# Mobile Microscopic Sensors for High-Resolution in vivo Diagnostics*

Tad Hogg and Philip J. Kuekes
Hewlett-Packard Laboratories

October 18, 2006

**Abstract**

Molecular electronics and nanoscale chemical sensors could enable constructing microscopic sensors capable of detecting patterns of chemicals in a fluid. Information from a large number of such devices flowing passively in the bloodstream allows estimating properties of tiny chemical sources in a macroscopic tissue volume. We use estimates of plausible device capabilities to evaluate their performance for typical chemicals released into the blood by tissues in response to localized injury or infection. We find the devices can readily discriminate a single cell-sized chemical source from the background chemical concentration, providing high-resolution sensing in both time and space. By contrast, such a chemical source would be difficult to distinguish from background when diluted throughout the blood volume as obtained with a blood sample.

keywords: molecular electronics; nanoscale chemical sensors; microfluidics; medical diagnostics; nanorobotics

*to appear in *Nanomedicine: Nanotechnology, Biology, and Medicine*

# 1 Background

Microscopic machines, with sizes comparable to bacteria, and nanoscale-structured materials inside the body could significantly improve disease diagnosis and treatment [1, 2, 3]. Many of these proposed applications require machines with significant capabilities for molecular recognition and manipulation, computation, communication and in vivo locomotion. They also involve coordinated behaviors among large numbers of devices, which must therefore be individually cheap to manufacture for practical use. Unfortunately, fabricating such machines is well beyond current manufacturing capabilities.

Instead, identifying applications requiring only minimal device capabilities indicates early uses for the technology. We thus consider microscopic devices, about one micron in size, built from components whose fabrication is already demonstrated in laboratory settings, specifically molecular electronics and nanowire chemical sensing [4, 5, 6, 7, 8, 9, 10]. These components will enable microscopic devices to detect chemicals at low concentrations in fluids and provide modest computation. But the devices will lack the locomotion and communication capabilities required for more advanced applications.

An important question is whether such limited devices would nevertheless provide significant benefits for some biomedical applications. In particular, without locomotion the devices cannot independently move into regions of interest, e.g., by following chemical gradients. And without transmission, devices could not communicate to compare sensor readings with those of neighboring devices nor report in vivo observations in real-time to an experimenter or physician. Instead, such devices would be limited to passively monitoring chemicals presented to them in their microenvironments, performing modest pattern-recognition computations on their own observations and storing the results in their memories for later retrieval.

A compelling application for large numbers of such devices is high-resolution detection of patterns of chemicals released into the bloodstream from cell-sized sources in biological tissues. The small size of the devices gives them access to individual cells, their large numbers allows simultaneous monitoring of many cells, and the speed of molecular electronics and sensors allows millisecond time resolution. However, without independent locomotion devices must either be directly placed in regions of interest or moved to them via other means. An example of the former approach is embedding the devices in biofilms to monitor chemical signals exchanged among the bacteria. In this paper we consider a more ambitious application using the latter approach to placing the devices: high-resolution in vivo chemical sensing in large multicellular organisms via their circulatory system. Our focus on collecting information from within small blood vessels throughout a tissue volume complements other applications of nanoscale devices for monitoring chemical behaviors within individual cells [11].

Microscopic sensors could detect localized high concentrations of chemicals that are too low to distinguish from background concentrations when diluted in the whole blood volume as obtained with a sample. Moreover, if the interesting event consists of the joint expression of several chemicals, each of which also occurs from separate sources, the device's pattern recognition capability could identify the spatial locality, which would not be apparent when the chemicals are mixed throughout the blood volume. The pattern could consist of chemicals expressed over some spatial extent along a small vessel, which the

devices could determine by comparing their detections with previous events stored in their memories. The sensor information would also indicate changes (e.g., response to an external stimulus such as introduction of a drug) that would be impractical to obtain from, say, repeated blood samples.

We describe a specific protocol for using the limited microscopic devices in the next section. The remainder of the paper evaluates its performance in a challenging diagnostic scenario to illustrate the capability of these devices.

# 2 Protocol for High-Resolution Sensing

Micron-scale devices are small enough to move through even tiny blood vessels. Hence, even without independent locomotion, injecting many devices into the bloodstream would allow them to pass as close to the cells of a tissue as circulating blood cells, typically within a few cell diameters of every cell in the tissue. While moving passively with the fluid flow of the bloodstream, the devices can monitor for preprogrammed patterns of chemicals released into the fluid. A population of such devices could simultaneously monitor for chemicals released into the blood by any of a large number of individual cells in a tissue volume. This application exploits the combination of features these devices provide: small size, large numbers, chemical sensing even at low concentrations and modest computation.

In our protocol, the devices store their observations and the time they occur based on an internal clock in their memories. After circulating in the bloodstream for a predetermined time, the devices are retrieved via a filtration process and their memories read. This read out process can use optical, acoustic, electrical or other methods in controlled laboratory settings, allowing easier access to the devices than during their time in vivo.

Finally, a conventional computer uses the data retrieved from all the devices to estimate properties of the chemical sources in the tissue. These properties include the spatial structure and concentrations of the chemical sources encountered by the devices. Estimating these values is a computational inference problem whose results will depend on the sophistication of the inference algorithm and the prior knowledge of likely values.

Inferring properties of the tissue from a collection of observations from microscopic devices contrasts with the reconstruction problem of computerized tomography [12]. In tomography, the data consists of integrals of the quantity of interest (e.g., absorption of x-rays) over a large set of lines with known geometry selected by the experimenter. The microscopic sensors, on the other hand, can record individual sensor events as they pass through vessels distributed throughout the tissue, providing more information than just a single aggregate value such as the total number of sensor events. However, the precise path of each sensor through the tissue, i.e., which vessel branches it took and the locations of those vessels, will not be known.

