

Factor Graph-Based Biomolecular Circuit Analysis for Designing Forward Error Correcting Biosensors

Yang Liu, *Student Member, IEEE*, and Shantanu Chakrabartty, *Member, IEEE*

Abstract—We previously reported the fabrication and the verification of novel biomolecular transistors where electrical conductivity of a “polyaniline nanowires” channel is controlled by antigen-antibody interactions. In this paper, we present a simulation framework for analyzing the reliability of biosensor circuits constructed by using these biomolecular transistors. At the core of the proposed framework is a library of electrical circuit models that capture the stochastic interaction between biomolecules and their variability to environmental conditions and experimental protocols. Reliability analysis is then performed by exploiting probabilistic dependencies between multiple circuit elements by using a factor graph-based decoding technique. The proposed computational approach facilitates rapid evaluation of forward error correction (FEC) strategies for biosensors without resorting to painstaking and time-consuming experimental procedures. The analysis presented in this paper shows that an asymmetric FEC biosensor code outperforms a repetition FEC biosensor code which has been proposed for microarray technology. In addition, we also show that the proposed analysis leads to a novel “co-detection” protocol that could be used for reliable detection of trace quantities of pathogens.

Index Terms—Biomolecular circuits, biosensors, computer-aided design (CAD), factor graph, forward error correction co-detection, polyaniline, reliability.

I. INTRODUCTION

BIOMOLECULAR circuits, which operate by converting the binding of an analyte with a biological recognition layer into a measurable signal, form the core of many biosensor systems. Unfortunately, the biomolecular interactions are stochastic in nature and lead to significant variability under small changes in experimental conditions. This is particularly true for affinity-based biosensors [1]. As a result, achieving high reliability for detecting pathogens in the presence of experimental and biosensor artifacts poses a significant challenge. We had proposed the use of the forward error correction (FEC) principle to address this challenge and had demonstrated the functionality of basic biomolecular circuit elements that can be used for constructing an FEC biosensor [2]. However, evaluating the reliability of different FEC biosensor configurations using direct experimental techniques has proven to be impractical due to the

time-consuming and labor-intensive nature of sample preparation and handling. This has motivated us to develop an analytical framework for evaluating the performance of different biosensor FEC topologies and is the main focus of this paper.

The model biosensor chosen to describe the proposed framework is a lateral flow immunosensor which utilizes polyaniline nanowires-based transducers to convert binding between a target antigen and its antibody into an electrical signal. The principle was extended to implement biomolecular logic circuits (AND and OR) and their functionality was verified by using experimental results [2]. In this paper, we will first develop equivalent circuit models that capture the stochastic response of respective biomolecular circuit elements. A similar analytical approach for modeling the biomolecular stochastic interaction has been reported in [1]. In this work, two kinds of biosensor noise sources were identified and modeled: 1) systematic noise that was induced by sample handling errors and errors introduced by device artifacts and 2) random noise due to probabilistic biomolecular binding interactions [3], [4]. In comparison, our approach directly uses a system identification approach for approximating the experimental data using a log-linear model [5] and extracts the parameters of the noise sources based on the regression and random errors.

We then assemble these composite models to simulate the response of different FEC biosensor structures and evaluate the performance of different decoding algorithms for improving the reliability of biosensors in the presence of systematic and random noises. In the literature, FEC principles have been used for improving the reliability of nanoscale systems, which suffer from similar computational artifacts as biosensors. Some of the examples include the design of fault-tolerant circuits [6] and next-generation flash memory [7]. Other work in the related area includes the application of an FEC principle for understanding biological systems [8], [9]. For example, error-correcting codes were proposed for a DNA microarray, where redundant gene spotting was used to reduce drop-out errors [10]. This work was extended in [11] where an error-control coding scheme and quality control were integrated for fabricating DNA microarrays by using multiplexed DNA strands. However, the FEC techniques proposed in the aforementioned work use a simple repetition procedure [10], [11] for which an optimal decoding algorithm was proposed. Repetition techniques have been used in other microarray designs [12], [13] where they have been shown to greatly increase the power of the experiments to measure gene expression changes. In this paper, we will show that for our model biosensor, an asymmetric encoding method and its corresponding factor graph decoding algorithm outperforms the repetition code for achieving higher reliability of detection.

Manuscript received July 15, 2008; revised October 31, 2008. Current version published May 22, 2009. This work was supported by a grant from the National Science Foundation ECCS-0622056. This paper was recommended by Associate Editor Z. Wang.

The authors are with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824 USA.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBCAS.2009.2014247

Factor graph is a generalization of the ‘‘Tanner graph’’ [14] that graphically captures the interdependencies between different variables of a global function. Factor graphs have been shown to be general enough to describe many popular algorithms that are extensively used for learning and signal processing [15]. In this paper, reliability analysis of the biomolecular circuits is performed by using a factor graph message passing approach that produces confidence scores that are proportional to the concentration of pathogens present in the sample. However, unlike the conventional factor graph decoding approach, the analysis presented in this paper uses an asymmetric message passing algorithm which takes into account the type of biomolecular logic functions that can be realized. Our analysis show that an asymmetric code outperforms a repetition code in reducing the number of detection errors. Also, based on a simulation study, we observe and describe a novel ‘‘co-detection’’ protocol that uses the nonlinear encoding-decoding property of the asymmetric code to detect trace quantities of a pathogen in a given sample.

This paper is organized as follows. Section II describes the operating principle of the biomolecular circuits and summarizes the analytical model obtained by using experimental data. Section III describes the construction of biosensor encoder based on the biomolecular circuit models obtained in Section II. Section III also describes the factor graph decoding algorithm that is used for reliability analysis. Section IV presents behavioral simulation results by using the analytical framework and Section V discusses the simulation results in detail with a description of the novel ‘‘co-detection’’ protocol. Section VI concludes this paper with some final remarks.

II. BIOMOLECULAR CIRCUIT MODELS

In this section, we use our previously reported experimental results [2] to extract circuit models that capture their systematic and random noises. The biomolecular circuits were constructed on a porous membrane instead of silicon or other rigid substrate and transporting the analyte through the membrane is achieved by using capillary force. The operating principle of a single biomolecular transistor based on the lateral flow principle is illustrated in Fig. 1 which shows a cross-sectional view of the membrane. The details about the principle can be found in [16] and are briefly described here. The transistor consists of two silver electrodes (equivalent to a source and drain of a metal-oxide semiconductor field-effect transistor (MOSFET)) and a layer of antibody between the electrodes (equivalent to a channel of a MOSFET). The gap between two electrodes on the membrane is open initially. When the analyte is applied on the biosensor, the antigen in the analyte first hybridizes with its target antibody which then conjugates with polyaniline conductive nanowires. The antigen-antibody-polyaniline hybrid then migrates into the membrane gap due to the capillary force and interacts with the immobilized antibodies to form a nanowire bridge across the electrode. The affinity between the antigen-antibody interaction and, hence, the concentration of the nanowires determines the conductance of the channel formed between the electrodes. The scanning electron microscopy (SEM) images shown in each reaction phase (see Fig. 1) show a change in membrane texture and

