

RESEARCH ARTICLE

# State Estimation of the Time-Varying and Spatially Localized Concentration of Signal Molecules from the Stochastic Adsorption Dynamics on the Carbon Nanotube-Based Sensors and Its Application to Tumor Cell Detection

Hong Jang<sup>1</sup>, Jay H. Lee<sup>1\*</sup>, Richard D. Braatz<sup>2</sup>

**1** Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea, **2** Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America

\* [jayhlee@kaist.ac.kr](mailto:jayhlee@kaist.ac.kr)



CrossMark  
click for updates

OPEN ACCESS

**Citation:** Jang H, Lee JH, Braatz RD (2015) State Estimation of the Time-Varying and Spatially Localized Concentration of Signal Molecules from the Stochastic Adsorption Dynamics on the Carbon Nanotube-Based Sensors and Its Application to Tumor Cell Detection. PLoS ONE 10(11): e0141930. doi:10.1371/journal.pone.0141930

**Editor:** Adam R Hall, Wake Forest University School of Medicine, UNITED STATES

**Received:** July 26, 2015

**Accepted:** October 14, 2015

**Published:** November 3, 2015

**Copyright:** © 2015 Jang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by the Advanced Biomass R&D Center(ABC) of Global Frontier Project funded by the Ministry of Science, ICT and Future Planning (ABC-2011-0031354). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

This paper addresses a problem of estimating time-varying, local concentrations of signal molecules with a carbon-nanotube (CNT)-based sensor array system, which sends signals triggered by monomolecular adsorption/desorption events of proximate molecules on the surfaces of the sensors. Such sensors work on nano-scale phenomena and show inherently stochastic non-Gaussian behavior, which is best represented by the chemical master equation (CME) describing the time evolution of the probabilities for all the possible number of adsorbed molecules. In the CME, the adsorption rate on each sensor is linearly proportional to the local concentration in the bulk phase. State estimators are proposed for these types of sensors that fully address their stochastic nature. For CNT-based sensors motivated by tumor cell detection, the particle filter, which is nonparametric and can handle non-Gaussian distributions, is compared to a Kalman filter that approximates the underlying distributions by Gaussians. In addition, the second-order generalized pseudo Bayesian estimation (GPB2) algorithm and the Markov chain Monte Carlo (MCMC) algorithm are incorporated into KF and PF respectively, for detecting latent drift in the concentration affected by different states of a cell.

## Introduction

Recently, several near-infrared (nIR) fluorescent sensors based on single-walled carbon nanotubes (SWNTs) have been developed for detecting biomolecules in the human body [1–9]. In response to a continuous incident light source, the SWNT-based sensors detect stepwise changes in emitted light intensity triggered by monomolecular adsorption and desorption (i.e.,

adsorption and desorption at a single-molecular level) of a trace of proximate molecules on the surfaces of the sensors. The nIR fluorescence can penetrate more deeply into tissues than visible fluorescence without photobleaching or overlapping with autofluorescence from biological substrates [10, 11]. Furthermore, compared with small fluorescent probes [12–22], non-diffusive SWNTs allow for a precise spatial resolution at the micrometer scale. As a result of these advantages, SWNTs can act as effective sensing platforms for real-time, direct and selective detection *in vivo*. In particular, for nitric oxide (NO) and hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>),  $\mu\text{M}$  level concentration could be detected successfully by using this sensing platform and resolve several questions about local generation upon growth factor stimulation and the signalling mechanism in a living cell [7, 8].

This sensor technology presents some challenges as well as opportunities. A sensor array system where multiple CNT-based sensors distributed on a small area potentially can be used to follow time-varying, local concentrations of target molecules *in vivo* and in real time with precise spatial resolution. In turn, precise spatiotemporal control of these molecules may become feasible with the advent of appropriate actuators. Challenges in the experimental side include selective sensor design for target molecules in a desired detection range and actuator design for the spatiotemporal control at micro-scale. On the system's side, an immediate challenge is the development of an on-line state estimation method that can effectively extract concentration information from the stochastic adsorption data.

Some methods have been proposed for quantifying local concentrations of signal molecules near CNT-based sensors [23–25]. The estimation task is made challenging by the fact that the adsorption/desorption event is highly stochastic given a small number of molecules involved at the nanoscale sensor's surface. Conventional methods like least squares are limited in terms of performance for such problems. For a more accurate estimation, chemical master equation (CME) describing the evolution of the probability distribution among all possible adsorption states (i.e., the number of adsorbed molecules on the sensor) has been used in the estimation formulation. Based on the exact solution of the CME, maximum likelihood estimation (MLE) has been proposed [23–25]. However, the previous works assumed a constant concentration and performed the estimation with a batch set of data, which is not realistic for a sensor system working in a real-time environment in which concentrations show dynamic, time-varying behavior. What is needed is a full state estimation method that can fully and recursively utilize the information coming from the sensors to follow the local concentration in real time.

Bayesian methods have been a popular choice for state estimation of stochastic systems owing to its flexible, convenient formulation and theoretical rigor. For Gaussian systems, only the first two moments of the probability density function (PDF) have to be followed and the Kalman filter (KF) provides a simple solution to the problem. However, data from the CNT-based sensor system shows highly non-Gaussian characteristics that follow convolved binomial distributions [24]. For highly non-Gaussian systems, a class of sequential Monte Carlo methods known as particle filters (PFs) can be attractive as a nonparametric method that can handle any distribution shape [26]. The PF methods represent the required posterior PDF as a set of random samples and associated weights.

This article mainly proposes an effective recursive state estimator for estimating time-varying, local concentrations of signal molecules using the stochastic adsorption and desorption time-profiles onto the surface of the CNT-based sensors. By tracking the concentration of the signal molecules with the help of a rigorously formulated stochastic state estimator, we can gain further insights into their roles in biological systems or the effects of other species on them. The stochastic nature of the adsorption and desorption at the molecular level brings in the chemical master equation (CME) at the sensor level and makes the problem a challenging one that cannot be easily handled by the conventional state estimation techniques. Hence, the state estimation problem studied in this article has not been addressed before in the literature.

To test the feasibility and potentials of the proposed method, we test it in the context of a sensing problem, which is admittedly simplistic and artificial but still is inspired by the real biological problem. Given the known parameters in the model, performances of the KF and PF methods are examined in terms of both accuracy of estimated local concentration of the signal molecule and computational cost. The nano-sensors have previously been used for detecting and measuring signal molecules in human body, to follow the concentrations of signal molecules like nitric oxide (NO) and hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>), which are consistently generated from enzymes in vascular endothelial cells to regulate various physiological and pathological processes [6,23–25]. Their concentration levels are known to be affected significantly by cell states, the switching behavior of which is simplistically represented by a hidden Markov model in our case study. To solve the simulated estimation problem, KF and PF are designed with the second-order generalized pseudo-Bayesian estimation (GPB2) algorithm and the Markov chain Monte Carlo (MCMC) algorithm respectively, for the Markov jump system with nano-sensors. Their performances are compared for the case of a single sensor as well as of multiple sensors.

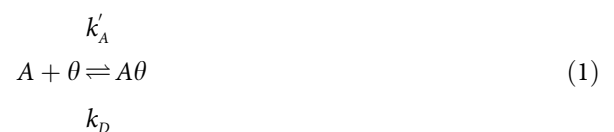
## Methods

### Single-molecule Sensor System

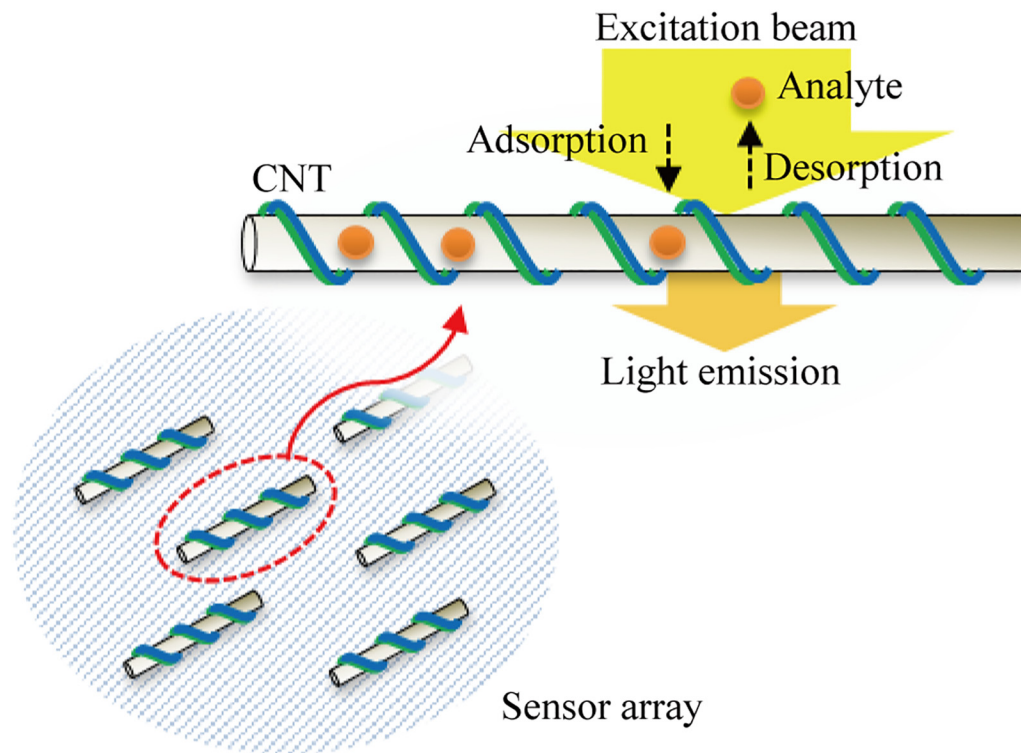
**Carbon nanotube-based sensor.** The basic mechanism of SWNT-based sensors is optical detection of discretized light intensity changes induced by adsorption and desorption of target molecules on the sensor’s surface at nano-scale. To enhance the sensitivity and selectivity for target molecules, usually present at the micromolar (μM) concentration level, the SWNT surface is functionalized by wrapping the nanotube with various polymers such as collagen [7] or certain DNA sequences [8] (Fig 1). The variation in the SWNT wrapping controls the adsorption rates of different analytes present. For example, collagen-SWNTs have shown different, selective time-profiles of adsorption and desorption events for H<sub>2</sub>O<sub>2</sub>, H<sup>+</sup>, and Fe(CN)<sub>6</sub><sup>3-</sup> in different concentration ranges [6]. Importantly, all time-profile data had reversible features, which indicate adsorption and desorption rates of similar magnitudes.

