearson signal detection theory, a molecular signal detection model is developed and applied to the important TNF–NF-κB molecular signaling system.

Analysis of the experimental data using the proposed approach reveals that it gives biologically plausible and quantitative pictures of the complex signal detection behaviors of both normal and abnormal cells. In the TNF–NF-κB system, the model shows that when the TNF signal is strong in normal cells, the signal detection probability is higher, compared to the weak TNF signal detection. This is a biologically reasonable result. The model also provides interesting quantitative predictions regarding signal detection changes in abnormal cells. For example, for the false alarm probability of 10%, we observe a large detection probability reduction from 98% to 49%, even when the TNF signal is strong in A20-deficient cells. Considering the involvement of the mutation or loss of A20 in several pathological conditions [8]-[10], the proposed molecular signal detection model and its predictions may have implications for developing proper therapeutics.

Finally, it should be noted that the developed Neyman-Pearson molecular decision making framework in the presence of noise may be applicable to various other organisms such as bacteria, viruses, yeast, etc. [13].

References

[1] M. Ozen, T. Lipniacki, A. Levchenko, E. S. Emamian & A. Abdi, "Modeling and measurement of signaling outcomes affecting decision making in noisy intracellular networks using machine learning methods," *Integrative Biology*, vol. 12, pp. 122–138, 2020.

[2] I. Habibi, R. Cheong, T. Lipniacki, A. Levchenko, E. S. Emamian & A. Abdi, "Computation and measurement of cell decision making errors using single cell data," *PLoS Computational Biology*, e1005436, 2017.

[3] R. Cheong, A. Rhee, C. J. Wang, I. Nemenman & A. Levchenko, "Information transduction capacity of noisy biochemical signaling networks," *Science*, vol. 334, pp. 354-358, 2011.

[4] A. Emadi, M. Ozen & A. Abdi, "A hybrid model to study how late long term potentiation is affected by faulty molecules in an intraneuronal signaling network regulating transcription factor CREB," *Integrative Biology*, vol. 14, pp. 111-125, 2022.

[5] A. Emadi & A. Abdi, "A study of how abnormalities of the CREB protein affect a neuronal system and its signals: Modeling and analysis using experimental data," in *Proc. IEEE Signal Processing in Medicine and Biology Symp.*, Philadelphia, PA, 2022, pp. 1-6.

[6] A. Sitaramayya, *Signal Transduction: Pathways, Mechanisms and Diseases*. Springer, 2009.

[7] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice-Hall, 1998.

[8] L. Catrysse, M. Farhang Ghahremani, L. Vereecke, S. A. Youssef, C. Mc Guire, M. Sze, et al., "A20 prevents chronic liver inflammation and cancer by protecting hepatocytes from death," *Cell Death & Disease*, vol. 7, e2250, 2016.

[9] S. G. Hymowitz & I. E. Wertz, "A20: From ubiquitin editing to tumour suppression," *Nature Reviews Cancer*, vol. 10, pp. 332-341, 2010.

[10] E. G. Lee, D. L. Boone, S. Chai, S. L. Libby, M. Chien, J. P. Lodolce, et al., "Failure to regulate TNF-induced NF-κB and cell death responses in A20-deficient mice," *Science*, vol. 289, pp. 2350-2354, 2000.

[11] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. John Wiley & Sons, 2001.

[12] R. Cheong, C. J. Wang & A. Levchenko, "High content cell screening in a microfluidic device," *Molecular & Cellular Proteomics*, vol. 8. pp.433-442, 2009.

[13] G. Balazsi, A. van Oudenaarden & J. J. Collins, "Cellular decision making and biological noise: From microbes to mammals," *Cell*, vol. 144, pp. 910-925, 2011.