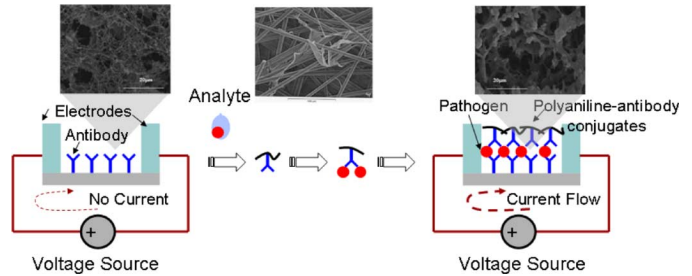


Fig. 1. Principle of a single biomolecular transistor.

it is attributed to the formation of the antibody-antigen-antibody-polyaniline complex connecting the electrodes. We have fabricated and characterized the response of a single biomolecular transistor by using the *B. cereus* antibody with respect to different pathogen concentrations. Fig. 2 shows the measured conductance across the biosensor electrodes as the concentration of pathogen (*B. cereus*) in the sample is varied. The measured conductance is normalized with respect to the ‘‘control’’ conductance (measured when a sample containing no pathogens is applied) and shows a clear discrimination between pathogenic and nonpathogenic cases. The plot in Fig. 2 also shows a monotonic increase in conductance with an increase in pathogen concentration (given in colony forming units per milliliter (CFU/mL)). This response can be approximated by using a log-linear model (shown by a dotted line in Fig. 2), and the regression error is used to approximate the systematic error. Fig. 2 also shows error bars that are computed by using multiple experimental runs, and they are used for estimating the random errors. The systematic and random errors are included as the additive noise component G_n in the equivalent circuit model of the biomolecular transistor (shown in Fig. 3). X_B denotes the concentration of the *B. cereus*, and G_B, G_o are respective conductances whose definitions and large signal models are given in Tables I and II. Based on the large signal models given in Table I, it can be shown that a single biosensor acts like a ‘‘pathogen concentration’’-controlled resistor that is similar to the operation of a MOSFET transistor biased in weak inversion.

The biosensor principle was extended to implement two variants of logical operations (AND and OR) [2]. Fig. 4 shows the structure of an AND gate (marked by 1) and OR gate (marked by 2), constructed by using the antigen-antibody-polyaniline complex. An AND operation is achieved by cascading two different antibodies in between the biosensor electrodes. Thus, in an ideal condition, conduction between the electrodes occurs only when both pathogens are present in the sample [for completing the polyaniline bridge as shown in Fig. 4 (1)]. An OR operation is achieved by immobilizing a mixture of antibodies between the electrodes. Thus, in an ideal condition, a polyaniline nanowire bridge is formed when either of the antigens is present.

Figs. 5 and 6 show the transient responses measured by using the AND and OR logic gates. The conductance measured across the electrodes stabilizes 100 s after the application of the analyte and constitutes the steady state of the biosensor logic gates. The transient behavior is attributed to the dynamics of the polyaniline sandwich in the presence of analyte flow, adhesion, and capillary force. However, in this paper, only the steady-state conductance will be used for modeling and the derivation of tran-

TABLE I
EQUIVALENT CIRCUIT EQUATIONS FOR FUNDAMENTAL BUILDING BLOCKS

Fundamental building blocks	Circuit Equation
Single biomolecular transistor	$G(X) = G_0 + \kappa \log \frac{X}{X_0} + G_n$
OR logic gate	$G(X_B, X_E) = G_{OR} + \kappa_B^{OR} \log \frac{X_B}{X_{0B}^{OR}} + \kappa_E^{OR} \log \frac{X_E}{X_{0E}^{OR}} + G_n$
AND logic gate	$G(X_B, X_E) = G_{AND} + \kappa_B^{AND} \log \frac{X_B}{X_{0B}^{AND}} + \kappa_E^{AND} \log \frac{X_E}{X_{0E}^{AND}} + \kappa_{EB}^{AND} \log \frac{X_E + X_B}{X_{0EB}^{AND}} + G_n$

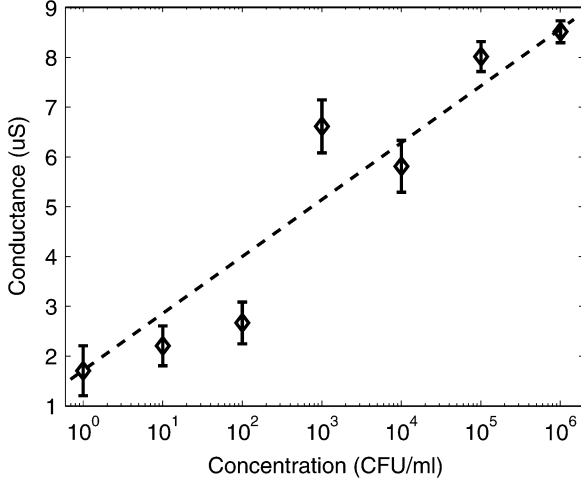


Fig. 2. Conductance measurement of a *B. cereus* single biomolecular transistor.

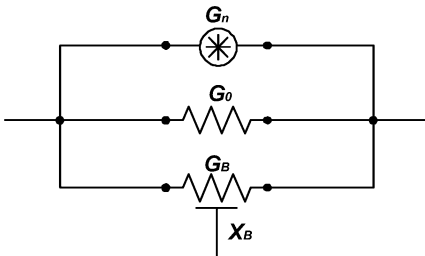


Fig. 3. Circuit model of a single biomolecular transistor.

sient models is deferred for future publications. Fig. 7 shows that the steady-state conductance for the AND gate is highest when both pathogens (*B. cereus* and *E. coli*) are present in the analyte compared to the conductance measured when either one or none of the pathogens are present. The response of the OR gate (Fig. 8) shows that the stabilized conductance is almost the same when either of the pathogens or both of them are present while the “control” condition leads to a lower conductance compared to other logical states. The responses of the logic gates for different concentrations of pathogens are also shown in Figs. 7 and 8 and have been used to derive their equivalent circuit models that incorporate the inherent noise sources. The corresponding circuit models for OR and AND gates are shown in Fig. 9 and their respective mathematical representations are provided in Table I. Table II summarizes typical values of the model parameters that have been extracted by using the measured results and have been used for simulations presented in this paper.

TABLE II
PARAMETER MEANING AND TYPICAL VALUES OF TABLE I

Parameter	Meaning	Value
G_0	“control” transconductance	$1.24 \mu S$
κ	sensitivity factor	$1.2 \mu S$
X_0	detection constant	$0.1 CFU/mL$
G_{OR}	“control” transconductance	$13.6 \mu S$
κ_B^{OR}	sensitivity factor of <i>B. cereus</i>	$0.15 \mu S$
X_{0B}^{OR}	detection constant for <i>B. cereus</i>	$0.76 \mu S$
κ_E^{OR}	sensitivity factor of <i>E. coli</i>	$0.09 \mu S$
X_{0E}^{OR}	detection constant for <i>E. coli</i>	$8.5 \times 10^{-4} CFU/mL$
G_{AND}	“control” transconductance	$13.1 \mu S$
κ_B^{AND}	sensitivity factor of <i>B. cereus</i>	$3.4 \mu S$
X_{0B}^{AND}	detection constant for <i>B. cereus</i>	$10^3 CFU/mL$
κ_E^{AND}	sensitivity factor of <i>E. coli</i>	$0.45 \mu S$
X_{0E}^{AND}	detection constant for <i>E. coli</i>	$4.6 \times 10^2 CFU/mL$
κ_{EB}^{AND}	sensitivity factor of coupling effect	$0.4 \mu S$
X_{0EB}^{AND}	detection constant of coupling effect	$1.2 \times 10^3 CFU/mL$
G_n	conductance induced by noise	variable

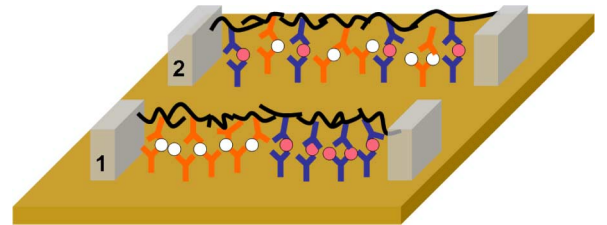


Fig. 4. Schematic illustration of biomolecular logic gates [AND gate (marked by 1), OR gate (2)].