The maximum number of adsorbed molecules is experimentally found to be around 10 [8], and this number is consistent with the maximum number of excitons (an excitation is an electron and positive hole pair, which remain near each other due to electrostatic Coulomb force and is free to move through a semiconducting material) diffusion-limited segments on the SWNT [1] for which an average length is about 1~2 μm. So several SWNT-based sensors can be placed in a small area less than 10 μm<sup>2</sup> [8]. Fig 1 shows an example of a sensor array system depicted as sensors randomly distributed on a small area of neighborhood. With this array system, the objective is to estimate a time profile of the local concentration of target molecules with high accuracy.

**Stochastic adsorption model.** The number of adsorbed molecules is assumed to be read at every sampling time from the sensors, which are distributed in a sufficiently small area of a same concentration level. In developing a sensor model, free target molecules *A* in its surrounding liquid phase are assumed to adsorb onto unoccupied sites of the nanotube segment *θ* to form bound molecules *Aθ* through reversible adsorption:



where  $k'_A$  [s<sup>-1</sup>] and  $k_D$  [s<sup>-1</sup>] are adsorption and desorption rate constants, respectively. The



**Fig 1. An example of a SWNT-based sensor array system.**

doi:10.1371/journal.pone.0141930.g001

corresponding rates are expressed as

$$r_A = k'_A N_{\theta} \tag{2}$$

$$r_D = k_D N_{A\theta} \tag{3}$$

where  $N_{\theta}$  is the number of empty sites and  $N_{A\theta}$  is the number of occupied sites. The adsorption rate can be considered to be a first-order function of the local concentration of the surrounding target molecules  $C(t)$  [24],

$$k'_A = k_A C(t) \tag{4}$$

where  $k_A$  is a constant factor in the adsorption coefficient.

These equations connect the sensor information (i.e., the number of adsorbed molecules) to the concentration in the surrounding media. If the adsorption/desorption events could be deterministic, a continuum (or average) model for the sensor can be formed by one differential equation for the number of adsorbed molecules  $N_{A\theta} \in [0, N_T]$  as a continuous variable with an initial value of  $N_{A\theta,0}$ ,

$$\frac{dN_{A\theta}(t)}{dt} = k_A C(t)(N_T - N_{A\theta}(t)) - k_D N_{A\theta}(t), N_{A\theta}(0) = N_{A\theta,0} \tag{5}$$

A recursive form of the solution obtained by considering the previous measurement  $N_{A\theta}(t_{k-1})$  as an initial condition and integrating the equation for one sample interval assuming  $C(t_k)$

remains constant over the interval is

$$N_{A\theta}(t_k) = \frac{N_T + \left(N_{A\theta}(t_{k-1}) \left(1 + \frac{k_D}{k_A C(t_k)}\right) - N_T\right) e^{-(k_A C(t_k) + k_D)\Delta t}}{1 + \frac{k_D}{k_A C(t_k)}} \tag{6}$$

where  $k$  is the index for the time step and  $\Delta t$  is the size of the sample time step size, which is set sufficiently small for the approximation to be accurate.

In actuality, the adsorption reaction on the sensor surface is highly stochastic because only a very small number of molecules ( $\sim 10$ ) are involved. Hence, significant fluctuations occur from the average behavior described in (6). In this case, use of the chemical master equation (CME) composed of differential equations describing the evolution of the probabilities for all possible discrete states of the system is more appropriate [27]. Then, the state of the system is defined as the discrete number of adsorbed molecules  $\tilde{N}_{A\theta} \in [0, N_T]$ , resulting in  $N_T + 1$  total possible states. The probability of being in each state is denoted by  $P_i = \Pr(\tilde{N}_{A\theta} = i) \in [0, 1]$ , where  $i$  is the number of adsorbed molecules. The CME, along with the appropriate boundary equation, can be expressed by  $N_T + 1$  ordinary differential equations (ODEs):

$$\frac{dP_0(t)}{dt} = -P_0(t)[k'_A N_T] + P_1(t)[k_D] \tag{7}$$

$$\begin{aligned} \frac{dP_i(t)}{dt} &= P_{i-1}(t)[k'_A(N_T - (i - 1))] - P_i(t)[k_D i + k'_A(N_T - i)] + P_{i+1}(t)[k_D(i + 1)], \quad i \\ &= 1, 2, \dots, N_T - 1 \end{aligned} \tag{8}$$

$$\frac{dP_{N_T}(t)}{dt} = P_{N_T-1}(t)[k'_A] - P_{N_T}(t)[k_D N_T] \tag{9}$$

The monomolecular reaction systems, which were studied by [28], provide a path to an analytical solution of the CME. The adsorption/desorption process can be considered as a monomolecular reaction system with only two species (e.g. adsorbed molecules on sensor surface and desorbed molecules in bulk). For such a system, the probability distribution of the CME is described by a binomial distribution with time-varying parameters. More specifically, the number of adsorbed molecules  $N_{A\theta}$  at a time  $t_k$  is a random variable distributed as a binomial with the number of trials equal to  $N_T$  and probability parameter equal to  $\lambda(t_k)$ , which is related to  $N_{A\theta}(t_k)$  calculated from the continuum Eq (6) divided by  $N_T$  as

$$\tilde{N}_{A\theta}(t_k) \sim \mathcal{B}(N_T, \lambda(t_k)) \tag{10}$$

$$\Pr(\tilde{N}_{A\theta}(t_k) = i) = \binom{N_T}{i} (\lambda(t_k))^i (1 - \lambda(t_k))^{N_T-i} \tag{11}$$

$$\lambda(t_k) = \frac{N_{A\theta}(t_k)}{N_T} \tag{12}$$

The local concentration of target molecule  $C(t_k)$  enters the probability distribution of Eq (10) through  $N_{A\theta}(t_k)$  of Eq (6) appearing in Eq (12) for  $\lambda(t_k)$ .

For monomolecular adsorption, the overall population can be divided into two subsets, representing occupied sites and unoccupied sites on the sensor. With some previously measured value  $\hat{N}_{A\theta}(t_{k-1})$ , the distribution at the next time step can be derived as the convolution of two

binomial distributions applicable to the “fully occupied” and “empty” subsets, which are of size  $\hat{N}_{A\theta}(t_{k-1})$  and  $1 - \hat{N}_{A\theta}(t_{k-1})$  respectively:

$$\tilde{N}_{A\theta}(t_k) \sim B(\hat{N}_{A\theta}(t_{k-1}), \lambda^F(t_k)) * B(N_T - \hat{N}_{A\theta}(t_{k-1}), \lambda^E(t_k)) \tag{13}$$

$$\lambda^F(t_k) = \frac{1 + \frac{k_D}{k_A C(t_k)} e^{-(k_A C(t_k) + k_D)\Delta t}}{1 + \frac{k_D}{k_A C(t_k)}} \tag{14}$$

$$\lambda^E(t_k) = \frac{1 - e^{-(k_A C(t_k) + k_D)\Delta t}}{1 + \frac{k_D}{k_A C(t_k)}} \tag{15}$$

The first binomial distribution can be derived from (10)–(12) by assuming the sites are fully occupied initially, and the second binomial distribution can be derived from (10)–(12) by considering the initial state as being empty [28]. If the expression for  $N_{A\theta}(t_k)$  obtained by setting  $N_{A\theta}(t_{k-1}) = N_T$  (“fully occupied”) in (6) is further substituted into (12), the probability parameter  $\lambda(t_k)$  becomes  $\lambda^F(t_k)$  of (14) and  $N_T$  cancels out. If the same substitution is carried out by setting  $N_{A\theta}(t_{k-1}) = 0$  (“fully empty”) in (6),  $\lambda(t_k)$  becomes  $\lambda^E(t_k)$  of (15).

### Recursive State Estimation Design

Based on the observation model proposed in Section 2, the overall system for state estimation can be generally described by the discrete-time state space model,

$$x_k = f(x_{k-1}) + w_k \tag{16}$$

$$y_{k,j} \sim p(y_{k,j} | x_k, y_{k-1,j}), \quad \forall j = 1, \dots, N_s \tag{17}$$

where  $x_k$  is a single state indicating the local concentration  $C(t_k)$  in the neighborhood;  $w_k$  is zero-mean white noise;  $y_{k,j}$  is the measurement of the number of adsorbed molecules  $\tilde{N}_{A\theta,j}(t_k)$  onto the surface of the  $j$ th sensor;  $N_s$  is the number of sensors in the neighbourhood,  $f(\cdot)$  represents the state transition function which can describe production, degradation, mass transport, biological reactions, etc. of the signal molecules; and  $p(\cdot)$  denotes the probability distribution represented by the convolution of the two binomial distributions, as in (13), which describes the stochastic adsorption reaction model. The expression involves both  $x_k$  and  $y_{k-1,j}$  (corresponding to  $C(t_k)$  and  $\hat{N}_{A\theta,j}(t_{k-1})$ , respectively), which explains the use of the notation  $p(y_{k,j} | x_k, y_{k-1,j})$ . The available information at time step  $k$  is the set of measurements  $Y_k = \{y_i \in \mathbb{R}^{N_s} : i = 1, \dots, k\}$ . Note that other biological effects on the concentration are not considered in the model (16) and (17). The “cell state” as an example of such effects will be included as a hidden Markov state in the later part of this article.