In our protocol, the microscopic sensors do not require in vivo locomotion or communication, which considerably simplifies their fabrication. On the other hand, the lack of these capabilities reduces performance compared with more advanced technologies. An important question is how well this protocol can perform. For instance, without locomotion, the devices cannot follow concentration gradients and will spend only a brief time near a small source emitting chemicals into the bloodstream, which limits the number of molecules

they can detect to distinguish the source from background concentration. Similarly, without communication, a device will have no explicit ability to correlate its observations with those of other nearby devices, thereby increasing the difficulty of the subsequent computational inference on the collected data. To quantitatively address these questions, the next sections of this paper examine the ability to infer existence of chemical sources from the information collected from many microscopic sensors moving passively through the vessels within a macroscopic tissue volume.

# 3 Methods for Evaluating Device Performance

Because the microscopic sensors can not yet be fabricated, estimates of their performance use models of both the devices and their task environment. Moreover, quantitative biophysical properties of many microenvironments are not precisely known so models necessarily use rough estimates of a plausible range of values. Such studies can not yet be validated with physical experiments. Nevertheless observed behaviors of microorganisms, which face the same physical constraints as future microscopic sensors, give guidelines for feasible behaviors.

One modeling approach focuses on collective behaviors in a highly simplified environment, such as cellular automata. For example, a two-dimensional scenario demonstrates structure assembly [13] from local rules, but does not include physical behaviors such as fluid flow. Distributed controls for swarms [14] are well-suited to microscopic devices with their limited physical and computational capabilities and large numbers. These studies show how local interactions lead to interesting collective behaviors, though generally assume the devices can communicate either directly or through changes they make to their environments [14, 15].

Simulations including some physical properties of microscopic environments give plausible quantitative performance estimates with various assumptions about the task environment. As one example, a two-dimensional simulation of chemotaxis [16] indicates the ability to find microscopic chemical sources. A more elaborate simulator [17] includes three-dimensional motions in viscous fluids, Brownian motion, and numerous cell-sized objects in the fluid, though without accounting for how these objects change the fluid flow.

Another approach to device behaviors employs a stochastic mathematical framework for distributed computational systems [18, 19]. This method directly evaluates the average behaviors of many devices, which would otherwise need numerous repeated runs of a simulation. This approach is best suited for simple control strategies, with minimal dependencies on events in individual device histories. Microscopic devices, with limited computational and communication capabilities, are likely to be used with relatively simple reactive controls for which this analytic approach is ideally suited. One example is following chemical gradients in a one-dimensional geometry without fluid flow [20].

Cellular automata, swarms, physically-based simulations and stochastic analysis are useful tools for evaluating the behaviors of microscopic devices. In particular, minimally-capable devices which simply record events they encounter while passively moving with the fluid do not require elaborate control programs. Thus we can estimate the device behavior with these various approaches within a given model of the task environment. In our

case, the variation in task environment properties, such as fluid flow speeds and chemical concentrations, affects performance more significantly than details of the device models. To examine likely performance of our protocol for limited-capability microscopic devices we employ plausible estimates of device behaviors and task environment properties in a model of behavior simple enough to solve directly without extensive simulations. This approach is adequate to estimate the ability to detect small chemical sources and the corresponding false positive detections, while requiring only a simple inference algorithm on the data collected from the devices.

## 3.1 Device Properties

For microscopic devices in small vessels, viscosity dominates the fluid motion, with different physical behaviors than seen with larger organisms and robots [21, 22, 23, 24]. In the approximate treatment considered here, we suppose the devices in the small vessels move with the average fluid velocity, denoted as $v_{\rm avg}$.

Nanoscale chemical sensors involve the selective binding of molecules to receptors which alters the electrical characteristics of nanoscale wires. Such sensors can detect femtomolar concentrations [8, 9], with performance primarily limited by the time required for molecules to reach the sensor. Thus for sensing chemicals at low concentrations, statistical fluctuations in the number of molecules encountered is a dominant source of sensor noise. This noise must be considered in the choice of inference algorithm for the collected data.

Microscopic sensors and bacteria face similar physical constraints in detecting chemicals [25]. The diffusive capture rate $\gamma$ for a sphere of radius $a$ in a region with concentration $C$ is [26]

$$\gamma = 4\pi DaC \tag{1}$$

Even when sensors cover only a relatively small fraction of the device surface, the capture rate is almost this large due to the relatively long time a diffusing molecule remains near the surface once it gets close to it [26]. Devices with nonspherical shapes have similar capture rates. Thus Eq. (1) is a reasonable approximation for a variety of designs of micron-scale devices.

With the relevant fluid speeds, chemical concentrations, and source sizes for the task environment described below, the sensors will pass through high concentrations near the sources on millisecond time scales. Checking for patterns of detected chemicals requires on-board computation. In our protocol, most of the computation to interpret sensor results occurs in conventional computers after the devices are retrieved. The computation required in the devices themselves is modest. Recognizing and storing a chemical detection involves at least a few arithmetic operations to compare sensor counts to threshold values stored in memory. An estimate on the required computational capability is about 100 elementary logic operations and memory accesses within a $10$ms measurement time, i.e., about $10^4$ logic operations per second. This rate is well within the capabilities of molecular electronics [10], and much slower than the speed of conventional computers.

Operating the devices requires a power source. Specifically, each logic operation in current electronic circuits use $10^4 - 10^5$ times the thermal noise level $k_B T = 4 \times 10^{-21}$J at the fluid temperature of Table 1, where $k_B$ is the Boltzmann constant. Near term molecular

electronics could reduce this to $\approx 10^3 k_B T$, in which case $10^4$ operations per second uses a bit less than $0.1$pW.

For tasks of limited duration, the devices could use an on-board fuel source. Otherwise, the devices could use energy available in their environment such as converting vibrations to electrical energy [27] or using chemical reactions. For instance, typical concentrations of glucose and oxygen in the bloodstream could generate $\approx 1000$pW continuously, limited primarily by the diffusion rate of these molecules to the device [1].