III. DESIGN AND SIMULATION OF FORWARD ERROR CORRECTING BIOSENSOR CIRCUITS

In this section, we utilize the biomolecular circuit models to evaluate different FEC topologies that can improve the reliability of the model biosensor. A system-level architecture of an FEC biosensor is shown in Fig. 10, which consists of an encoder that is an ensemble of N biomolecular circuits that convert the biological binding into a change in conductance, a channel that introduces random and systematic errors, and a decoder that uses the noisy measurements to produce probability estimates indicating the presence or absence of pathogens in

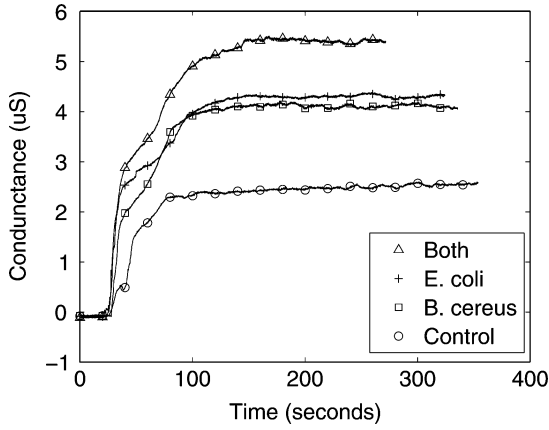


Fig. 5. Typical transient response of the AND gate by using model pathogens (*B. cereus* and *E. coli*).

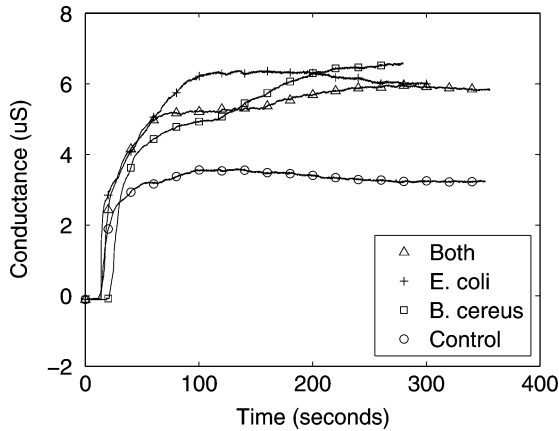


Fig. 6. Typical transient response of the OR gate by using model pathogens (*B. cereus* and *E. coli*).

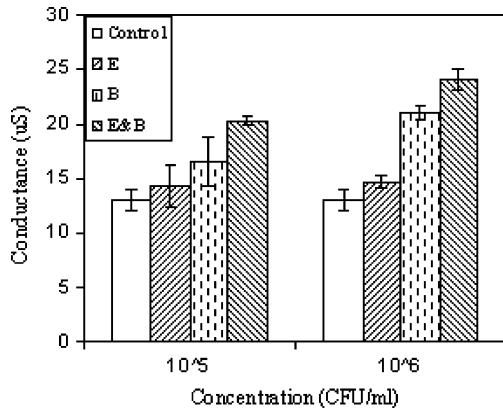


Fig. 7. Conductance measurement of the AND gate by using model pathogens (*B. cereus* and *E. coli*).

a sample. The decoder processes the noisy measured conductances $G_i, i = 1, \dots, N$ and produces posteriori probability estimates $P(X_k | G_1, \dots, G_N), k = 1, \dots, N$, where $X_k \in \{0, 1\}$ is a Boolean variable corresponding to the logical operation of the k th biomolecular circuit. In the mathematical treatment that follows, it will be assumed that the channel is the only source of randomness and the encoding/decoding operation is perfectly reliable.

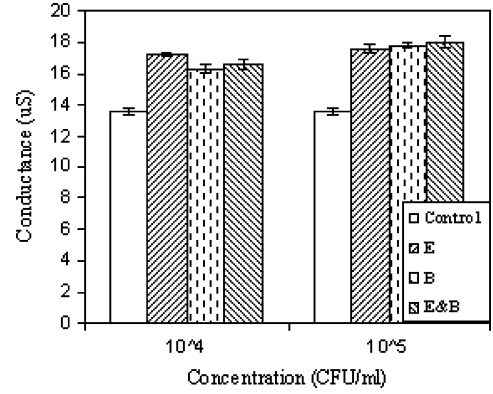


Fig. 8. Conductance measurement of the OR gate by using model pathogens (*B. cereus* and *E. coli*).

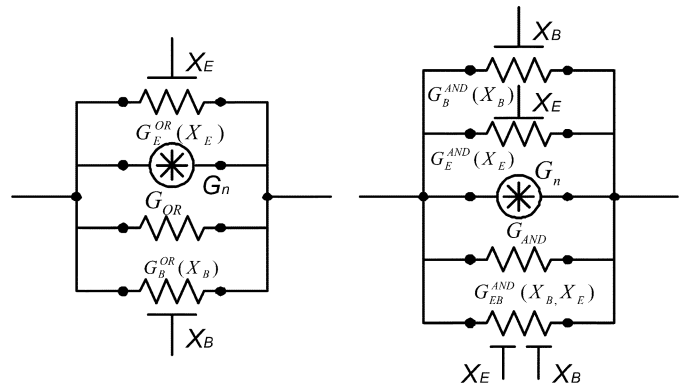


Fig. 9. Circuit model of the OR and AND gates.

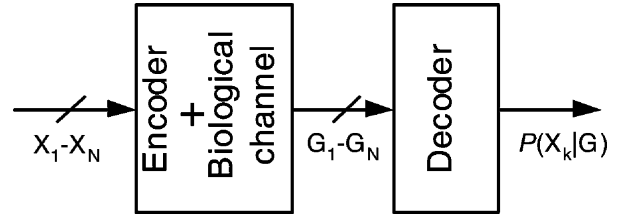


Fig. 10. Framework of FEC biosensors.