The above model can be extended to a multiple-state (vector  $x$ ) system where concentrations at different spatial locations are measured by separate sets of CNT sensors, which can be useful in cases where one deals with a spatially distributed concentration profile and/or multiple signal molecules over a large sensing area. In this case, the concentrations and therefore the measured data at different locations can be correlated through the mass transfer phenomena, which can be represented by mass transport models such as diffusion equation [29]. To communicate the essence of the problem in a simple and transparent manner, this article focuses on estimation of concentration at a single location, using single or multiple sensors.

**Kalman filter.** The Bayesian approach offers a systematic way to combine prior knowledge, state and observation models, and measurement information into an informative

estimate of the state (i.e., a *posteriori* probability density function (PDF) of the state  $p(x_k|Y_k)$ ). For linear Gaussian systems, the Kalman filter (KF) enables a recursive construction of the exact PDF of the state estimate, which is parameterized by the mean and covariance. Kalman filtering can be applied to the exact probability distribution model (13) by approximation of the exact PDF by a Gaussian distribution function.

The binomial distribution  $B(n, p)$  has the mean of  $np$  and the variance of  $np(1 - p)$  and can be approximated by a normal distribution with the same mean and variance,  $\mathcal{N}(np, np(1 - p))$  [30]. In this work, the two binomial distributions in the exact observation model can be approximated by

$$\tilde{N}_{A\theta_j}(t_k) \sim \mathcal{N}(\mu_j^F, (\sigma_j^F)^2) * \mathcal{N}(\mu_j^E, (\sigma_j^E)^2) \tag{18}$$

$$\mu_j^F = \hat{N}_{A\theta_j}(t_{k-1})\lambda^F(t_k) \tag{19}$$

$$\sigma_j^F = \hat{N}_{A\theta_j}(t_{k-1})\lambda^F(t_k)(1 - \lambda^F(t_k)) \tag{20}$$

$$\mu_j^E = (N_T - \hat{N}_{A\theta_j}(t_{k-1}))\lambda^E(t_k) \tag{21}$$

$$\sigma_j^E = (N_T - \hat{N}_{A\theta_j}(t_{k-1}))\lambda^E(t_k)(1 - \lambda^E(t_k)) \tag{22}$$

Convolution of the two Gaussian distributions  $\mathcal{N}(\mu_1, \sigma_1^2)$  and  $\mathcal{N}(\mu_2, \sigma_2^2)$  is a Gaussian distribution with  $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$  [31], so the observation model can be approximated by

$$\tilde{N}_{A\theta_j}(t_k) \sim \mathcal{N}(\mu_j^F + \mu_j^E, (\sigma_j^F)^2 + (\sigma_j^E)^2) \tag{23}$$

Hence, the Gaussian-approximated observation model for the  $j$ th sensor is defined by

$$y_{k,j} = h(x_{k,j}, y_{k-1,j}) + v_{k,j}, \quad \forall j = 1, \dots, N_s \tag{24}$$

where  $h(x_{k,j}, y_{k-1,j})$  is same as the mean in (23) and  $v_{k,j}$  is zero-mean Gaussian noise with the variance of (23).

The KF method can be summarized in a recursion of prediction and correction steps, starting from an initial guess defined by the mean  $\hat{x}_{1|1}$  and the covariance  $P_{1|1}$ . Given the posterior mean and covariance of  $x_{k_0}$ , the mean and covariance of the prior PDF of next state  $x_{k+1}$  is

$$\hat{x}_{k|k-1} = \bar{A}_{k-1}\hat{x}_{k-1|k-1} \tag{25}$$

$$P_{k|k-1} = \bar{A}_{k-1}P_{k-1|k-1}\bar{A}_{k-1}^T + Q \tag{26}$$

where  $P$  is the covariance of the state and  $Q$  is the covariance of the process noise  $w_k$ , and  $\bar{A}_{k-1}$  follows from the linearization

$$\bar{A}_{k-1} = \left[ \frac{\partial f}{\partial x} \right]_{x=\hat{x}_{k-1|k-1}} \tag{27}$$

The mean and covariance of the posterior PDF is

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \mathbf{L}_k \left( \mathbf{y}_k - \begin{bmatrix} h(\hat{x}_{k|k-1,1}, \mathbf{y}_{k-1,1}) \\ \vdots \\ h(\hat{x}_{k|k-1,N_s}, \mathbf{y}_{k-1,N_s}) \end{bmatrix} \right) \quad (28)$$

$$\mathbf{L}_k = P_{k|k-1} \bar{\mathbf{C}}_k^T (\mathbf{R}_k + \bar{\mathbf{C}}_k P_{k|k-1} \bar{\mathbf{C}}_k^T)^{-1} \quad (29)$$

$$P_{k|k} = P_{k|k-1} - \mathbf{L}_k \bar{\mathbf{C}}_k P_{k|k-1} \quad (30)$$

where  $\mathbf{L}_k \in \mathbb{R}^{N_s}$  is the Kalman gain matrix and  $\bar{\mathbf{C}}_k \in \mathbb{R}^{N_s}$  follows from the linearization

$$[\bar{\mathbf{C}}_k]_j = \left[ \frac{\partial h(x_{k,j}, y_{k-1,j})}{\partial x_{k,j}} \right]_{x=\hat{x}_{k|k-1}} \quad (31)$$

The covariance matrix  $\mathbf{R}_k \in \mathbb{R}^{N_s \times N_s}$  for the measurement noise can be defined by a diagonal matrix

$$[\mathbf{R}_k]_{jj} = r_{k,j} \quad (32)$$

where  $r_{k,j}$  is same with the variance in (23) for the  $j$ th sensor.

**Particle filter.** To use the non-Gaussian observation model (17) directly, a sampling-based approach known as particle filtering (PF) can be used. PF is based on a discrete weighted approximation of the true posterior PDF with a set of random samples (particles). If the number of samples becomes extremely large, the approximation converges to the true posterior PDF.

The sequential importance sampling (SIS) algorithm is considered as the current standard of PF [26]. The first step of the algorithm is an initialization of  $N$  particles and their weights, denoted by  $\{x_k^i, w_k^i; i = 1, \dots, M\}$ . In this step, each particle is sampled from the initial PDF  $p(x_1) \sim \mathcal{N}(\bar{x}_1, Q_1)$  and the associated weight is initialized to  $1/M$ . After the initialization, the importance sampling step and the weight update step are repeated. In the importance sampling step,  $x_k^i, i = 1, \dots, M$  are sampled from an importance density  $q(x_k|Y_k)$  which is a user-defined choice. The importance density is commonly chosen as the prior PDF,

$$q(x_k|Y_k) = p(x_k|x_{k-1}) \quad (33)$$

In the weight update step, the weight  $w_k^i$  for each  $i = 1, \dots, M$  is updated with

$$w_k^i \propto w_{k-1}^i \frac{\prod_{j=1}^{N_s} p(y_{k,j}|x_{k,j}^i, y_{k-1,j}) p(x_k|x_{k-1})}{q(x_k|Y_k)} \quad (34)$$

If (33) is substituted into (34), the weight update equation is described as

$$w_k^i \propto w_{k-1}^i \prod_{j=1}^{N_s} p(y_{k,j}|x_{k,j}^i, y_{k-1,j}) \quad (35)$$

Based on the samples and normalized weights, the posterior PDF can be approximated as

$$p(x_k|Y_{k-1}) = \sum_{i=1}^M w_k^i \delta(x_k - x_k^i) \quad (36)$$

where  $\delta(\cdot)$  is the Dirac delta function. The estimated value is commonly calculated as a



weighted mean,

$$\hat{x}_{k|k} = \sum_{i=1}^M w_k^i x_k^i \quad (37)$$

In addition, a resampling step can be added to mitigate the degeneracy problem [26]. The degeneracy problem refers to the growing number of samples having negligible weights with iterations. The resampling step eliminates the samples with small weights and concentrates the calculation on those samples with large weights whenever a significant degeneracy problem is detected. After generating a new set of  $x_k^i$  for  $i = 1, \dots, M$  by resampling, the weights are reset to  $1/M$  as in the initialization step. After resampling, the estimated value  $\hat{x}_{k|k}$  is calculated as a mean of  $x_k^i$  for  $i = 1, \dots, M$ .

## Results and Discussion

This case study for testing the two approaches is motivated by the problem of detecting tumor cells through NO and H<sub>2</sub>O<sub>2</sub> signal molecules. NO generated from vascular endothelial NO synthase (eNOS) correlates with stimulation of angiogenesis. This activity is intimately linked with metastasis of tumor cells since their survival and proliferation are highly dependent on adequate supply of O<sub>2</sub> and nutrients from blood vessels by diffusion [32–34]. Membrane-associated NADPH oxidases are also found in vascular endothelial as well as smooth muscle cells, and generate H<sub>2</sub>O<sub>2</sub> as an important signal molecule in angiogenesis. Produced H<sub>2</sub>O<sub>2</sub> can activate signalling pathways to stimulate tumor cell proliferation and migration [35–37]. Knowledge of how concentrations of these signal molecules change as a cell changes its state can help understand their biological roles in tumor cell growth, which in turn can lead to advances in medical treatments.