## 3.2 Task Environment Properties

Tissue microenvironments vary considerably in many properties relevant to the performance of our sensing protocol. These include the nature of the significant chemical sources, the density and geometry of vessels passing through the tissue and fluid flow rates. As a specific example illustrating the capabilities of passive motion by microscopic sensors, we consider a task environment consisting of a macroscopic volume of tissue containing a few microscopic sources producing a particular chemical (or combination of chemicals) while the rest of the tissue does not produce this chemical, or only produces it at much lower background concentrations. For simplicity, we treat the chemical sources as spheres of the same size and uniformly distributed throughout the tissue volume, with parameters given in Table 1.

| parameter | value |
|---|---|
| **vessels and tissue** | |
| vessel radius | $R = 5\mu$m |
| vessel length | $L = 1000\mu$m |
| number density of vessels in tissue | $\rho_{\text{vessel}} = 500/\text{mm}^3$ |
| tissue volume | $V = 1\text{cm}^3$ |
| **fluid** | |
| average fluid velocity | $v_{\text{avg}} = 1000\mu\text{m/s}$ |
| fluid temperature | $T = 310$K |
| **sensors** | |
| sensor radius | $a = 1\mu$m |
| number density of sensors in vessels | $\rho_{\text{sensor}} = 200/\text{mm}^3$ |

Table 1: Parameters for the sensing task.

The density of sensor devices in Table 1 corresponds to $n_{\text{sensor}} = 10^9$ devices in the entire $5$-liter blood volume of a typical adult. This many devices occupy only about $10^{-6}$ of the vessel volume and have a total mass of about $4$mg. By comparison, blood cells occupy $20\% - 40\%$ of the blood volume. Thus in operation, the typical spacing between devices will be many times their own size so their primary hydrodynamic interactions will be with the blood cells and vessel walls rather than with other devices.

We consider a scenario in which sensors can be coarsely localized to a macroscopic volume $V$ of interest, of about $1$cm in size. This coarse localization could be due to a distinctive chemical environment (e.g., high oxygen concentrations in the lungs), an externally

supplied signal (e.g., ultrasound) detectable by devices passing through vessels within the volume, or a combination of both methods. In the first case, devices would have sensors not only for the specific chemicals of interest, but also for whatever marker chemicals distinguish the desired operation tissue from others in the body. If normally occurring chemicals are not sufficiently discriminating, specific marker chemicals could be injected into the tissue of interest. This method could mark fairly small volumes, depending on the size of the injection needle and rate at which the marker chemical degrades as it moves away from the injection site. The second approach to coarse localization, via external signals, requires devices able to detect the signals, thereby somewhat increasing their fabrication complexity. Nevertheless, inbound communication *to* the devices of just a few bits (in this case, whether or not to be active) is much simpler than having the devices able to transmit information or receive instructions at high bit rates.

With the coarse localization, the devices could be designed to be active only when they detect they are in the specified region. Outside the tissue volume, the devices passively circulate without recording chemical events. Such behavior illustrates an advantage of devices with computational capability: simple changes in the control program installed in the devices could alter the conditions under which they become active rather than requiring different hardware. While not necessary for operation, coarse localization reduces false positive detections when the diagnostic of interest is known to be in such a region. Alternatively, the devices could always be active and record the coarse localization information in their memories as they encounter it. In either case, the coarse spatial localization could aid subsequent interpretation of chemical source detections.

We focus on chemical detection in the small blood vessels, i.e., those small enough to allow chemicals from the sources in the tissue to diffuse into, and with flow slow enough to give reasonable detection probability before a sensor moves far past the source. A rough model of the geometry is each of the small vessels has length $L$ and is randomly placed in the tissue. Let $\rho_{\text{vessel}}$ be the number density of such vessels in the tissue volume. The $\rho_{\text{vessel}}V$ small vessels in the tissue volume occupy about a fraction $\rho_{\text{vessel}}\pi R^2 L \approx 4\%$ of the volume with the parameters of Table 1.

To evaluate high-resolution sensing, we suppose the chemical sources are small, with size $R_{\text{source}}$ as small as individual cells, i.e., about $10\mu$m. The rate sensors enter small vessels within the tissue volume (via the branching from larger vessels) is

$$\omega_{\text{sensor}} = \rho_{\text{sensor}}\rho_{\text{vessel}}V\pi R^2 v_{\text{avg}} \approx 8 \times 10^3/\text{s} \tag{2}$$

where $\rho_{\text{sensor}}$ is the number density of sensor devices within the vessel volume, and the numerical value uses the parameters of Table 1. This is also the rate sensors leave the small vessels as they merge to form larger vessels.

## 3.3 Chemical Properties

Many chemicals of interest will have sufficiently low concentrations that the devices are likely to only encounter a few molecules while passing near the source. For example, one scenario for microscopic sensors is detecting small regions of infection or injury before they become serious. In this case, the chemicals at the source arise from the initial immunological response, whereby the injured area produces chemicals which enter nearby

| parameter | value |
|---|---|
| diffusion coefficient | $D = 100\mu\text{m}^2/\text{s}$ |
| concentration near source | $C_{\text{source}} = 2\,\text{molecule}/\mu\text{m}^3$ |
| background concentration | $C_{\text{background}} = 6 \times 10^{-3}\,\text{molecule}/\mu\text{m}^3$ |
| chemical production flux | $F = 20\,\text{molecule}/\text{s}/\mu\text{m}^2$ |
| chemical source radius | $R_{\text{source}} = 10\mu\text{m}$ |

Table 2: Parameters for chemicals and a source. The diffusion coefficient and chemical concentrations correspond to a typical 10 kilodalton chemokine molecule, with mass concentrations near the source and background (i.e., far from the source) equal to $3 \times 10^{-8}$g/ml and $10^{-10}$g/ml, respectively. The chemical production flux at the surface of the source is chosen to give the measured steady-state concentration at the surface of the source, taken to be a sphere with radius $R_{\text{source}}$.

small blood vessels to recruit white blood cells [28]. Table 2 gives typical properties of such chemicals, with concentrations well above the demonstrated sensitivity of nanoscale chemical sensors [8, 9].