The posteriori estimate $P(X_k | G_1, \dots, G_N)$ can be simplified according to

$$\begin{aligned}
 P(X_k | G_1, \dots, G_N) &= \propto \sum_{\sim X_k} P(X_1, \dots, X_N, G_1, \dots, G_N) \\
 &= \sum_{\sim X_k} \prod_{j=1}^N P(G_j | X_k) P(X_1, \dots, X_N) \\
 &= \sum_{\sim X_k} f(X_1, \dots, X_N) \prod_{j=1}^N P(G_j | X_k) \quad (1)
 \end{aligned}$$

where $\prod_{j=1}^N P(G_j | X_k)$ models the response of the biomolecular circuit elements and $f(X_1, \dots, X_N)$ represents a Boolean function that captures the logical dependency between the variables $X_k, k = 1, \dots, N$ and, hence, models the structure of the

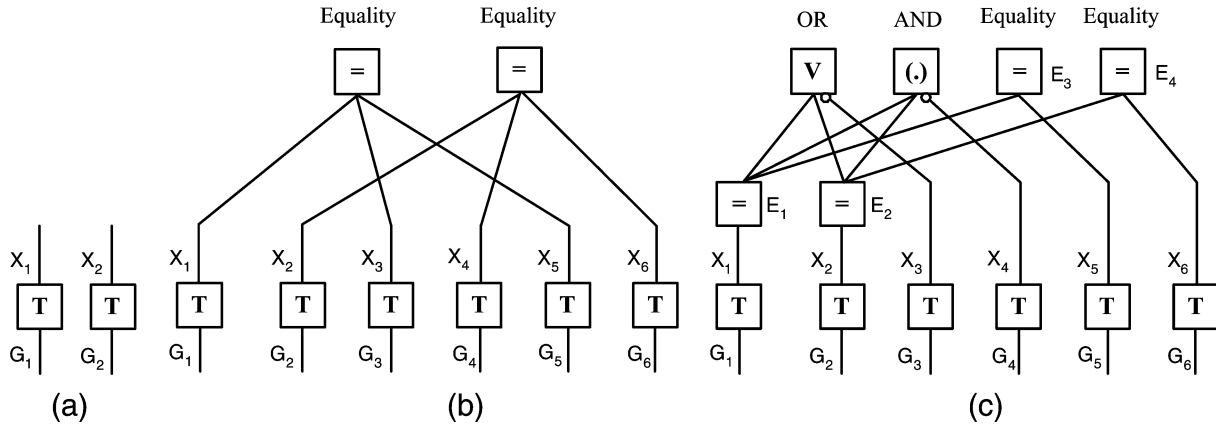


Fig. 11. Forney-style factor graph models of FEC biosensors. (a) Uncoded biosensor. (b) (6,2) biosensor repetition code. (c) (6,2) biosensor asymmetric code.

TABLE III
(6,2) REPETITION CODE

X_1	X_2	X_3	X_4	X_5	X_6
0	0	0	0	0	0
0	1	0	1	0	1
1	0	1	0	1	0
1	1	1	1	1	1

TABLE IV
(6,2) ASYMMETRIC CODE

X_1	X_2	X_3	X_4	X_5	X_6
0	0	0	0	0	0
0	1	1	0	0	1
1	0	1	0	1	0
1	1	1	1	1	1

encoder. The encoder function $f(\cdot)$ can be represented in a tabular form where each table entry represents a state of the variables X_1, \dots, X_N for which $f(X_1, \dots, X_N) = 1$ and zero otherwise. We will provide some specific examples of the encoder function in the next section. Equation (1) also describes the decoding algorithm used for computing the *a posteriori* probability estimates. However, marginalization of a large number of variables according to (1) is computationally intensive [15] and a factor graph-based decoding algorithm is therefore used in this paper.

A. Biomolecular Circuit Encoder

The simplest encoding method is the “repetition” code where biomolecular transistors that detect single pathogen are replicated multiple times. For example, biomolecular transistors specific to two model pathogens *B. cereus* and *E. coli* are replicated three times, respectively. In this case, X_1, \dots, X_6 will be used to represent the Boolean variables corresponding to the output of each biomolecular transistor. The resulting encoder function $f(X_1, X_2, \dots, X_6)$ is summarized in Table III. This encoder is denoted as a (6,2) repetition code which implies that six measurements are independently performed (sequentially or in parallel) to detect two possible pathogens.

Another form of encoding function that will be the focus of this paper uses the biomolecular OR and AND logic circuits. In conventional FEC codes used in communications and storage systems, XOR operation is utilized to obtain linear codes which by construction are symmetric. Unfortunately, XOR logic using the proposed biosensor principle is unrealizable and, hence, our encoding (also referred to as an asymmetric code) function will only be based on OR and AND biomolecular circuits. One specific instance of a (6,2) encoder function is summarized in Table IV, where X_1, X_2 are Boolean variables corresponding to the absence or presence of *B. cereus* and *E. coli*. The variable X_3 represents a logical OR operation between X_1 and X_2 and variable

X_4 corresponds to a logical AND operation. The variables X_5 and X_6 are repetition of variables X_1 and X_2 . The structure of the code can also be conveniently represented as a factor graph (shown in Fig. 11) which has also been used for the decoding procedure.

B. Factor Graph Model and Soft Decoding Algorithm

The factor graph representations are used for efficient computation of the *a posteriori* probabilities by marginalizing variables according to (1). Computation on factor graphs proceeds by using distributed algorithms which are known as the “sum-product” algorithms in the literature [17]. In this section, we will describe the “sum-product” message passing algorithm for factor graphs corresponding to the biomolecular circuits. For a general treatment of message passing algorithms, the readers can be referred to [15].

A Forney-style factor graph [18] corresponding to an uncoded biosensor, a (6,2) repetition code biosensor, and a (6,2) asymmetric code biosensor is shown in Fig. 11(a)–(c). For the uncoded case, the factor graph in Fig. 11(a) consists of two transducer nodes (T) whose inputs are the conductance measured from two biosensors specific to *E. coli* and *B. cereus*. The transducer node captures the relationship between the measured conductances G_1, \dots, G_N and the indicator variables X_1, \dots, X_N . Since estimations of the presence of pathogens are directly based on the measurement, there is no coupling between the two transducer nodes, implying that the detection of pathogens (*E. coli* and *B. cereus*) is performed independently. For the (6,2) repetition code whose factor graph is shown in Fig. 11(b), the transducer nodes corresponding to each of the two pathogens are repeated twice. In this case, however, the measurements are coupled and the dependency is depicted in the factor graph using Equality nodes (=). In the (6,2) asymmetric code factor graph shown in Fig. 11(c), some of the transducer nodes also model the biomolecular OR and