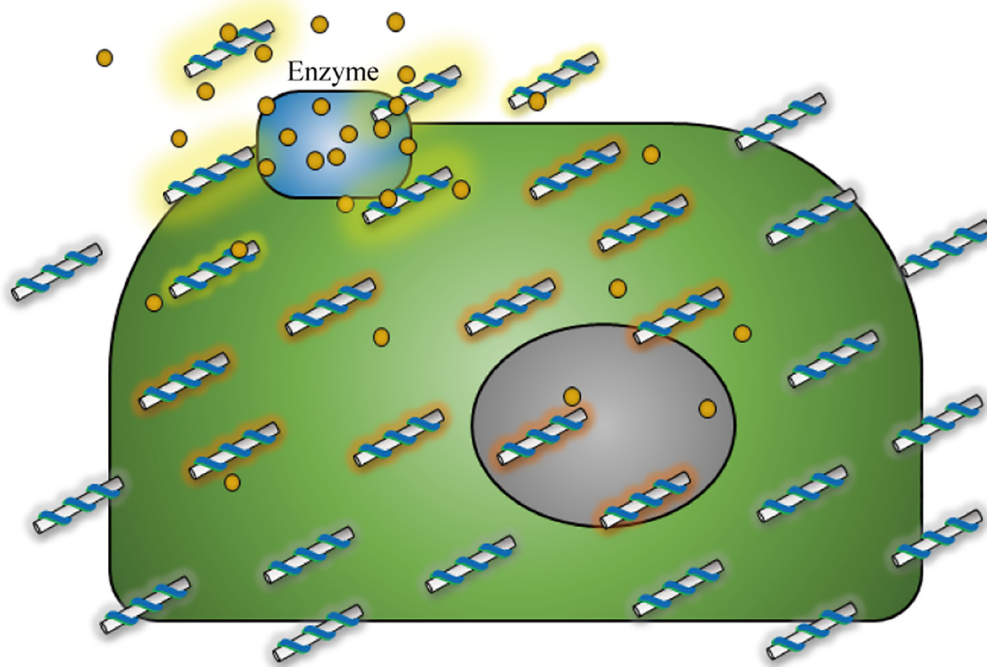
The estimation of the concentration of signal molecules from a normal cell is examined first, and then the more complex case of a cell transitioning from a normal state to a tumor state is considered.

### Estimating the concentration of signal molecules from a normal vascular endothelial cell

This section develops a state estimation problem for the signal molecules (NO or H<sub>2</sub>O<sub>2</sub>) from vascular endothelial cells. The width and length of the endothelial cells is more than 10 μm [38], which indicates that dozens of SWNT-based sensors can be placed on a single vascular endothelial cell and send multiple stochastic monomolecular adsorption data [7] (Fig 2). Among them, sensors near the enzymes generating signal molecules, where frequent adsorption/desorption events are detected, can be selected and used in the estimation of the local concentration of the signal molecules. The small area proximate to the generator of the signal molecules can be considered as a neighborhood sharing same local concentration that represents the cell state as a whole.

It is difficult to obtain from an experimental setup a large dataset that includes sufficient, representative stochastic variations to render a fair and thorough evaluation of estimation performance. Alternatively, representative stochastic adsorption datasets can be generated from kinetic Monte Carlo (KMC) simulations. Each KMC simulation run can be viewed as a realization of the stochastic system that is described by the CME [39]. The adsorption/desorption process involves fairly simple molecular level events and Zhang et al. 2010 [8] showed that experimental data for this system was well described by the KMC simulation.

In this particular simulation study, the number of adsorbed molecules on the sensor is allowed to range from 0 to 10, so the number of possible discrete states is 11. The length of



**Fig 2. Generation of signal molecules from a membrane-associated enzyme and their detection from a sensor array system on a single cell.**

doi:10.1371/journal.pone.0141930.g002

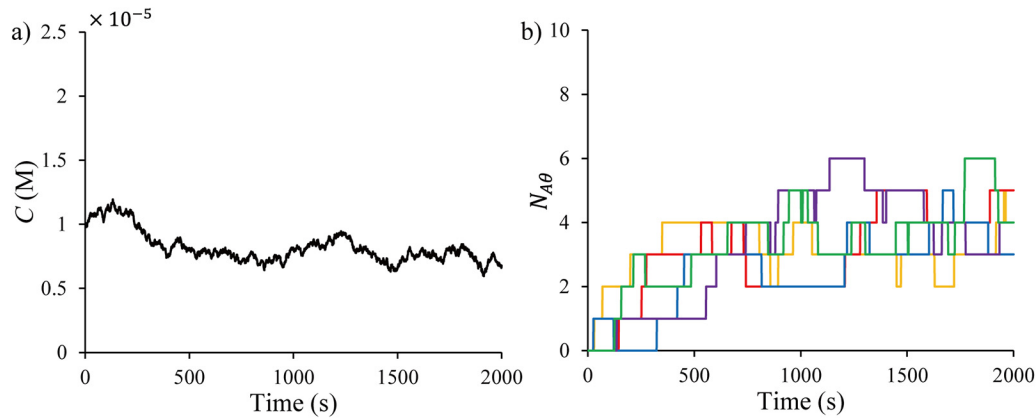
each run is 2000 s and the sampling time interval is 1 s. The starting state is assumed to be 0 (empty of molecules). Adsorption/desorption parameters,  $k_A$  and  $k_D$ , are chosen as  $100 \text{ M}^{-1}\text{s}^{-1}$  and  $0.001 \text{ s}^{-1}$  respectively, which are taken from [6]. In the normal vascular endothelial cell, the signal molecules are released consistently from the enzyme at a low concentration level ( $\sim 10 \text{ }\mu\text{M}$ ) [40]. These dynamics can be simply described as an integrated white noise process,

$$x_k = x_{k-1} + w_k \quad (38)$$

Fig 3A shows an example time-varying concentration with  $w_k \sim \mathcal{N}(0, 10^{-16})$  in (38) and Fig 3B shows associated five representative realizations of the time profile of the number of adsorbed molecules.

The adsorption data show five distinct time profiles for the same local concentration profile, indicating significant stochastic characteristics of the sensor system at the nano-scale. Other features of the stochastic data shown in Fig 3B are the stepwise (discretized) and reversible variations. In addition, each time profile shows several stationary regions after unpredicted transitions even as the local concentration changes continuously at the bulk phase, which indicates the information-poor characteristic of the dataset, for which accurate modelling is essential for accurate state estimation.

We consider the state estimation problem with increasing number of sensors in the neighborhood measuring the same local concentration. KF and PF (with 200 particles) are compared with increasing number of sensors. KF is based on the Gaussian-approximated observation model, while PF considers the full non-Gaussian stochastic model resulting in the non-Gaussian posterior PDFs as shown in Fig 4. The posterior PDFs are kernel densities reconstructed from the particles of PF for a specific dataset at  $t = 100 \text{ s}$ ,  $500 \text{ s}$ ,  $1000 \text{ s}$ , and  $1500 \text{ s}$ . The overall PDFs show non-Gaussian distributions that are mostly positively skewed.



**Fig 3. Time plots for (a) the local concentration and (b) associated stochastic adsorption data for a single sensor, generated by KMC simulations.**

doi:10.1371/journal.pone.0141930.g003

Performance of the two estimation methods can be compared by observing how well the estimates track the true concentration profile throughout the run time from a wrong initial guess ( $\bar{x}_1 = 2 \times 10^{-5}$ ,  $Q_1 = 1 \times 10^{-5}$ ). Plots of the estimates for the 1-sensor and 5-sensor cases are shown in Fig 5. For the 1-sensor case, the PF estimates follow the true concentration more closely than the KF (Fig 5A). For the 5-sensor case, the gap between the PF and KF estimates is reduced as long time (Fig 5B).

For quantitative comparison, the root-mean-square-errors (RMSEs) of the estimated concentrations per run are averaged over 100 runs that generated different adsorption/desorption data from different local concentration profiles (2000-sample dataset per one run). The averaged RMSE is defined by

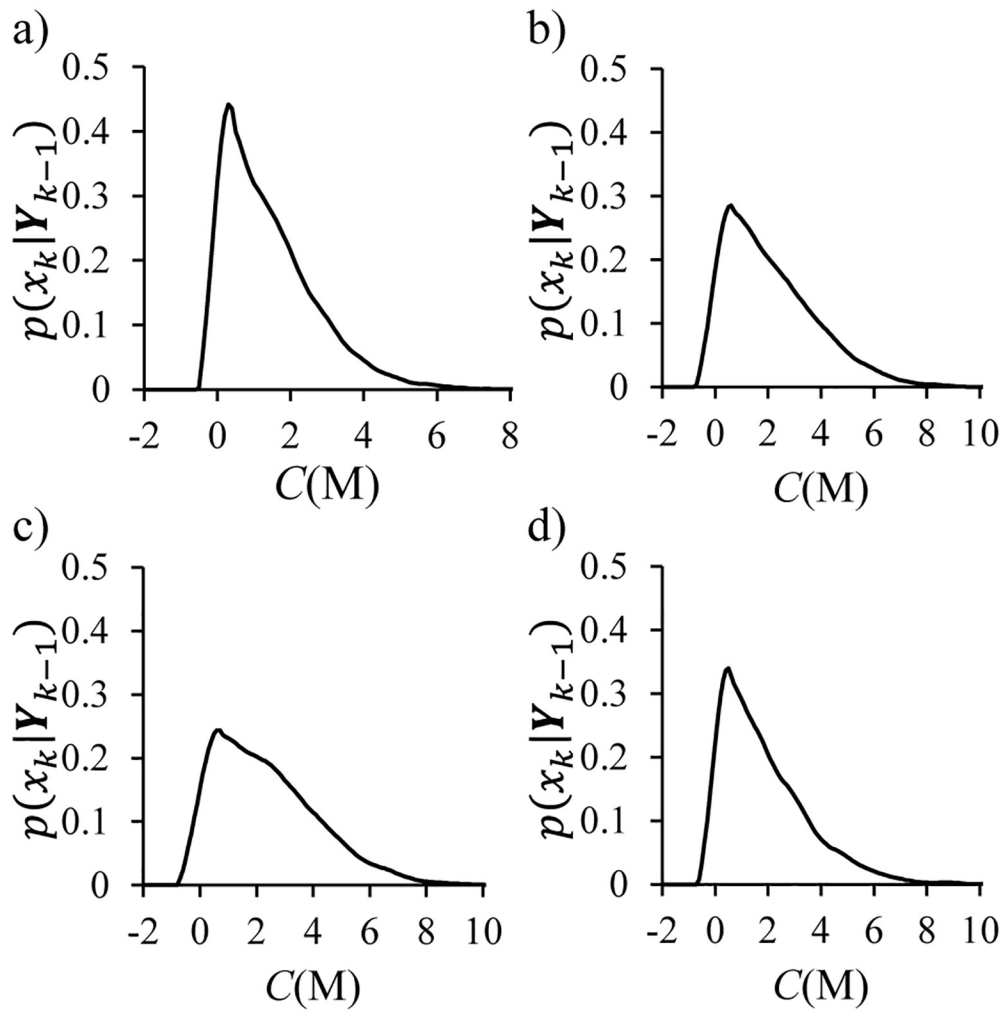
$$RMSE = \frac{\sum_{i=1}^{N_R} \sqrt{\frac{\sum_{k=1}^{N_T} (C_{true,i,k} \times 10^5 - \hat{C}_{i,k} \times 10^5)^2}{N_S}}}{N_R} \tag{39}$$

where  $N_S$  is the number of samples in one run,  $N_R$  is the number of runs,  $C_{true,i,k}$  is the true local concentration value, and  $\hat{C}_{i,k}$  is the estimate for the  $k$ th sample time of the  $i$ th run.