With these parameters, Eq. (1) gives the capture rates $\gamma \approx 8/\text{s}$ at the background concentration and $\approx 2300/\text{s}$ near the source. Detection over a time interval $\Delta t$ is a Poisson process with mean number of detections $\gamma \Delta t$.

We consider static chemical concentration from the source(s). The concentration satisfies the time-independent diffusion equation. Taking the source to be a sphere, the concentration depends only on the distance $r$ from the center of the sphere with $r \geq R_{\text{source}}$:

$$DrC''(r) + 2DC'(r) = r\alpha C(r) \tag{3}$$

where $D$ is the diffusion coefficient for the chemical, $C(r)$ is the concentration at distance $r$ from the source and $\alpha$ is the rate at which the chemical is removed from the tissue volume. Removal mechanisms include degradation of the chemical in the tissue as well as physical removal of the chemical entering the vessels via fluid flow. For instance, the characteristic time to flow through the small vessel to reach a larger merging branch and then move out of the tissue volume is $\approx L/v_{\text{avg}}$, i.e., about a second. With the parameters of Table 1, such vessels occupy a few percent of the tissue volume, giving an effective removal rate of $\alpha \approx 10^{-2}/\text{s}$. Aside from this removal term, the diffusion equation ignores changes in concentration within the vessels due to fluid flow. That is, we model the concentration in the tissue and assume the small fraction of tissue occupied by the vessels does not substantially alter the concentration profile near the source.

With a spherical source producing the chemical with a flux $F$, the solution to Eq. (3) is

$$C(r) = \frac{FR_{\text{source}}^2}{Dr(1 + R_{\text{source}}/\eta)} \exp\left(-\frac{r - R_{\text{source}}}{\eta}\right) \tag{4}$$

for $r \geq R_{\text{source}}$ with $\eta = \sqrt{D/\alpha}$ characterizing the range from the source beyond which the concentration decreases exponentially with distance. With our parameters, $\eta$ is $100\mu$m. With multiple sources, the total concentration is the sum of that from each of the sources.

With the choices of Table 2, the flux $F$ corresponds to a total production $4\pi R_{\text{source}}^2 F = 2 \times 10^4$molecule/s and $C(R_{\text{source}})$ matches the value $C_{\text{source}}$ in the table. As a point of comparison for this flux applied to a single cell, a eucaryotic cell has $\sim 10^7$ ribosomes and production of a protein molecule takes about a minute, allowing production of about $10^5$ protein molecules per second. While not all these molecules will be signal proteins released through the cell membrane, we see the flux used here is within the range possible for a single eucaryotic cell. Bacteria, of size only about $1\mu$m, will have lower maximum steady-state production rates, so the flux used here corresponds to the combined production rate of a small group of bacteria. This estimate of production rate based on the number of ribosomes gives an upper bound on the long-term average production rate. In addition, the signal proteins could be gradually accumulated and released in bursts, giving much higher local concentrations for short times. If the burst times are relatively short, e.g., less than a minute, the temporal resolution of passing sensors would be particularly useful since there would be some probability of a sensor passing during the high concentration produced by a burst. By contrast, attempting detection via a blood sample would give a temporal average of the release, which would be much lower than the high concentration during a burst.

# 4 Results for a Sensing Task Scenario

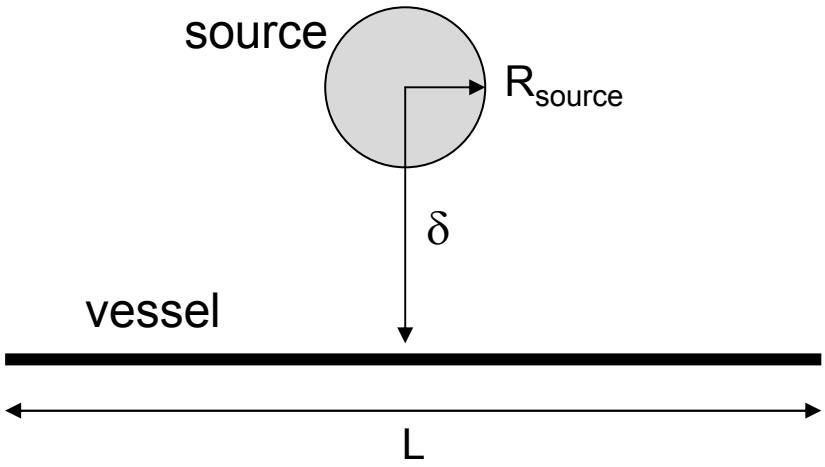


Figure 1: A vessel passing at distance $\delta$ from a source. A sensor, passing through the vessel at average speed $v_{\text{avg}}$, encounters changing concentration based on its distance from the source.