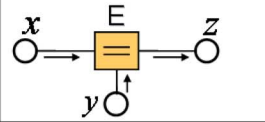
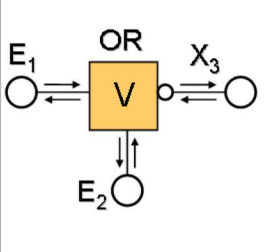
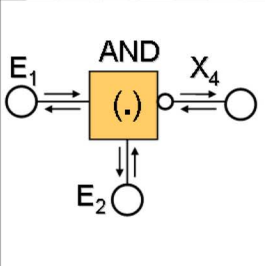
	$\begin{pmatrix} \mu_{E \rightarrow z}(0) \\ \mu_{E \rightarrow z}(1) \end{pmatrix} = \begin{pmatrix} \mu_{x \rightarrow E}(0)\mu_{y \rightarrow E}(0) \\ \mu_{x \rightarrow E}(1)\mu_{y \rightarrow E}(1) \end{pmatrix}$
	$\begin{pmatrix} \mu_{OR \rightarrow E_1}(0) \\ \mu_{OR \rightarrow E_1}(1) \end{pmatrix} = \begin{pmatrix} \mu_{E_2 \rightarrow OR}(0)\mu_{X_3 \rightarrow OR}(0) + \mu_{E_2 \rightarrow OR}(1)\mu_{X_3 \rightarrow OR}(1) \\ \mu_{E_2 \rightarrow OR}(0)\mu_{X_3 \rightarrow OR}(1) + \mu_{E_2 \rightarrow OR}(1)\mu_{X_3 \rightarrow OR}(1) \end{pmatrix}$ $\begin{pmatrix} \mu_{OR \rightarrow E_2}(0) \\ \mu_{OR \rightarrow E_2}(1) \end{pmatrix} = \begin{pmatrix} \mu_{X_3 \rightarrow OR}(0)\mu_{E_1 \rightarrow OR}(0) + \mu_{X_3 \rightarrow OR}(1)\mu_{E_1 \rightarrow OR}(1) \\ \mu_{X_3 \rightarrow OR}(0)\mu_{E_1 \rightarrow OR}(1) + \mu_{X_3 \rightarrow OR}(1)\mu_{E_1 \rightarrow OR}(1) \end{pmatrix}$ $\begin{pmatrix} \mu_{OR \rightarrow X_3}(0) \\ \mu_{OR \rightarrow X_3}(1) \end{pmatrix} = \begin{pmatrix} \mu_{E_1 \rightarrow OR}(0)\mu_{E_2 \rightarrow OR}(0) \\ \mu_{E_1 \rightarrow OR}(0)\mu_{E_2 \rightarrow OR}(1) + \mu_{E_1 \rightarrow OR}(1)\mu_{E_2 \rightarrow OR}(0) + \mu_{E_1 \rightarrow OR}(1)\mu_{E_2 \rightarrow OR}(1) \end{pmatrix}$
	$\begin{pmatrix} \mu_{AND \rightarrow E_1}(0) \\ \mu_{AND \rightarrow E_1}(1) \end{pmatrix} = \begin{pmatrix} \mu_{E_2 \rightarrow AND}(0)\mu_{X_4 \rightarrow AND}(0) + \mu_{E_2 \rightarrow AND}(1)\mu_{X_4 \rightarrow AND}(0) \\ \mu_{E_2 \rightarrow AND}(0)\mu_{X_4 \rightarrow AND}(0) + \mu_{E_2 \rightarrow AND}(1)\mu_{X_4 \rightarrow AND}(1) \end{pmatrix}$ $\begin{pmatrix} \mu_{AND \rightarrow E_2}(0) \\ \mu_{AND \rightarrow E_2}(1) \end{pmatrix} = \begin{pmatrix} \mu_{E_1 \rightarrow AND}(0)\mu_{X_4 \rightarrow AND}(0) + \mu_{E_1 \rightarrow AND}(1)\mu_{X_4 \rightarrow AND}(0) \\ \mu_{E_1 \rightarrow AND}(0)\mu_{X_4 \rightarrow AND}(0) + \mu_{E_1 \rightarrow AND}(1)\mu_{X_4 \rightarrow AND}(1) \end{pmatrix}$ $\begin{pmatrix} \mu_{AND \rightarrow X_4}(0) \\ \mu_{AND \rightarrow X_4}(1) \end{pmatrix} = \begin{pmatrix} \mu_{E_1 \rightarrow AND}(0)\mu_{E_2 \rightarrow AND}(0) + \mu_{E_1 \rightarrow AND}(0)\mu_{E_2 \rightarrow AND}(1) + \mu_{E_1 \rightarrow AND}(1)\mu_{E_2 \rightarrow AND}(0) \\ \mu_{E_1 \rightarrow AND}(1)\mu_{E_2 \rightarrow AND}(1) \end{pmatrix}$

Fig. 12. Sum-product message updated rules of Equality function node (top), OR node (center), and AND node (bottom).

AND circuits. Therefore, the edges in the factor graph that represent the functional dependencies between the nodes connect the pathogen indicator variables (X_1, \dots, X_6) using the AND, OR, and Equality nodes. Note that the AND and OR nodes are connected to exactly three edges, whereas the Equality nodes are connected to at least two edges.

In the biomolecular factor graph decoding algorithm, each node propagates messages along the edge to each of its immediate neighbors. These messages take the form of probability estimates that the node to which the message is being sent to is in state 0 or 1. For example, consider an Equality node E that is connected to three adjacent nodes x, y, z (shown in Fig. 12). The node E receives messages from nodes x and y denoted by $(\mu_{x \rightarrow E}(0), \mu_{x \rightarrow E}(1))^T$ and $(\mu_{y \rightarrow E}(0), \mu_{y \rightarrow E}(1))^T$. It then computes the message sent to node z , denoted by $(\mu_{E \rightarrow z}(0), \mu_{E \rightarrow z}(1))^T$ according to the equality constraints which ensures that there are only two valid states $(x, y, z) = (0, 0, 0), (1, 1, 1)$. The corresponding message passing rules for the Equality node are summarized in Fig. 12. Also summarized in Fig. 12 are message passing rules that correspond to OR and AND node. Since OR and AND operations are asymmetric (as opposed to an XOR operation that is symmetric) with respect to their inputs, the figure describes two sets of rules based on the direction of the message flow. The asymmetric message schedule is unique to the proposed FEC biosensor since it only uses AND, OR, and Equality logic functions for computation.

Decoding using the factor graph models shown in Fig. 11 begins by initializing the transducer nodes using the conductance measurements obtained from the biomolecular circuits. The transducer nodes first normalize the measurements according to

$$\mu_{G_k \rightarrow X_k}(1) = \frac{\exp^{\beta_k(G_k - G_0)}}{1 + \exp^{\beta_k(G_k - G_0)}} \quad (2)$$

where $\beta_k \in [0, 1]$ is a scaling factor that is heuristically chosen for the transducer element k . These normalized measurements are used as messages that are sent to the neighboring Equality, AND, and OR nodes. The Equality, AND, and OR nodes also compute messages locally and transmit them to their neighbors. Messages are propagated back and forth between the nodes for a predetermined number of iterations before a decision on the Boolean variables X_1, \dots, X_6 is made [15]. In algorithm 1, we summarize the complete message passing algorithm that is specific to the factor graph model shown in Fig. 11(c).

Algorithm 1 Calculate $P(X_1 | G_1, \dots, G_6)$, $P(X_2 | G_1, \dots, G_6)$

Require: Factor graph containing nodes X_1, X_2 and G_1, \dots, G_6 .

Require: Measured conductances G_1, \dots, G_6 .

Ensure: For each pair of nodes (i, j) connected by an edge in the factor graph initializes the messages from node i to j according to $(\mu_{i \rightarrow j}(0), \mu_{i \rightarrow j}(1)) = (0.5, 0.5)$ except for messages from G_k to X_k that is initialized as $(\mu_{G_k \rightarrow X_k}(0), \mu_{G_k \rightarrow X_k}(1)) = (1 - P(X_k | G_k), P(X_k | G_k))$ according to (2). No messages are passed from X_k to G_k .