In Table 1, for all cases, the RMSEs of the estimates from PF are smaller than from KF. The difference in the RMSEs of the two methods slowly decreases with increasing number of sensors in the neighborhood, while the computation time of PF is higher and increases more rapidly than that of KF. If the objective is only a nominal state estimate, the benefit of rigorous stochastic modelling in the state estimation is reduced when more information is contained in the dataset (through the use of multiple sensors). Of course, a disadvantage of the KF for any number of sensors is that it is not capable of estimating the non-Gaussian character of the distribution of the state estimates.

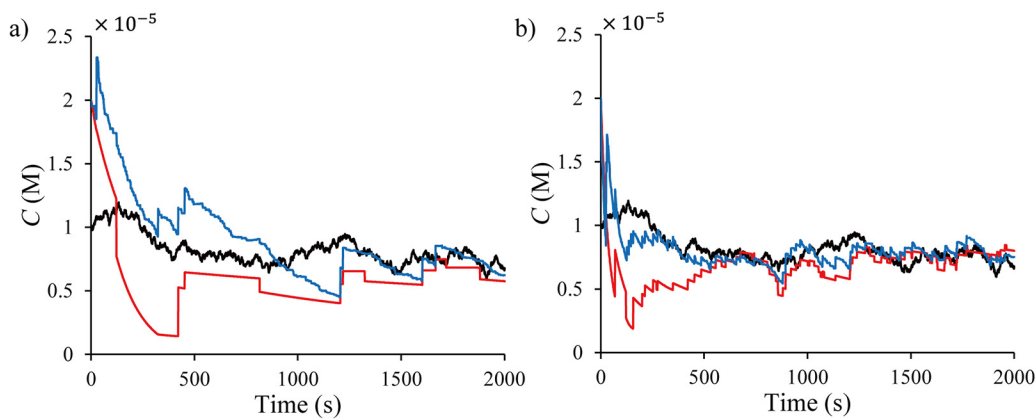
### Estimating the signal molecules from a cell having two states

In normal vascular endothelial cells, signal molecules generated from the enzymes are at a low concentration level ( $\sim 10 \mu M$ ). In tumor vascular endothelial cells, on the other hand, the expression levels and activities of eNOS are abnormally increased compared to the normal endothelial cells (Fig 6), and the elevated level of NO promotes tumor progression and metastasis by inducing angiogenesis as well as tumor cell invasion, proliferation, and migration [41,



**Fig 4. Kernel density reconstruction of the posterior PDFs from the PF at (a) 100 s, (b) 500 s, (c) 1000 s, and (d) 1500 s.**

doi:10.1371/journal.pone.0141930.g004



**Fig 5. Time plots of the concentration estimates from the KF (red - - -) and PF (blue - - -) for (a) 1 sensor and (b) 5 sensors.**

doi:10.1371/journal.pone.0141930.g005

**Table 1. Averaged RMSEs of estimates from the PF and KF and associated computation time (in seconds) with increasing number of sensors.**

	# of sensors	1	5	10	20
Average RMSE <sup>1</sup>	PF	0.5038	0.3129	0.2457	0.2074
	KF	0.5356	0.3333	0.2521	0.1972
Average computational time <sup>1,2</sup>	PF	0.0035	0.0135	0.0275	0.0519
	KF	<0.0001	0.0001	0.0001	0.0001

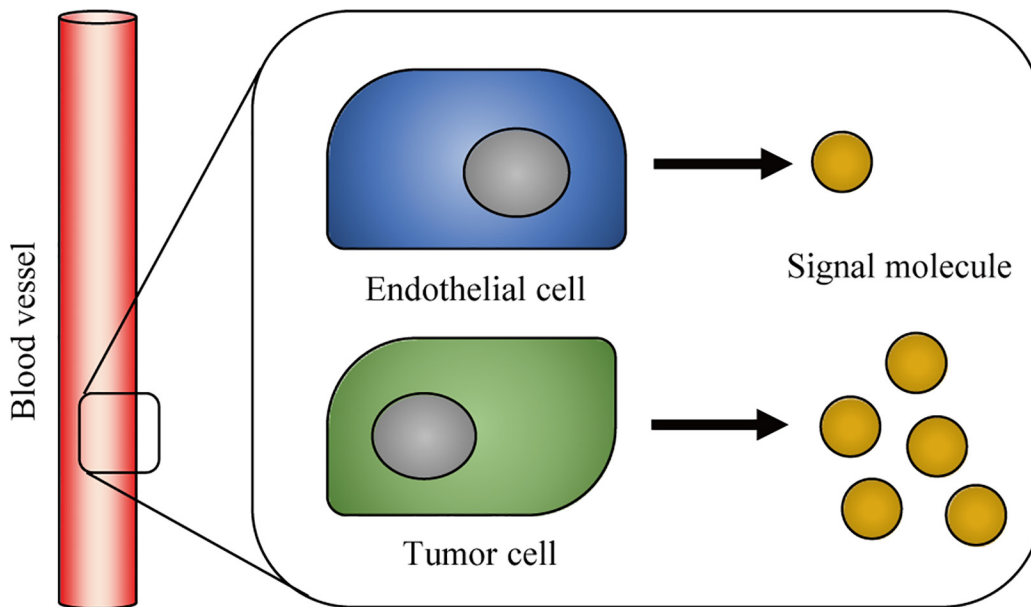
<sup>1</sup>The values are averaged over 100 runs.

<sup>2</sup>The computation time was recorded in seconds using a workstation with 3.40 GHz CPU and 8GB RAM.

doi:10.1371/journal.pone.0141930.t001

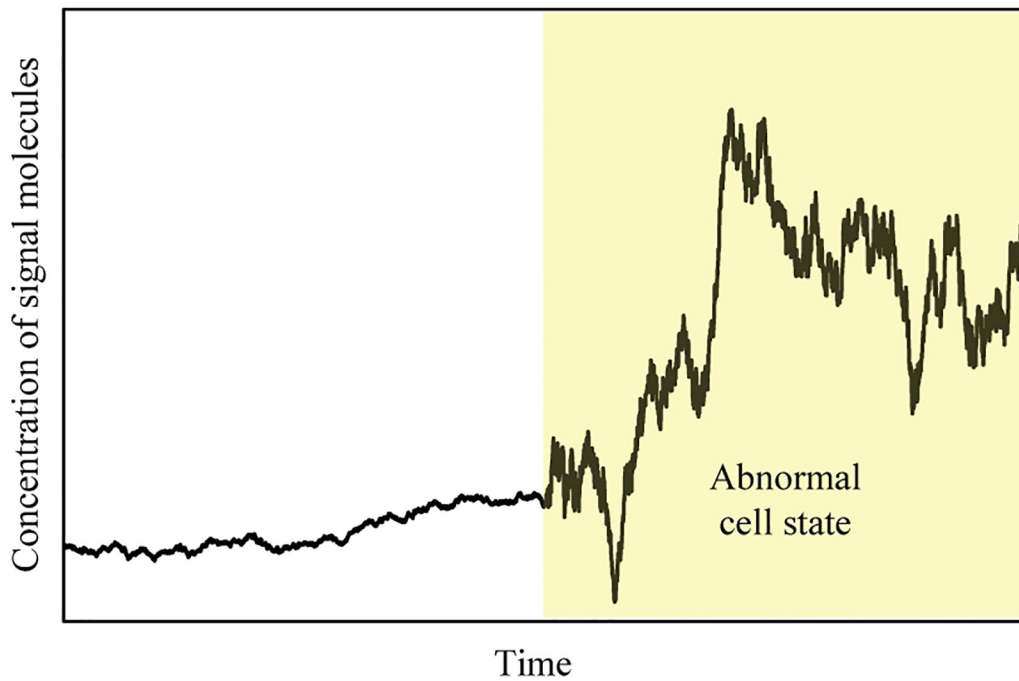
[42]. For H<sub>2</sub>O<sub>2</sub>, there is also a considerable variation among cells in the concentration level required to initiate a particular biological process. Moreover, it has been observed that different levels of H<sub>2</sub>O<sub>2</sub> can induce distinct responses within a cell. For example, overproduction of H<sub>2</sub>O<sub>2</sub> results in proliferation and migration of smooth muscle cells, contributing to atherosclerosis and restenosis [43].

**Hidden Markov model.** The concentration of the signal molecules affects and is affected by the state of the cell. For real-time state estimation, consideration of all the complex biological processes associated with different cell states is very challenging and linking with concentration variations of the signal molecules can easily become intractable. In addition, the signal molecules are small gaseous molecules showing very fast diffusion (with diffusion coefficients of around 10<sup>-5</sup> [cm<sup>2</sup>/s]) compared to cell activities in tissues [40]. In this context, we simplify the system to having two states: a normal state and an abnormal state. In the normal state, the signal molecules are released consistently at a low concentration level. In the abnormal state, the concentration of the signal molecules increases (drifts) rapidly to a new elevated



**Fig 6. Production of signal molecules from vascular endothelial cells.**

doi:10.1371/journal.pone.0141930.g006



**Fig 7. Variation of the concentration of signal molecules with different cell states.**

doi:10.1371/journal.pone.0141930.g007

level. Fig 7 illustrates a possible concentration variation in the vascular endothelial cell as it transitions to the abnormal state (based on data generated from the artificial Hidden Markov model).