A sensor moving in a small vessel past the source has the highest detection rate when closest to the source. To estimate the expected number of detections, consider a small vessel segment, of length $L$, passing a minimum distance $\delta \geq R_{\text{source}}$ from the source as shown in Fig. 1. Since $L$ is large compared to $\eta$, most counts will occur over a distance of about $\eta$, i.e., during a time interval of about $\eta/v_{\text{avg}} = 100$ms. Counts significantly above background occur only when $\delta$ is comparable to or smaller than the concentration decay range $\eta$ and in the portion of the vessel within about $\eta$ of the closest approach to the

source. The expected number of counts is estimated by integrating $C(r)$, with the detection rate Eq. (1), for a sensor moving with speed $v_{\text{avg}}$ through the vessel. Because most counts occur near the source and $L \gg \eta$, we can extend the integration range to infinity without significantly changing the result. Thus the expected number of counts from the source for a sensor in a vessel passing at distance $\delta \geq R_{\text{source}}$ from the source is

$$\begin{aligned} E(\delta) &\approx \int_{-\infty}^{\infty} 4\pi DaC\left(\sqrt{\delta^2 + (v_{\text{avg}}t)^2}\right) dt \\ &= 8\pi a \, \frac{F R_{\text{source}}^2 \, e^{R_{\text{source}}/\eta}}{v_{\text{avg}}(1 + R_{\text{source}}/\eta)} \, K_0(\delta/\eta) \end{aligned} \tag{5}$$

taking $t = 0$ to be the time of closest approach and where $K_0$ is the modified Bessel function of the second kind and order zero. $K_0(\delta/\eta)$ decreases exponentially for $\delta \gg \eta$, giving little chance of detection in vessels passing far from the source.

A device passing through a vessel near the source has about $\eta/v_{\text{avg}} = 100$ms with high concentration. Thus a simple criterion for detecting a source is to pick a threshold $K$ and consider a source is found if a sensor detects at least $K$ molecules in a $100$ms interval. With diffusive motion of the molecules, the actual number of counts is a Poisson distributed random process with mean value given by Eq. (5). The detection probability, i.e., having at least $K$ events when the expected number is $\mu$, is

$$\Pr(\mu, K) = 1 - e^{-\mu} \sum_{n=0}^{K-1} \frac{\mu^n}{n!}$$

With a random distribution of vessels in the tissue volume, the number of vessels within a distance $\delta$ of the source is proportional to the volume, i.e., proportional to $\delta^3$. That is, there are many more vessels far from the source than close to it, with the number at distance $\delta$ then proportional to the rate of volume increase, i.e., $\delta^2$. In our case, $\delta$ extends to the limit of the tissue volume, given by $(4/3)\pi\delta_{\text{max}}^3 = V$, i.e., $\delta_{\text{max}} \approx 6$mm, although the precise upper limit is not important since the concentration is too low to give any significant chance for detections when $\delta$ is large compared to $\eta$. Combining these observations with the rate $\omega_{\text{sensor}}$ at which sensors enter the tissue volume, the rate sensors detect the source using a threshold $K$ is

$$\omega_{\text{source}} = \omega_{\text{sensor}} \int_{R_{\text{source}}}^{\delta_{\text{max}}} \frac{\delta^2}{\Delta} \Pr(E(\delta), K) \, d\delta \tag{6}$$

where $\Delta = \int \delta^2 \, d\delta$ over the same limits for $\delta$ normalizes the distribution of vessel distances from the source, and $\omega_{\text{sensor}}$ is given by Eq. (2).

A sensor can also give a false positive, i.e., receiving enough counts to reach the threshold in $100$ms from just the background concentration. Although the expected number of such detections in a given $100$ms interval is small, the background concentration extends throughout the vessels in the tissue volume giving many opportunities for false positives. With the parameters of Table 2, the expected count from background in $100$ms is $E_{\text{background}} = 0.8$. Since a sensor spends $\approx L/v_{\text{avg}} = 1$s in a small vessel in the tissue volume, the sensor has about 10 independent $100$ms opportunities to accumulate counts toward the detection threshold $K$. The rate of false positive detections is then

$$\omega_{\text{background}} \approx 10 \; \omega_{\text{sensor}} \Pr(E_{\text{background}}, K) \tag{7}$$

For a diagnostic task, we pick a detection threshold $K$ and a time $t$ for sensors to accumulate counts. The expected number of sensors reporting detections from the source and from the background are then $\omega_{\text{source}}t$ and $\omega_{\text{background}}t$, respectively. The actual number is also a Poisson process, so another decision criterion for declaring a source detected is the minimum number of sensors $n$ reporting a detection. Since expected count rate near the source is significantly larger than the background rate, the contributions to the counts from the source and background are nearly independent, so the probability for $n$ sensors to reach the threshold number of counts to indicate a source is

$$\Pr\left((\omega_{\text{source}} + \omega_{\text{background}})t, n\right)$$

and similarly for the false positives with counts based only on $\omega_{\text{background}}$.

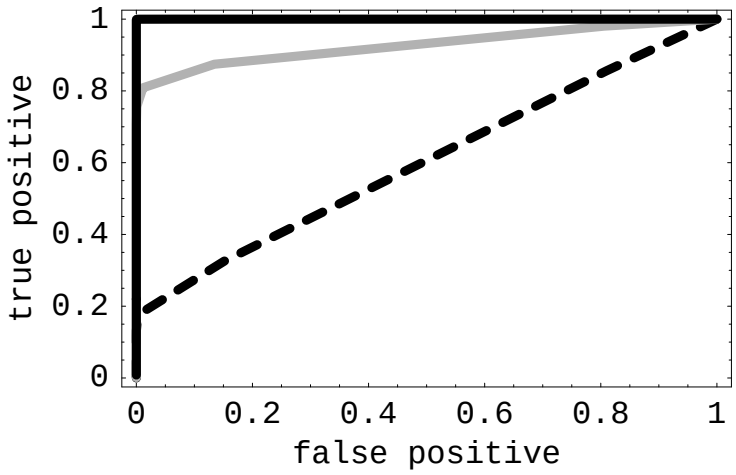


Figure 2: Probabilities of at least one sensor detecting a single source (true positive) and mistaking background concentration for a source (false positive) after sampling for $t = 1$, 10 and 60 seconds (dashed, gray and solid curves, respectively). Each curve corresponds to a range of values for the minimum count threshold, $K$, in a $100$ms interval that is taken to indicate detecting a source.