Ensure: Number of iterations $N = 1$

while $N \leq TotalIterations$ **do**

For each direction along the edge of the factor graph, calculate the messages according to the sum-product message updated rules shown in Fig. 12.

end while

$$P(X_1 = 1 | G_1, \dots, G_6) \leftarrow \mu_{G_1 \rightarrow X_1}(1)\mu_{E_1 \rightarrow X_1}(1)$$

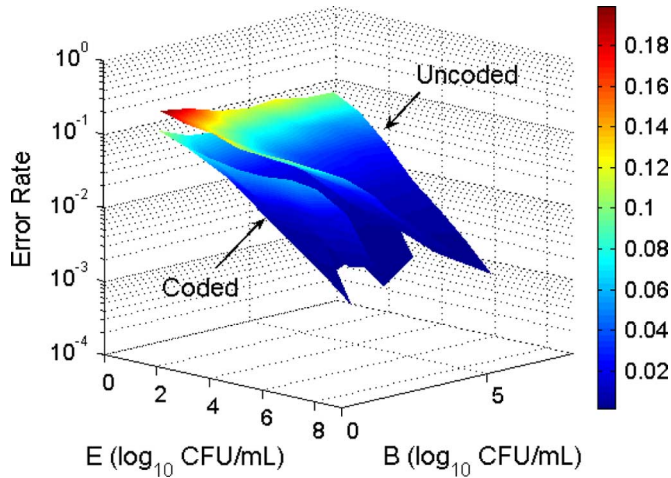


Fig. 13. DER curve of (6,2) repetition code. Each point in the error curve is based on multiple biosensor experiments for a given pathogen (*B. cereus* and *E. coli*) concentration.

return Normalized $P(X_1 = 1 | G_1, \dots, G_6)$.

$P(X_2 = 1 | G_1, \dots, G_6) \leftarrow \mu_{G_2 \rightarrow X_2}(1) \mu_{E_2 \rightarrow X_2}(1)$

return Normalized $P(X_2 = 1 | G_1, \dots, G_6)$.

IV. RESULTS

In this section, we present results that were obtained by using the algorithm 1 when applied to the factor graphs shown in Fig. 11. The biosensor encoder was simulated by using the biomolecular circuit models summarized in Table I. For different concentrations of pathogens (*B. cereus* and *E. coli*), these models produced conductance parameters that were then corrupted by measurement noise. The noise was modeled as a zero-mean additive white Gaussian noise (AWGN) whose variance was experimentally determined according to the procedure described in Section II. The noisy conductance parameters G_1, \dots, G_k were then presented as inputs to the factor graph model, and the probabilities of the presence of *B. cereus* and *E. coli* were estimated by using the message passing algorithm 1. The estimated probabilities were compared against a predetermined threshold (0.5) to obtain a yes/no answer indicating the presence or absence of pathogens in the sample. The simulation experiments were repeated 1000 times for each pathogen concentration level, and the detection error rate (DER) was determined by the occurrence of false negative and false positive.

Fig. 13 shows a two-dimensional DER (error rate for *E. coli* + error rate for *B. cereus*) curves obtained for a (6,2) biosensor repetition code (represented by the factor graph model in Fig. 11(b)) and compares it against the DER curves obtained for the uncoded case [represented by the factor graph model in Fig. 11(a)]. As expected, the DER is reduced when the concentration of pathogen (represented in CFU/mL) increases. Also as expected, the DER for the repetition code (due to larger redundancy) is lower than that of the uncoded case. Similar improvement is also obtained when a (6,2) biosensor asymmetric code (represented by the factor graph model in Fig. 11(c)) is compared against the uncoded case and it is shown in Fig. 14. Fig. 15

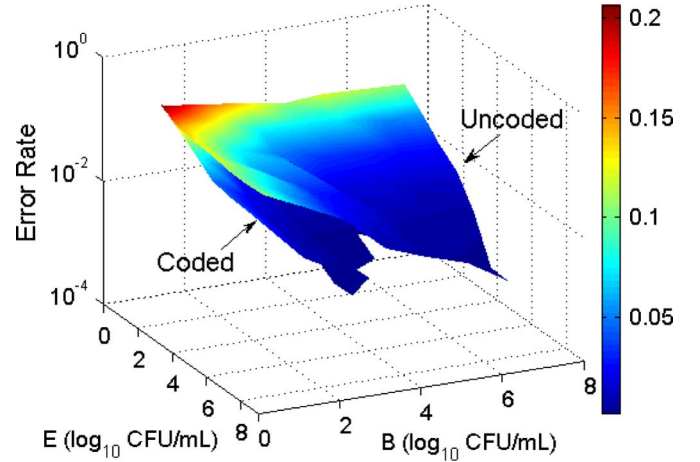


Fig. 14. DER curve of (6,2) asymmetric code.

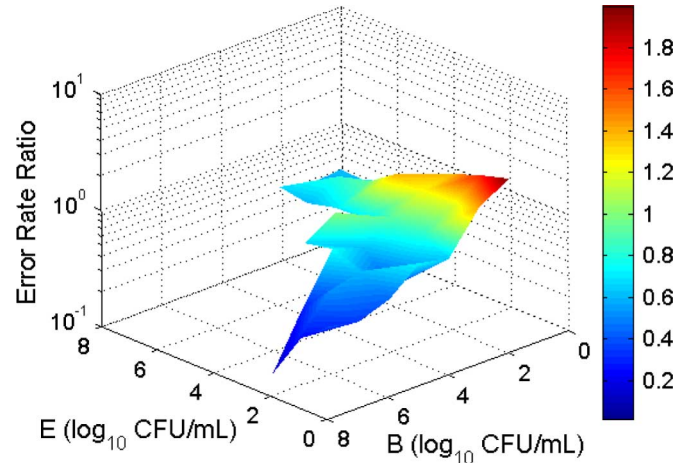


Fig. 15. Comparison of (6,2) asymmetric code and (6,2) repetition code. The error curve shows the ratio of the DER corresponding to the asymmetric code with the repetition code.

shows the ratio of the DER corresponding to the asymmetric code with the repetition code and demonstrates that except for ultra-low pathogen concentration levels, the performance of the asymmetric code biosensor is superior to that of the repetition code biosensor by a factor of 5.

Moreover, compared to the repetition code, an asymmetric biosensor code reveals a nonlinear relationship between two pathogen detections, which improves the performance of the biosensor in terms of reduced error rates. The nonlinear coupling relationship between two pathogen detection can be clearly seen if the DER for each pathogen (instead of total DER) is separately projected on a 2-D plot. This is shown in Fig. 16(a) and (b) for a repetition biosensor code and in Fig. 16(c) and (d) for the asymmetric biosensor code. Fig. 16(a) shows the colormap of the DER corresponding to *B. cereus* illustrating that the DER is independent of the *E. coli* concentration which is expected since there is no coupling between the two detection mechanisms. A similar DER plot for the *E. coli* is shown in Fig. 16(b). However, equivalent plots for the asymmetric code shown in Fig. 16(c) and (d) demonstrate strong coupling between the concentration of *E. coli*/*B. cereus*

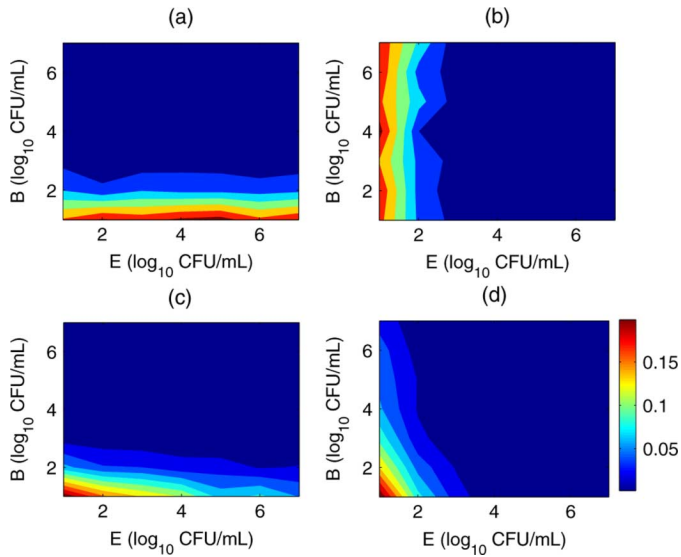


Fig. 16. 2-D projection of the DER curve. (a) and (b) Repetition code. (c) and (d) Asymmetric code. (a) and (c) represent the detection error rate of pathogen *B. cereus* where the nonlinear detection relationship between two pathogens is revealed in (c). (b) and (d) represent the detection error rate of pathogen *E. coli* where a nonlinear detection relationship between two pathogens is revealed in (d).