Such a pattern in Fig 7 can be characterized as a mixture of quiescent and drifting phases, which is called “intermittent drift.” A hidden Markov model (HMM) can be used for modelling such shifts in the disturbance pattern [44, 45]. HMM represents a useful class of statistical models where a hidden state,  $\mathcal{H} \in \{1, 2, \dots, H\}$  transitions probabilistically among possible states in a Markovian fashion. In this work, each member of the set  $\mathcal{H}$  represents a particular system state, for example, “normal cell” or “tumor cell” states. Mathematically, a finite-state Markov chain is a sequence of random integers,  $r_k$ , where the transition probability matrix  $\Pi$  has elements defined by

$$\pi_{ij} = \Pr(r_k = j | r_{k-1} = i), \quad i, j \in \mathcal{H} \tag{40}$$

$$\sum_{j=1}^{\mathcal{H}} \pi_{ij} = 1 \tag{41}$$

Based on the transition probability matrix, the intermittent drift in the concentration of the signal molecule,  $x_k$  can be described by

$$x_k = x_{k-1} + w_{r_k}, \quad r_k \in 1, 2 \tag{42}$$

$$\pi_{11} = \pi_{22} \approx 1 \tag{43}$$

$$\pi_{12} = \pi_{21} < 1 \tag{44}$$

where “1” indicates the normal cell and “2” indicates the tumor cell. The  $w_{r_k}$  is a white Gaussian noise with covariance  $Q_{r_k}$  defined by

$$Q_{r_k=1} \approx 0 \tag{45}$$

$$Q_{r_k=2} \gg Q_{r_k=1} \tag{46}$$

Since there is only a low probability of switching once the system enters a particular regime, a diagonally dominant  $\Pi$  is employed, as reflected in (43) and (44). Note that the actual regime is usually not known with complete certainty and must be inferred from measurements. Additional behavior could be incorporated into the model by introducing more hidden states (e.g. other transitional cell states or environmental effects on the local concentration) with appropriate accompanying stochastic models for them.

**Second-order generalized pseudo-Bayesian algorithm.** For using KF for a Markov jump system represented by (42), the generalized pseudo-Bayesian estimation algorithm of order 2 (GPB2) has been suggested as an effective sub-optimal filter [46]. Let  $\hat{x}_{k|k}(r_{k-1}, r_k)$  denote the estimate conditioned on the two most recent hidden state realizations. Similarly, the corresponding estimation error covariance is represented as  $P_{k|k}(r_{k-1}, r_k)$ . The main idea is to generate multiple Gaussian distributions from KF for all possible trajectories of the last two hidden states, and combine them into a single Gaussian distribution, parameterized by  $\{\hat{x}_{k|k}, P_{k|k}\}$ . A recursive scheme is characterized by two steps: “branching” and “merging.”

Starting with  $\{\hat{x}_{k-1|k-1}(r_{k-1}), P_{k-1|k-1}(r_{k-1})\}$ , the branching step is to obtain the set  $\{\hat{x}_{k|k}(r_{k-1}, r_k), P_{k|k}(r_{k-1}, r_k)\}$  through the prediction and correction steps of KF. The merging step involves the law of total probability and Bayes’ rule to collapse the products from the branching step as

$$\hat{x}_{k|k}(r_k) = \sum_{r_{k-1}=1}^2 \hat{x}_{k|k}(r_{k-1}, r_k) p(r_{k-1}|r_k, \mathbf{Y}_k) \tag{47}$$

$$P_{k|k}(r_k) = \sum_{r_{k-1}=1}^2 [\{\hat{x}_{k|k}(r_{k-1}, r_k) - \hat{x}_{k|k}(r_k)\}^2 + P_{k|k}(r_{k-1}, r_k)] p(r_{k-1}|r_k, \mathbf{Y}_k) \tag{48}$$

$$p(r_{k-1}|r_k, \mathbf{Y}_k) = \frac{1}{c_1} p(y_k|r_{k-1}, r_k, \mathbf{Y}_{k-1}) p(r_k|r_{k-1}) p(r_{k-1}|\mathbf{Y}_{k-1}) \tag{49}$$

where  $c_1$  is a constant ensuring that  $p(r_{k-1}|r_k, \mathbf{Y}_k)$  sums to unity, and  $p(y_k|r_{k-1}, r_k, \mathbf{Y}_{k-1})$  is related to the correction step of KF in the branching step. A point estimate is obtained from

$$\hat{x}_{k|k} = \sum_{r_{k-1}=1}^2 \hat{x}_{k|k}(r_k) p(r_k|\mathbf{Y}_k) \tag{50}$$

$$P_{k|k} = \sum_{r_{k-1}=1}^2 [\{\hat{x}_{k|k}(r_k) - \hat{x}_{k|k}\}^2 + P_{k|k}(r_k)] p(r_k|\mathbf{Y}_k) \tag{51}$$

$$p(r_k|\mathbf{Y}_k) = \frac{1}{c_2} \sum_{r_{k-1}=1}^2 p(y_k|r_{k-1}, r_k, \mathbf{Y}_{k-1}) p(r_k|r_{k-1}) p(r_{k-1}|\mathbf{Y}_{k-1}) \tag{52}$$

where  $c_2$  is a constant ensuring that  $p(r_k|\mathbf{Y}_k)$  sums to unity.

**Markov chain Monte Carlo algorithm.** Adapting PF to the Markov jump system is relatively simpler than KF. Starting with  $\{x_{k-1}^i(r_{k-2}, r_{k-1}), w_{k-1}^i(r_{k-2}, r_{k-1})\}$ , samples  $x_k^i(r_{k-1}, r_k)$  for  $i = 1, \dots, M$  are generated from the same importance density of (33) for all possible trajectory for the recent hidden Markov states,  $r_{k-1} = 1, 2$  and  $r_k = 1, 2$ . This approach is called the

Markov chain Monte Carlo (MCMC) algorithm [47]. The weight update Eq (35) is modified by including,  $p(r_k, r_{k-1} | Y_k)$ ,

$$w_k^i(r_{k-1}, r_k) \propto w_{k-1}^i(r_{k-1}, r_k) p(r_k, r_{k-1} | Y_k) \prod_{j=1}^{N_s} p(y_{k,j} | x_{k,j}^i(r_{k-1}, r_k), y_{k-1,j}) \quad (53)$$

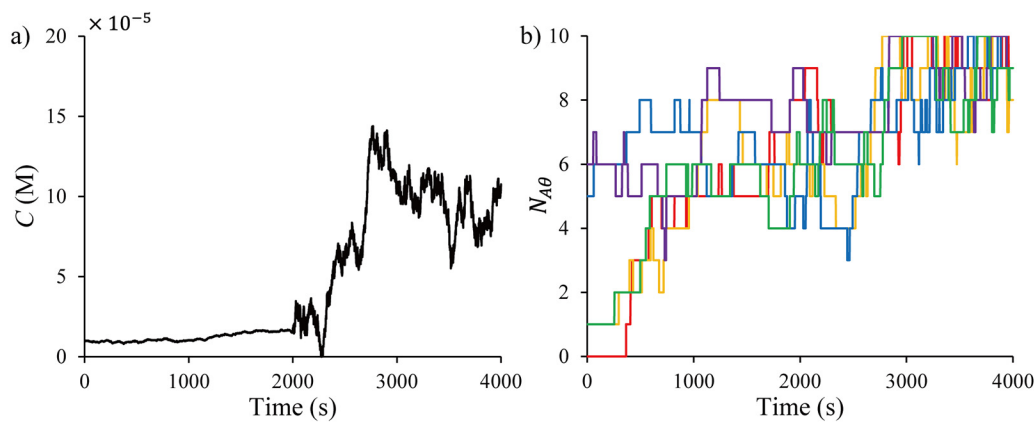
Finally, the point estimate can be obtained by

$$\hat{x}_{k|k} = \sum_{r_{k-1}=1}^2 \sum_{r_k=1}^2 \sum_{i=1}^M w_k^i(r_{k-1}, r_k) x_k^i(r_{k-1}, r_k) \quad (54)$$

**Detection of tumor cell activity.** As stated before, two regimes are considered in the system: the normal cell state and the tumor cell state. The objective is to detect a regime change through the local concentration variations of the signal molecules seen from the nano sensors. The reference work [7] investigates the effect of a growth factor, which stimulates cell growth, proliferation, and differentiation on the  $H_2O_2$  generation in living cells. From the 3000 s observation after the stimulation with the growth factor at  $t = 0$ , it was observed that the  $H_2O_2$  concentration level increased immediately and reached a maximum in the time range between 600 s and 1800 s. This observation indicates that the tumor cell activity and its effect on the local concentration of signal molecules can be prolonged for a long time (~ 30 min).

Based on this data, stochastic adsorption/desorption profiles were generated from KMC simulation. The number of adsorption sites on the sensor is 10 and the length of each run is 4000 s with the sampling time interval of 1 s. The starting state is assumed to be a random integer less than 10 (partly occupied) and the  $k_A$  and  $k_D$  are assumed to be  $100 M^{-1}s^{-1}$  and  $0.001 s^{-1}$ . Eq (42) is used in the state estimation as the state model. At the ‘normal cell’ state, the local concentration is stable and affected only by low-level noise ( $w_k \sim \mathcal{N}(0, 10^{-16})$ ). When the cell becomes a tumor cell, the local concentration becomes elevated by high-level noise ( $w_k \sim \mathcal{N}(0, 5 \times 10^{-14})$ ). The plots in Fig 8 show a representative concentration variation with the tumor cell activity for the time period from 2000 s to 4000 s and five different realizations of the associated time profile of the number of adsorbed molecules.