As an example of detection options, Fig. 2 shows the performance trade-off between true and false positive detections of a single source in the tissue volume after various sampling times with $n = 1$. The curves range from the lower-left corner (low detection rates) with a high threshold ($K = 100$) to the upper-right corner (high detection and high false positive rate) with a low threshold ($K = 1$). Sampling for only a minute gives excellent discrimination, e.g., $K = 15$ corresponds to the upper-left corner where sources are detected with over $99\%$ probability with the probability of a false positive only about $10^{-7}$.

For comparison, instead of using microscopic sensors, one could test for the chemical in a blood sample. This allows using chemical sensors outside the body, giving simpler fabrication and use. However, such a sample dilutes the chemical throughout the blood volume, resulting in considerably smaller concentrations than are available to microscopic sensors passing close to the source. As an example, suppose a single source described above produces the chemical for one day and all this production is delivered to the blood without any

degrading before a sample is taken. The source producing $\sim 2 \times 10^4$molecule$/$s then gives a concentration in the 5 liter blood volume of about $3 \times 10^{11}$molecule$/$m$^3$, which is less than $10^{-4}$ of the background concentration. In this case, the additional chemical released by the source would be undetectable against the variations in background concentration.

# 5 Extensions

The scenario described above illustrates behavior for one set of biologically relevant parameters. We can similarly consider other scenarios. For instance, with a smaller number of sensors, detection times would be correspondingly longer, or the sensors would only detect a larger number of sources. Thus, with $10^6$ sensors, a factor of 1000 fewer than in Table 1, achieving the discrimination shown in Fig. 2 would require about 1000 minutes, i.e., a day. Alternatively, for $10^6$ sensors with 1000 sources distributed randomly in the tissue volume $V$ instead of just one source, performance would be similar to that shown in the figure. Note that 1000 sources would occupy only a few millionths of the tissue volume.

Tissues vary in their density of small vessels and flow speed varies among these vessels. Uniform changes in these parameters alters both true and false positive detections similarly, so such changes mainly affect the required measurement time, as with changing the number of sensors. For example, Eq. (5) shows the counts from the source increase as fluid velocity decreases, due to the additional time a sensor spends near the source. Similarly, false positive detections also increase. Thus, tissue containing vessels with a different value of $v_{\rm avg}$ gives the same good performance by corresponding changes in the detection threshold $K$. A more challenging situation is if the changes are not uniform within the tissue volume. As an extreme case, suppose the few vessels passing near the source have fluid speed $v_{\rm avg}$ as given in Table 1 but the vast majority of vessels in the tissue volume have speeds only $1/100$ as fast, i.e., $10\mu$m$/$s. Due to the high discrimination seen in Fig. 2, the increase by a factor of 100 in the false positive rate still gives good performance: over $99\%$ probability to detect the source, using threshold $K = 15$, with the probability of a false positive rising to only about $10^{-5}$ compared with $10^{-7}$ in the original scenario. This example also illustrates how the sensor data indicates properties of the source neighborhood. In this case, a detection with $K = 15$ would likely have only a few counts more than this threshold. On the other hand, if the source were in a region of vessels with lower velocities, the observed counts would tend to be significantly higher than this threshold value. Thus analysis of the sensor data beyond simply whether the counts reach a threshold provides information on the source microenvironments.

The chemical properties of Table 2 are one example for a diagnostic application. Examining lower concentrations indicates the range of usefulness of these passive sensors. As the source concentration $C_{\rm source}$ decreases, suitable threshold choices giving good discrimination between true and false positives require increased observation time. Fig. 3 illustrates this behavior by showing the measurement times required for various choices of source and background concentrations. The figure shows measurement times ranging from 10 to $10^5$ seconds (i.e., about one day). As the source concentration decreases toward that of the background, the required measurement time grows very rapidly, so this diagnostic approach is no longer feasible. Instead, one would need to reduce the effective background

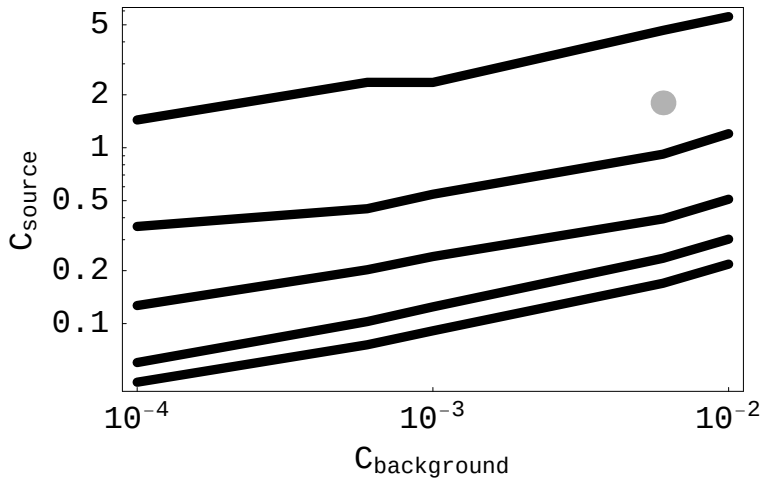


Figure 3: Combinations of source and background concentrations, on logarithmic scales, allowing good discrimination over differing measurement times, $t$. The curves correspond to $t = 10$, $10^2$, $10^3$, $10^4$ and $10^5$ seconds, from top to bottom. For these curves, good discrimination consists of choosing the count threshold $K$ in a $100$ms interval giving $99\%$ probability for true positive detection and at most $1\%$ probability for a false positive. Since the threshold $K$ is an integer, for most background concentrations, the smallest $K$ with false positive probability at most $1\%$ in fact gives a substantially smaller probability. Concentrations are measured in molecule$/\mu$m$^3$ with the gray point corresponding to parameters used in Table 2.

concentration or increase the time sensors spend near the source. For instance, better coarse localization could help, e.g., reducing the size from 1cm to 1mm would reduce false positive rate by a factor of 1000. Alternatively, testing for combinations of chemicals released by the source could also reduce the effective background rate. Finally, fabricating devices able to alter their surface to reversibly stick to vessel walls would allow time to collect more counts once they pass a threshold, thereby better distinguishing a source from background.