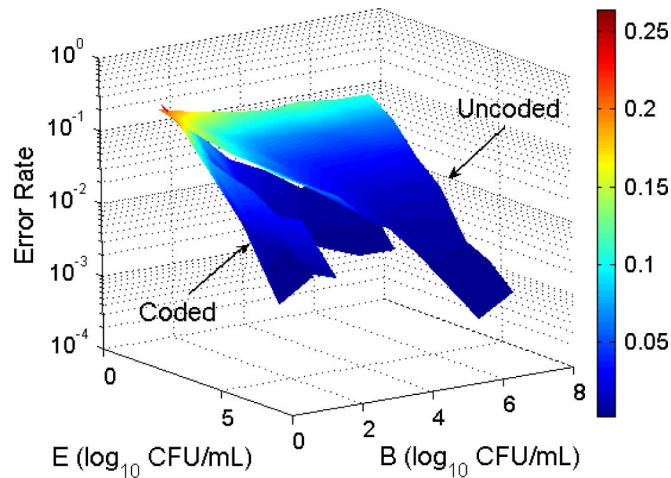


Fig. 17. DER curve of (10,2) extended asymmetric code.

in the input sample and the DER corresponding to *B. cereus*/*E. coli*. This nonlinear coupling relationship is unique to the proposed asymmetric biosensor codes because of constructed AND and OR biomolecular logic circuits. Also, it suggests a novel detection principle that a large concentration of one pathogen could, in fact, improve the detection performance of trace quantities of another pathogen if we properly design the biosensor encoder and decoder. We refer to this mutual coupling as the “co-detection” principle and it represents one of the benefits of a developed simulation environment where the different encoding-decoding techniques could yield novel methods of improving the reliability of biosensors.

We have also conducted experiments with different sizes of asymmetric code and demonstrated that the DER will consistently improve as the size of codes increases. It is illustrated in Fig. 17 which shows that the DER obtained for a (10,2) extended asymmetric code with the structure is shown in Table V

TABLE V
(10,2) ASYMMETRIC CODE

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
0	0	0	0	0	0	0	0	0	0
0	1	1	0	0	1	1	0	0	1
1	0	1	0	1	0	1	0	1	0
1	1	1	1	1	1	1	1	1	1

and compares it with an uncoded case. This illustrates that consistent improvement could be obtained if the asymmetric code is applied to large-scale immunosensors similar to DNA microarrays.

V. DISCUSSIONS

Several conclusions can be drawn from the simulation results presented in Section IV. First, embedding an encoding scheme, such as a repetition code on the biosensor, improves its reliability compared to the case when no encoding is applied. However, an equivalent asymmetric code offers better performance in terms of reduced DER as shown in Fig. 15. The error rate of repetition code is higher than that of the asymmetric code except for low concentrations of both pathogens. The deviation can be attributed to imperfect modeling of the logic gates (AND and OR) due to limited experimental data, especially at low pathogen concentration levels. We also believe that improving the response of logic gates (AND and OR) would improve the performance of the asymmetric code.

The use of logic gates in the biosensor encoder synthetically introduces the coupling between multiple conductance measurements. In contrast to most biosensor designs, where the objective is to obtain independent measurements and in the process suppress any cross-reactive phenomena. In our previous work, we demonstrated that a nonlinear classifier (support vector machine) can exploit the nonlinear interaction between pathogens and their target/nontarget antibodies to improve the detection performance. However, the training complexity significantly increases when the classifier has to model the side information present at the output of the biosensor logic gates. We believe that incorporating cross-reactive principles in the FEC biosensor encoder would enhance the side information available to the decoder to improve the detection reliability. The nonlinear side information also leads to the previously referred principle called “co-detection.” In the co-detection principle, a large concentration of known pathogen improves the detection of trace quantities of unknown pathogens. We attribute this effect to the nonlinear properties of the AND/OR gate formed by using antibodies corresponding to the known and unknown pathogen. When a large quantity of known pathogens is added, the conductive polyaniline bridge between the electrodes is partially formed. Thus, completion of the bridge could be achieved even when trace quantities of an unknown pathogen are present. In ordinary biosensor platforms, the probability of the occurrence of false positive might increase using the aforementioned treatment. This did not occur in our proposed asymmetric FEC biosensor thanks to the encoding/decoding principle. An experimental protocol that uses “co-detection” would therefore: 1) first identify easy-to-detect pathogens (pathogens that have high concentration levels in the sample);

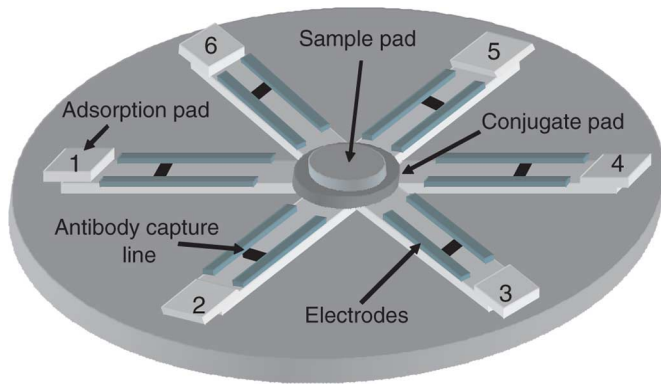


Fig. 18. Visualization of the multipathogen biosensor compact disc (CD) that can implement the encoding methods.

2) fabricate an asymmetric FEC biosensor that includes the antibody specific to identified antigen, then intentionally add large quantities of the identified pathogen into the sample which will enable trace detection of other pathogens using the “co-detection” protocol; and 3) repeat the procedure until all of the pathogens of interest have been screened.

The simulation study also show that the reliability of the FEC biosensor improves when the size of the asymmetric code is increased. Thus, a tradeoff exists between the improved reliability of the biosensor and redundant biomolecular circuit elements that have been added, which is also related to the biosensor cost. Future research will examine an equivalent concept of the “Shannon-Limit” that is applicable to biosensors. In this regard, a fundamental limit might exist, namely, the “biosensor channel capacity,” which could be a function of the detection technology being employed and determines the minimum amount of redundancy required to obtain perfect reliability. In this regard, the nature of the biological channel needs to be further investigated to model and understand the stochastic biomolecular interaction and how it contributes to the biosensor noise.