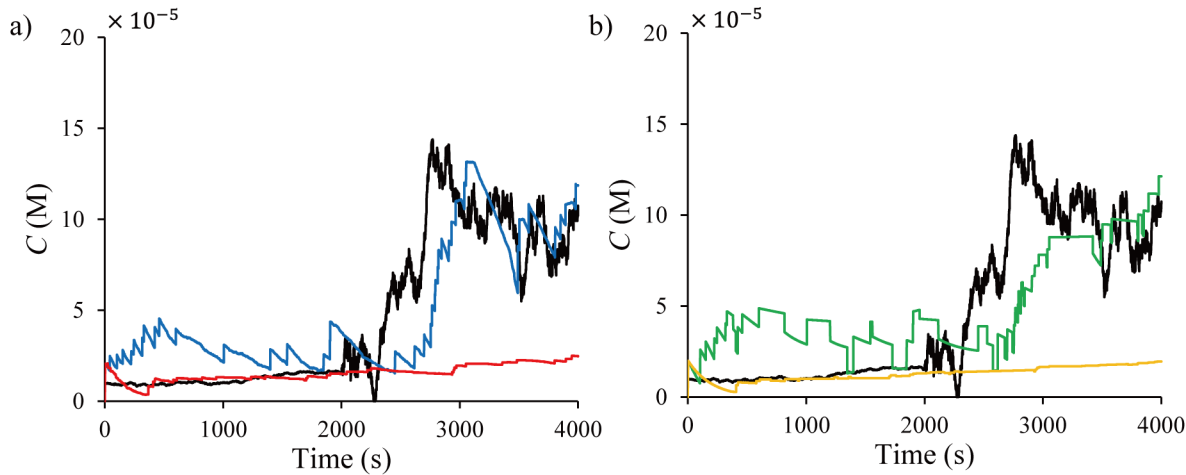
For this switching system, KF with the GPB2 algorithm (shortly, KF-GPB2) and PF with the MCMC algorithm (shortly, PF-MCMC) are designed and compared. The hidden Markov



**Fig 8. Time plots of (a) the local concentration affected by the tumor cell activity and (b) associated stochastic adsorption data, generated by running KMC simulations.**

doi:10.1371/journal.pone.0141930.g008





**Fig 9. Time plots of the concentration estimates from (a) PF (red - - -) and PF-MCMC (blue - - -) and (b) KF (orange - - -) and KF-GPB2 (green - - -) for one sensor.**

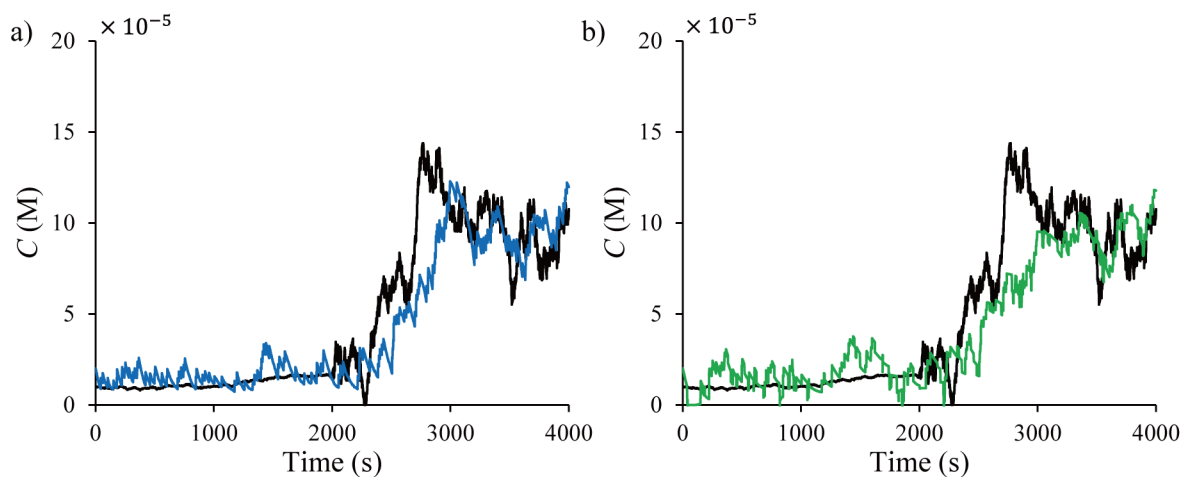
doi:10.1371/journal.pone.0141930.g009

chain used the state transition probability matrix

$$\Pi = \begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix} \tag{55}$$

Fig 9 shows the state estimates of PF, PF-MCMC, KF, and KF-GPB2 for one sensor. The basic PF and KF used in Section 4 cannot effectively follow the concentration drift caused by the tumor cell activity, while the PF-MCMC and KF-GPB2 estimators follow the drift much better. Among the methods, the estimates from PF-MCMC are closer to the true concentration dynamics and show less fluctuations compared to KF-GPB2.

In both PF-MCMC and KF-GPB2, the state estimates are improved further when more stochastic adsorption/desorption data are available, obtained from multiple sensors, as shown in Fig 10.



**Fig 10. Time plots of the concentration estimates from (a) PF-MCMC (blue - - -) and (b) KF-GPB2 (green - - -) for five sensors.**

doi:10.1371/journal.pone.0141930.g010

**Table 2. Averaged RMSEs and of the estimates from PF-MCMC and KF-GPB2 and associated computation time for increasing number of sensors.**

	# of sensors	1	5	10	20
Average RMSE <sup>1</sup>	PF-MCMC	2.3748	1.9180	1.7544	1.6190
	KF-GPB2	2.6095	2.0816	1.8984	1.7542
Average computational time <sup>1,2</sup>	PF-MCMC	0.0120	0.0474	0.0839	0.1637
	KF-GPB2	0.0016	0.0076	0.0126	0.0250

<sup>1</sup>The values are averaged over 100 runs.

<sup>2</sup>The computation time was recorded in seconds using a workstation with 3.40 GHz CPU, 8GB RAM.

doi:10.1371/journal.pone.0141930.t002

Table 2 shows the averaged RMSEs and computational time for PF-MCMC and KF-GPB2 based on 100 runs that generated different adsorption/desorption data from different tumor cell activities. In both methods, the RMSEs of the estimates decrease with increasing number of sensors in the neighborhood with PF-MCMC having smaller RMSE values than KF-GPB2. Though the computation time of PF-MCMC is larger than KF-GPB2, it is far less than the sampling time of 1 s.

In a real application, the neighborhood region proximate to the enzyme should be very small considering the short-life time and high diffusivity of the signal molecules. Therefore, less than 5 sensors might be valid in the state estimation for a single cell [7]. In this context, PF-MCMC can be recommended if accurate estimates of the local concentration of the signal molecules are needed with such limited information.

## Conclusions

Two stochastic state estimation methods—Kalman filtering (KF) and particle filtering (PF)—were investigated for estimating the time-varying local concentration of signal molecules from stochastic monomolecular adsorption/desorption data on the surface of the carbon-nanotube (CNT)-based sensors. In addition, the second-order generalized pseudo Bayesian estimation (GPB2) algorithm and the Markov chain Monte Carlo (MCMC) algorithm were incorporated into KF and PF respectively, for detecting latent drift in the concentration affected by different states of a cell. The stochastic nature of the adsorption data from each CNT-based sensor was fully modelled by using the chemical master equation (CME). In addition, intermittent concentration variations of the signal molecules were modelled by a hidden Markov model. Performances of the state estimators with the sensor array system were compared through a case study employing KMC simulation. The PF-MCMC combination showed the highest accuracy while having reasonable computation time.

Use of drugs affecting the production of signal molecules by inhibiting the associated enzyme or directly scavenging the signal molecules appears to be a promising strategy to inhibit angiogenesis and therefore tumor growth [48, 49]. In order to control the modification of signal molecules in a precise manner, further understanding of various factors involved such as the timing, concentration, and location is required. The proposed state estimators have promise in this endeavor.

## Supporting Information

**S1 Dataset. Original data including true concentration profiles, stochastic adsorption profiles and concentration estimates for plotting Figs 3, 5, 8, 9 and 10.**

(XLSX)

## Author Contributions

Analyzed the data: HJ JHL. Wrote the paper: HJ JHL RDB. Formulated the state estimation methods: HJ. Performed the simulation and estimation: HJ. Performed the numerical analysis: HJ.