# 6 Discussion

This paper described how passively moving microscopic chemical sensors could give high resolution estimates of chemically distinctive regions in vivo. The performance estimates show devices with limited capabilities – specifically, without locomotion or communication with other devices – can nevertheless perform well for rapidly detecting chemical sources as small as a single cell. The devices use their small size and large numbers to allow at least a few to get close to the source, where concentration is much higher than background.

The inference procedure could account for sensor failures, e.g., requiring detection by several sensors as independent confirmation of at least one source. Occasional spurious extra counts by the sensors amount to an increase in the effective background concentration. As long as these extra counts are infrequent, and not significantly clustered in time, such errors will not significantly affect the overall accuracy of the results.

The model examined here treats the small vessels and the chemical sources as randomly distributed in the tissue volume. Systematic variation in the density and organization of the vessels will increase the variation in detected values. For instance, the tissue could have correlations between vessel density and the chemical sources (e.g., if those chemicals enhance or inhibit growth of new vessels). More accurate inferences require better models of how the chemicals move through the tissue to nearby blood vessels where they can then reach the sensors. Such movement of the chemicals could be anisotropic and include advection due to microfluidic flows in addition to diffusion. When several chemicals are involved, they could react, e.g., to form dimers, changing the concentrations with distance from the source. Furthermore, sensors will not move past sources at uniform speeds, as assumed with Eq. (5), but instead will change speed depending on their location within the vessel and interaction with cells in the fluid, giving further variation in the count rate. Nevertheless, the simple model discussed here indicates the devices could have high discrimination for sources as small as single cells. Some unmodeled sources of variability could be addressed by extending the sensing time or using more sophisticated inference methods. Moreover, with coarse localization during operation, the devices themselves could estimate some of this variation (e.g., changes in density of vessels in different tissue regions), and these estimates could be used to improve the inference instead of relying on average or estimated values for the tissue structure.

The high-resolution sensing scenario described in this paper is only one possible application of the devices. Other applications are of two types. First, we can use more sophisticated inference from the data after the devices are retrieved to estimate further properties of the chemical sources and their environments. Second, additional hardware capabilities could allow collecting more information, communicating some information to external observers during operation, or taking actions based on sensor observations.

For example, correlations in the measurements could distinguish a strong source from many weak sources spread throughout the tissue volume producing the chemical at the same total rate. Specifically, the strong source would give high count rates for a few sensors (those moving through vessels that pass near the source). On the other hand, multiple weak sources would have some detection in a larger fraction of the sensors, since more sensors would pass close enough to some source to reach the threshold of detection. More specific discrimination of these cases would be evident from the temporal distribution of counts.

Including nanoscale sensors for fluid motion [29] on the devices would allow mapping the in vivo microfluidic behavior in small vessels. Such fluid flow sensors would allow correlating chemical detections with properties of the flow and the vessel geometry (e.g., branching and changes in vessel size or permeability to fluids). Detailed information on fluid behavior would also give design constraints on more elaborate devices with locomotion capability.

This paper considered homogeneous devices with the same control program and capabilities. More generally, devices could have various capabilities. For instance, some devices could have multiple specialized sensors for specific, rare chemicals, while others have more computational capability to evaluate spatial or temporal patterns of chemical activity. Initially homogeneous devices could also develop temporary or permanent differences based on their history, e.g., the level of their energy reserves or remaining amount of free memory influencing their behavior. Heterogeneity involving devices of different sizes

can also be useful when the diagnostic task involves chemical sources on several biological scales from individual cells to tissues [30, 31].

As an example of an application for taking action based on sensor readings, the devices could carry specific drugs to release only near cells matching a prespecified chemical profile [1, 32] as an extension of a recent in vitro demonstration of this capability using DNA computers [33]. Devices determining when to take such actions could combine chemical information with other sensing modalities, such as optical scattering in cells, as has been demonstrated to distinguish some cancer from normal cells in vitro [34].

Safety is an important design criterion for medical applications of microscopic sensors, both in terms of the devices themselves and subsequent actions based on diagnostic results. Thus, evaluation of the protocol should consider its accuracy allowing for sensor errors, failures of individual devices or variations in environmental parameters. For the distributed sensing discussed in this paper, aggregation of many devices' measurements provides robustness against these variations, e.g., as illustrated using DNA computing to respond to chemical patterns [33]. Physically, the devices must be compatible with their biological environment [35, 36], for at least enough time to complete their task. Appropriately engineered surfaces [35] should prevent unwanted inflammation or immune system reactions during their operation. However, even if individual devices are inert, too large a number in the circulation would be harmful. From Table 1, sensors occupy a fraction $(4/3)\pi a^3 \rho_{\mathrm{sensor}} \approx 10^{-6}$ of the volume inside the vessels. This value is well below the fraction, about $10^{-3}$, of micron-size particles experimentally demonstrated to be safely tolerated in the circulatory system of at least some mammals [35]. Thus the number of sensors used in the protocol of this paper is unlikely to be a safety issue.

The estimates obtained in this paper with plausible biophysical parameters show relatively modest molecular hardware capabilities could provide useful in vivo sensing capabilities. The usefulness of such capability depends on the types of chemical events of interest, and particularly the extent to which local spatial or temporal variations in the chemicals give information beyond single measurements, e.g., to reduce background detections, thereby benefiting from the computational capabilities of the devices. Using these sensors in research studies of tissue microenvironments will enable better inferences from their data, and quantify the further benefits possible with more capable devices.

# References


[1] Robert A. Freitas Jr. *Nanomedicine*, volume I: Basic Capabilities. Landes Bioscience, Georgetown, TX, 1999. Available at www.nanomedicine.com/NMI.htm.