For the lateral-flow immunosensor that was taken as a model biosensor in this study, the FEC principle can be realized by adding redundant paths for the flow of analyte toward the biomolecular logic gates. One possible realization is illustrated in Fig. 18, where the sample is first applied to a sample pad and a conjugate pad where the antigen-antibody-polyaniline complex is formed. The complex then splits into parallel flow paths and propagates to different antibody capture lines where the biomolecular logic gates/transistors are immobilized. Conductimetric potentiostats [19], [20] are then used to measure the conductance across the electrodes of the biomolecular circuits. The measured conductance is then processed by a digital signal processor or analog decoder chip, which implements the factor graph decoding algorithm and flags the presence or absence of target pathogens.

VI. CONCLUSION

In this paper, we have presented an analytic framework for analyzing the reliability of biosensors constructed by using biomolecular circuit models whose parameters were obtained by using experimental data. The models were augmented to capture systematic and random noises introduced by variability

in experimental methods as well as due to stochastic interaction between biomolecules. Reliability analysis is performed by exploiting probabilistic dependencies between the circuit elements using a factor graph-based decoding technique. Using the simulation framework, we demonstrated the efficacy of an asymmetric biosensor code as a potential candidate for improving the reliability of biosensors. We also reported a novel “co-detection” principle based on the property of the asymmetric code. The principle exploits the nonlinear coupling between different biomolecular circuits and prescribes an experimental protocol that can be used for trace detection of pathogens in a given sample. We believe that the analytical framework proposed in this paper will serve as an important design tool for circuit designers and information theorists for evaluating the performance of different encoding and decoding principles in biosensor systems.

ACKNOWLEDGMENT

The authors would like to acknowledge Dr. E. C. Alocilja and her research group for their assistance in conducting the biosensor experiments.

REFERENCES

- [1] A. Hassibi, “Integrated microarrays,” Ph.D. dissertation, Stanford University, Stanford, CA, 2006.
- [2] Y. Liu, S. Chakrabarty, and E. C. Alocilja, “Fundamental building blocks for molecular bio-wire based forward-error correcting biosensors,” *Nanotechnology*, vol. 18, no. 42, p. 424017(6pp), Oct. 2007.
- [3] A. Hassibi, H. Vikalo, and A. Hajimiri, “On noise processes and limits of performance in biosensors,” *J. Appl. Phys.*, vol. 102, no. 1, p. 014909(12pp), 2007.
- [4] A. Hassibi, S. Zahedi, R. Navid, R. W. Dutton, and T. H. Lee, “Biological shot-noise and quantum-limited signal-to-noise ratio in affinity-based biosensors,” *J. Appl. Phys.*, vol. 97, no. 8, p. 084701(10pp), 2005.
- [5] R. Christensen, “Linear and log-linear models,” *J. Amer. Statist. Assoc.*, vol. 95, 2000.
- [6] P. J. Kuekes, W. Robinett, G. Seroussi, and R. S. Williams, “Defect-tolerant interconnect to nanoelectronic circuits: Internally redundant demultiplexers based on error-correcting codes,” *Nanotechnology*, vol. 16, no. 6, pp. 869–882, 2005.
- [7] S. Gregori, A. Cabrini, O. Khouri, and G. Torelli, “On-chip error correcting techniques for new-generation flash memories,” *Proc. IEEE*, vol. 91, no. 4, pp. 602–616, Apr. 2003.
- [8] L. Fedichkin, E. Katz, and V. Privman, “Error correction and digitalization concepts in biochemical computing,” *J. Comput. Theor. Nanosci.*, vol. 5, pp. 36–43, 2008.
- [9] E. E. May, A. M. Johnston, W. E. Hart, J. Watson, R. J. Pryor, and M. D. Rintoul, “Detection and reconstruction of error control codes for engineered and biological regulatory systems,” Sandia Nat. Labs, SAND Rep.: 2003-3963 Oct. 2003.
- [10] A. H. Khan, A. Ossadchi, R. M. Leahy, and D. J. Smith, “Error-correcting microarray design,” *Genomics*, vol. 81, no. 2, pp. 157–165, 2003.
- [11] O. Milenkovic, “Error and quality control coding for DNA microarrays,” presented at the Inaugural Workshop Center for Information Theory and Application, San Diego, CA, Feb. 2006.
- [12] S. Draghici, *Data Analysis Tools for DNA Microarrays*. Boca Raton, FL: Chapman and Hall/CRC Press, 2003.
- [13] M. L. Lee, F. C. Kuo, G. A. Whitmore, and J. Sklar, “Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations,” *Proc. National Academy Sciences United States of America*, vol. 97, no. 18, pp. 9834–9839, Aug. 2000.
- [14] R. M. Tanner, “A recursive approach to low complexity codes,” *IEEE Trans. Inf. Theory*, vol. IT-27, no. 5, pp. 533–547, Sep. 1981.
- [15] H.-A. Loeliger, “An introduction to factor graphs,” *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.
- [16] S. Pal, E. C. Alocilja, and F. P. Downes, “Nanowire labeled direct-charge transfer biosensor for detecting bacillus species,” *Biosensors Bioelectron.*, vol. 22, no. 9–10, pp. 2329–2336, Apr. 2007.

- [17] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.
- [18] G. D. Forney, Jr., "Codes on graphs: Normal realizations," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 520–548, Feb. 2001.
- [19] A. Gore, S. Chakrabartty, S. Pal, and E. C. Alcilja, "A multichannel femtoampere-sensitivity potentiostat array for biosensing applications," *IEEE Trans. Circuits Syst. I*, vol. 53, no. 11, pp. 2357–2363, Nov. 2006.
- [20] Y. Liu, A. Gore, S. Chakrabartty, and E. C. Alcilja, "Characterization of sub-systems of a molecular bio-wire based biosensor device," *Micromol. Acta*, vol. 163, no. 1–2, pp. 49–56, 2008.



Yang Liu (S'05) received the B.S. degree in automatic control from Hefei University of Technology, Hefei, China, in 2001, the M.S. degree in electrical engineering from Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei, in 2004, and is currently pursuing the Ph.D. degree in electrical and computer engineering at Michigan State University, East Lansing.

His current research interests include bioelectronics and biosensors, hybrid biological/solid-state devices, and mixed-signal very-large scale integrated design for biomedical and life science applications.

Dr. Liu is a recipient of "The Entrepreneurial Faculty for the 21st Century University" fellowship, 2008–2009, awarded by Michigan State University. He is a member of Sigma Xi.



Shantanu Chakrabartty (M'96) received the B.Tech. degree from the Indian Institute of Technology, Delhi, in 1996, and the M.S. and Ph.D. degrees in electrical engineering from Johns Hopkins University, Baltimore, MD, in 2001 and 2004, respectively.

Currently, he is an Assistant Professor in the Department of Electrical and Computer Engineering at Michigan State University, East Lansing. From 1996 to 1999, he was with Qualcomm Inc., San Diego, CA, and during 2002, he was a Visiting Researcher at the

University of Tokyo, Tokyo, Japan. His current research interests include low-power analog and digital very-large scale integrated systems, hardware implementation of machine-learning algorithms with applications to biosensors, and biomedical instrumentation.

Dr. Chakrabartty was a recipient of The Catalyst foundation fellowship from 1999 to 2004 and received the best undergraduate thesis award in 1996. He is currently a member of the IEEE BioCAS technical committee, IEEE Circuits and Systems Sensors technical committee, and is an Associate Editor of *Advances in Artificial Neural Systems* (Hindawi publications).