## References

1. Cognet L, Tsybouski DA, Rocha J-DR, Doyle CD, Tour JM, Weisman RB. Stepwise quenching of exciton fluorescence in carbon nanotubes by single-molecule reactions. *Science*. 2007; 316(5830):1465–8. PMID: [17556581](#)
2. Saito S, Zettl A. *Carbon Nanotubes: Quantum Cylinders of Graphene*: Elsevier; 2008.
3. Barone PW, Baik S, Heller DA, Strano MS. Near-infrared optical sensors based on single-walled carbon nanotubes. *Nat Mater*. 2005; 4(1):86–92. PMID: [15592477](#)
4. Heller DA, Jeng ES, Yeung T-K, Martinez BM, Moll AE, Gastala JB, et al. Optical detection of DNA conformational polymorphism on single-walled carbon nanotubes. *Science*. 2006; 311(5760):508–11. PMID: [16439657](#)
5. Kim JH, Ahn JH, Barone PW, Jin H, Zhang J, Heller DA, et al. A luciferase/single-walled carbon nanotube conjugate for near-infrared fluorescent detection of cellular ATP. *Angew Chem Int Edit*. 2010; 122(8):1498–501.
6. Jin H, Heller DA, Kim J-H, Strano MS. Stochastic analysis of stepwise fluorescence quenching reactions on single-walled carbon nanotubes: single molecule sensors. *Nano Letters*. 2008; 8(12):4299–304. doi: [10.1021/nl802010z](#) PMID: [19367966](#)
7. Jin H, Heller DA, Kalbacova M, Kim J-H, Zhang J, Boghossian AA, et al. Detection of single-molecule H<sub>2</sub>O<sub>2</sub> signalling from epidermal growth factor receptor using fluorescent single-walled carbon nanotubes. *Nat Nanotechnol*. 2010; 5(4):302–9. doi: [10.1038/nnano.2010.24](#) PMID: [20208549](#)
8. Zhang J, Boghossian AA, Barone PW, Rwei A, Kim J-H, Lin D, et al. Single molecule detection of nitric oxide enabled by d(AT)<sub>15</sub> DNA adsorbed to near infrared fluorescent single-walled carbon nanotubes. *J Am Chem Soc*. 2010; 133(3):567–81. doi: [10.1021/ja1084942](#) PMID: [21142158](#)
9. Kim J-H, Heller DA, Jin H, Barone PW, Song C, Zhang J, et al. The rational design of nitric oxide selectivity in single-walled carbon nanotube near-infrared fluorescence sensors for biological detection. *Nat Chem*. 2009; 1(6):473–81. doi: [10.1038/nchem.332](#) PMID: [21378915](#)
10. O'Connell MJ, Bachilo SM, Huffman CB, Moore VC, Strano MS, Haroz EH, et al. Band gap fluorescence from individual single-walled carbon nanotubes. *Science*. 2002; 297(5581):593–6. PMID: [12142535](#)
11. Heller DA, Baik S, Eurell TE, Strano MS. Single-walled carbon nanotube spectroscopy in live cells: towards long-term labels and optical sensors. *Adv Mater*. 2005; 17(23):2793–9.
12. Kelm M, Dahmann R, Wink D, Feelisch M. The nitric oxide/superoxide assay insights into the biological chemistry of the NO/O<sub>2</sub> interaction. *J Biol Chem*. 1997; 272(15):9922–32. PMID: [9092531](#)
13. Staniek K, Nohl H. H<sub>2</sub>O<sub>2</sub> detection from intact mitochondria as a measure for one-electron reduction of dioxygen requires a non-invasive assay system. *BBA-Bioenergetics*. 1999; 1413(2):70–80. PMID: [10514548](#)
14. Grau M, Hendgen-Cotta UB, Brouzos P, Drexhage C, Rassaf T, Lauer T, et al. Recent methodological advances in the analysis of nitrite in the human circulation: nitrite as a biochemical parameter of the L-arginine/NO pathway. *J Chromatogr B*. 2007; 851(1):106–23.
15. Kojima H, Nakatsubo N, Kikuchi K, Kawahara S, Kirino Y, Nagoshi H, et al. Detection and imaging of nitric oxide with novel fluorescent indicators: diaminofluoresceins. *Anal Chem*. 1998; 70(13):2446–53. PMID: [9666719](#)
16. Cathcart R, Schwiens E, Ames BN. Detection of picomole levels of hydroperoxides using a fluorescent dichlorofluorescein assay. *Anal Biochem*. 1983; 134(1):111–6. PMID: [6660480](#)
17. Belousov VV, Fradkov AF, Lukyanov KA, Staroverov DB, Shakhbazov KS, Terskikh AV, et al. Genetically encoded fluorescent indicator for intracellular hydrogen peroxide. *Nat Methods*. 2006; 3(4):281–6. PMID: [16554833](#)
18. Casanova D, Bouzigues C, Nguyen T-L, Ramodiharilafy RO, Bouzahir-Sima L, Gacoin T, et al. Single europium-doped nanoparticles measure temporal pattern of reactive oxygen species production inside cells. *Nat Nanotechnol*. 2009; 4(9):581–5. doi: [10.1038/nnano.2009.200](#) PMID: [19734931](#)
19. Lim MH, Xu D, Lippard SJ. Visualization of nitric oxide in living cells by a copper-based fluorescent probe. *Nat Chem Biol*. 2006; 2(7):375–80. PMID: [16732295](#)

20. Lim MH, Lippard SJ. Metal-based turn-on fluorescent probes for sensing nitric oxide. *Accounts Chem Res.* 2007; 40(1):41–51.
21. Robinson JK, Bollinger MJ, Birks JW. Luminol/H<sub>2</sub>O<sub>2</sub> chemiluminescence detector for the analysis of nitric oxide in exhaled breath. *Anal Chem.* 1999; 71(22):5131–6. PMID: [10575964](#)
22. Sasaki E, Kojima H, Nishimatsu H, Urano Y, Kikuchi K, Hirata Y, et al. Highly sensitive near-infrared fluorescent probes for nitric oxide and their application to isolated organs. *J Am Chem Soc.* 2005; 127(11):3684–5. PMID: [15771488](#)
23. Boghossian AA, Zhang J, Le Floch-Yin FT, Ulissi ZW, Bojo P, Han J-H, et al. The chemical dynamics of nanosensors capable of single-molecule detection. *J Chem Phys.* 2011; 135(8):084124. doi: [10.1063/1.3606496](#) PMID: [21895176](#)
24. Ulissi ZW, Zhang J, Boghossian AA, Reuel NF, Shimizu SF, Braatz RD, et al. Applicability of birth–death Markov modeling for single-molecule counting using single-walled carbon nanotube fluorescent sensor arrays. *J Phys Chem Lett.* 2011; 2(14):1690–4.
25. Ulissi ZW, Strano MS, Braatz RD. Control of nano and microchemical systems. *Comput Chem Eng.* 2013; 51:149–56.
26. Arulampalam MS, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE T Signal Proces.* 2002; 50(2):174–88.
27. Fichthorn KA, Weinberg WH. Theoretical foundations of dynamical Monte Carlo simulations. *J Chem Phys.* 1991; 95(2):1090–6.
28. Jahnke T, Huisinga W. Solving the chemical master equation for monomolecular reaction systems analytically. *J Math Biol.* 2007; 54(1):1–26. PMID: [16953443](#)
29. Jang H, Lee JH, Braatz RD. Estimation of local concentration from measurements of stochastic adsorption dynamics using carbon nanotube-based sensors. *Korean J Chem Eng.* 2015:1–13.
30. Hayter A. *Probability and Statistics for Engineers and Scientists*: Cengage Learning; 2012.
31. Bracewell RN, Bracewell R. *The Fourier Transform and Its Applications*: McGraw-Hill New York; 1986.
32. Fukumura D, Kashiwagi S, Jain RK. The role of nitric oxide in tumour progression. *Nat Rev Cancer.* 2006; 6(7):521–34. PMID: [16794635](#)
33. Jenkins D, Charles I, Thomsen LLe, Moss D, Holmes L, Baylis S, et al. Roles of nitric oxide in tumor growth. *P Natl Acad Sci.* 1995; 92(10):4392–6.
34. Aicher A, Heeschen C, Mildner-Rihm C, Urbich C, Ihling C, Technau-Ihling K, et al. Essential role of endothelial nitric oxide synthase for mobilization of stem and progenitor cells. *Nat Med.* 2003; 9(11):1370–6. PMID: [14556003](#)
35. Veal EA, Day AM, Morgan BA. Hydrogen peroxide sensing and signaling. *Mol Cell.* 2007; 26(1):1–14. PMID: [17434122](#)
36. Li P-F, Dietz R, von Harsdorf R. Differential effect of hydrogen peroxide and superoxide anion on apoptosis and proliferation of vascular smooth muscle cells. *Circulation.* 1997; 96(10):3602–9. PMID: [9396461](#)
37. Irani K. Oxidant signaling in vascular cell growth, death, and survival a review of the roles of reactive oxygen species in smooth muscle and endothelial cell mitogenic and apoptotic signaling. *Circ Res.* 2000; 87(3):179–83. PMID: [10926866](#)
38. Garipcan B, Maenz S, Pham T, Settmacher U, Jandt KD, Zanow J, et al. Image analysis of endothelial microstructure and endothelial cell dimensions of human arteries—a preliminary study. *Adv Eng Mater.* 2011; 13(1-2):B54–B7.
39. Gillespie DT. Exact stochastic simulation of coupled chemical reactions. *J Phys Chem.* 1977; 81(25):2340–61.
40. Truskey G, Yuan F, Katz D. *Transport Phenomena in Biological Systems*. Pearson Prentice Hall. 2009.
41. Jones MK, Tsugawa K, Tarnawski AS, Baatar D. Dual actions of nitric oxide on angiogenesis: possible roles of PKC, ERK, and AP-1. *Biochem Bioph Res Co.* 2004; 318(2):520–8.
42. Shizukuda Y, Tang S, Yokota R, Ware JA. Vascular endothelial growth factor–induced endothelial cell migration and proliferation depend on a nitric oxide–mediated decrease in protein kinase C $\delta$  activity. *Circ Res.* 1999; 85(3):247–56. PMID: [10436167](#)
43. Moses MA, Klagsbrun M, Shing Y. The role of growth factors in vascular cell development and differentiation. *Int Rev Cytol.* 1995; 161:1–48. PMID: [7558689](#)
44. Wong WC, Lee JH. Fault detection and diagnosis using hidden markov disturbance models. *Ind Eng Chem Res.* 2010; 49(17):7901–8.

45. Wong WC, Lee JH. Realistic disturbance modeling using hidden Markov models: applications in model-based process control. *J Process Contr.* 2009; 19(9):1438–50.
46. Bar-Shalom Y, Li X-R. Estimation and tracking-principles, techniques, and software. Norwood, MA: Artech House, Inc, 1993.
47. Doucet A, Gordon NJ, Krishnamurthy V. Particle filters for state estimation of jump Markov linear systems. *IEEE T Signal Proces.* 2001; 49(3):613–24.
48. Morbidelli L, Donnini S, Ziche M. Role of nitric oxide in tumor angiogenesis. In: Kirsch M, Black PM, editors. *Angiogenesis in Brain Tumors*: Springer; 2004. p. 155–67.
49. Czubryt MP, Pierce GN. The role of hydrogen peroxide as a signaling molecule. In: Dhalla NS, Hryshko LV, Kardami E, Singal PK, editors. *Signal Transduction and Cardiac Hypertrophy*: Springer; 2003. p. 249–63.