[2] Kelly Morris. Macrodoctor, come meet the nanodoctors. *The Lancet*, 357:778, March 10 2001.

[3] Balazs L. Keszler, Istvan J. Majoros, and James R. Baker Jr. Molecular engineering in nanotechnology: Structure and composition of multifunctional devices for medical application. In *Proc. of the Ninth Foresight Conference on Molecular Nanotechnology*, 2001.

[4] Jose Berna et al. Macroscopic transport by synthetic molecular machines. *Nature Materials*, 4:704–710, 2005.

[5] C. P. Collier et al. Electronically configurable molecular-based logic gates. *Science*, 285:391–394, 1999.

[6] H. G. Craighead. Nanoelectromechanical systems. *Science*, 290:1532–1535, 2000.

[7] J. Fritz et al. Translating biomolecular recognition into nanomechanics. *Science*, 288:316–318, 2000.

[8] Fernando Patolsky and Charles M. Lieber. Nanowire nanosensors. *Materials Today*, 8:20–28, April 2005.

[9] Paul E. Sheehan and Lloyd J. Whitman. Detection limits for nanoscale biosensors. *Nano Letters*, 5(4):803–807, 2005.

[10] S.-Y. Wang and R. Stanley Williams, editors. *Nanoelectronics*, volume 80. Springer, March 2005. Special issue of *Applied Physics A*.

[11] Tuan Vo-Dinh, Paul Kasili, and Musundi Wabuyele. Nanoprobes and nanobiosensors for monitoring and imaging individual living cells. *Nanomedicine: Nanotechnology, Biology, and Medicine*, 2:22–30, 2006.

[12] Frank Natterer. *The Mathematics of Computerized Tomography*. Soc. for Industrial and Applied Math (SIAM), Philadelphia, 2001.

[13] Daniel Arbuckle and Aristides A. G. Requicha. Active self-assembly. In *Proc. of the IEEE Intl. Conf. on Robotics and Automation*, pages 896–901, 2004.

[14] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, Oxford, 1999.

[15] Owen Holland and Chris Melhuish. Stigmergy, self-organization and sorting in collective robotics. *Artificial Life*, 5:173–202, 1999.

[16] Amit Dhariwal, Gaurav S. Sukhatme, and Aristides A. G. Requicha. Bacterium-inspired robots for environmental monitoring. In *Proc. of the IEEE Intl. Conf. on Robotics and Automation*, 2004.

[17] Adriano Cavalcanti and Robert A. Freitas Jr. Autonomous multi-robot sensor-based cooperation for nanomedicine. *Intl. J. of Nonlinear Sciences and Numerical Simulation*, 3:743–746, 2002.

[18] Tad Hogg and Bernardo A. Huberman. Dynamics of large autonomous computational systems. In Kagan Tumer and David Wolpert, editors, *Collectives and the Design of Complex Systems*, pages 295–315. Springer, New York, 2004.

[19] Kristina Lerman et al. A macroscopic analytical model of collaboration in distributed robotic systems. *Artificial Life*, 7:375–393, 2001.

[20] Aram Galstyan, Tad Hogg, and Kristina Lerman. Modeling and mathematical analysis of swarms of microscopic robots. In P. Arabshahi and A. Martinoli, editors, *Proc. of the IEEE Swarm Intelligence Symposium (SIS2005)*, pages 201–208, 2005.

[21] E. M. Purcell. Life at low Reynolds number. *American Journal of Physics*, 45:3–11, 1977.

[22] Steven Vogel. *Life in Moving Fluids*. Princeton Univ. Press, 2nd edition, 1994.

[23] Y. C. Fung. *Biomechanics: Circulation*. Springer, NY, 2nd edition, 1997.

[24] George E. M. Karniadakis and Ali Beskok. *Micro Flows: Fundamentals and Simulation*. Springer, Berlin, 2002.

[25] Howard C. Berg and Edward M. Purcell. Physics of chemoreception. *Biophysical Journal*, 20:193–219, 1977.

[26] Howard C. Berg. *Random Walks in Biology*. Princeton Univ. Press, 2nd edition, 1993.

[27] Zhong Lin Wang and Jinhui Song. Piezoelectric nanogenerators based on zinc oxide nanowire arrays. *Science*, 312:242–246, 2006.

[28] Charles A. Janeway et al. *Immunobiology: The Immune System in Health and Disease*. Garland, 5th edition, 2001.

[29] Shankar Ghosh et al. Carbon nanotube flow sensors. *Science*, 299:1042–1044, 2003.

[30] D. Sretavan, W. Chang, C. Keller, and M. Kliot. Microscale surgery on axons for nerve injury treatment. *Neurosurgery*, 57(4):635–646, 2005.

[31] Tad Hogg and David W. Sretavan. Controlling tiny multi-scale robots for nerve repair. In *Proc. of the 20th Natl. Conf. on Artificial Intelligence (AAAI2005)*, pages 1286–1291. AAAI Press, 2005.

[32] Robert A. Freitas Jr. Pharmacytes: An ideal vehicle for targeted drug delivery. *Journal of Nanoscience and Nanotechnology*, 6:2769–2775, 2006.

[33] Yaakov Benenson, Binyamin Gil, Uri Ben-Dor, Rivka Adar, and Ehud Shapiro. An autonomous molecular computer for logical control of gene expression. *Nature*, 429:423–429, 2004.

[34] P. L. Gourley et al. Ultrafast nanolaser flow device for detecting cancer in single cells. *Biomed Microdevices*, 7:331–339, 2005.

[35] Robert A. Freitas Jr. *Nanomedicine*, volume IIA: Biocompatibility. Landes Bioscience, Georgetown, TX, 2003. Available at www.nanomedicine.com/NMIIA.htm.

[36] Andre Nel et al. Toxic potential of materials at the nanolevel. *Science*, 311:622–627, 2006.