

A Mathematical Channel Optimization Formula for Active Transport Molecular Communication

Nariman Farsad, Andrew W. Eckford
Department of Computer Science and Engineering
York University
4700 Keele Street, Toronto, ON, Canada M3J 1P3
Email: nariman@cse.yorku.ca, aeckford@yorku.ca

Satoshi Hiyama
Research Laboratories
NTT DOCOMO Inc.
Yokosuka, Kanagawa, Japan
Email: hiyama@nttdocomo.co.jp

Abstract—In this paper, a mathematical optimization formula for estimating the optimal channel dimensions of active transport molecular communication is presented. More specifically, rectangular channels with constant microtubule (MT) concentration are considered. It is shown, both using our formula and using Monte Carlo simulations, that square-shaped channels are optimal. Furthermore, when the value of time per channel use is on the order of a few minutes, which is the range of interest for a lot of potential applications such as diagnostic chips for healthcare, it is shown that our optimization formula can quickly and accurately estimate the optimal channel dimensions.

I. INTRODUCTION

Molecular communication, which is used in nature for intercellular and intracellular communication, relies on particles such as molecules for transmission of information. These information carrying molecules propagate in fluidic microchannels, either passively or actively. In passive transportation, the information molecules diffuse in the channel via Brownian motion from the transmitter to the receiver, while in active transport, the information molecules are transported as cargoes using a molecular motor system which could consist of static kinesin motors and microtubule (MT) carries. Moreover, molecular communication has attracted attention as a communication method in nanonetworks [1], [2].

Related works in this field include a general formulation of molecular communication as a timing channel under Brownian motion [3], [4], an analysis of information transfer rates using molecular motors [5], [6], mathematical channel models for continuous diffusion [7], binary concentration-encoded molecular communication [8]. Concerning active transport molecular communication, in [9], [10] basic channel design is considered for designing molecular sorters and rectifiers. Our previous work includes a simple mathematical transport model for active transport propagation [11], and optimization of the transmission zone and vesicular encapsulation [12].

In this work we extend our earlier work on transmission zone design, and consider the optimization of channel shape in active transport molecular communication. Design and optimization of these channels in wet labs are extremely time consuming and laborious. Therefore, previous works have relied on computer simulations for design and optimization. However, computer simulations can also become computationally intensive and tedious in solving design problems. In this

paper, we derive a mathematical optimization formula that can be used to estimate the optimal channel dimensions, thereby eliminating the need for computationally intensive simulations. In [13], using computer simulations we present general rules for channel design and optimization, where only a single MT is used for transporting the information particles. In practice, however, the number of MTs inside the channel is related to the concentration of the MTs. In this work, we extend our previous work and consider the more realistic framework in which the concentration of MTs dictates the number of MTs inside the channel. Our contributions are as follows:

- We present a mathematical optimization formula that can be used to estimate the optimal channel dimensions quickly. Although, finding the optimal channel dimensions is also possible through computer simulations, such simulations would be computationally intensive and extremely time consuming.
- Using Monte Carlo simulations we demonstrate that for molecular communication systems with constant concentration of MTs, the optimal channel is generally square-shaped. We show that this result is in agreement with the estimates of our optimization formula.
- We show that for small values of time per channel use (in the order of a few minutes), our optimization formula can accurately estimate the optimal channel dimensions. Although the optimization formula loses its accuracy for larger values of time per channel use, this result is significant since numerous potential applications such as diagnostic chips for healthcare, rely on small values of time per channel use.

In [13], it is shown that for channels employing a single MT as the carrier of the information particles, the smaller the perimeter of the channel, the higher the channel capacity. It is observed that the channel capacity is higher in this case because MTs on average follow the walls of the channel closely. Therefore, the smaller the channel dimensions, the higher the information transmission rate. However, in this work, using the proposed formula we show that for channels with constant concentration of MTs (i.e. multiple MTs could be inside the channel), the ratio of the volume to the perimeter of the channel becomes the important factor in optimization.

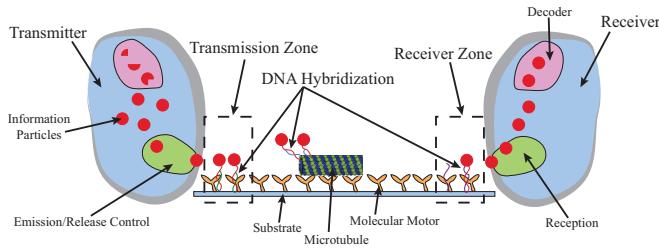


Fig. 1. Depiction of the communication environment.

In this case, the rule for optimization is non-trivial, and depending on the parameters larger channel dimensions could be the optimal.

II. ACTIVE TRANSPORT MOLECULAR COMMUNICATION ENVIRONMENT

Our molecular communication environment is similar to the ones in [14], [15], [16]. We use a rectangular propagation environment (with slightly rounded corners), consisting of a *transmission zone* and a *receiver zone*. Message-bearing particles originate at the transmission zone. These particles are initially assumed to be anchored to the transmission zone until MT filaments, moving over molecular motor tracks that cover the whole environment, pick up and transport the information carrying particles from the transmission zone to the receiver zone.

As in [16], we assume that the microchannel environment is lined with static kinesin molecular motors, and that these motors cause MT filaments to propagate along their surface. Furthermore, we assume deoxyribonucleic acid (DNA) hybridization bonds are used to anchor, load and unload the information particles. This process is summarized in Figure 1 and the reader is referred to [16] for detailed explanation.

Previous works have relied on computer simulations to study these systems. In [9], [10] the motion of MT filaments over stationary kinesin molecular motors was simulated, and in [14], [15] a complete simulation of the molecular communication environment was presented. In this work, to verify the accuracy of our proposed optimization formula, we use computer simulations similar to the ones used in [14], [9]. In particular, we simulate the molecular communication system by employing the mathematical models of the motion of the MTs.

In the rest of this paper, we assume the channel is rectangular with the transmission zone on the left and the receiver zone on the right side of the channel, as shown in Figure 2. We assume the width and length of the channel, as well as the size of the transmission and the receiver zones, are variable. However, we assume for a given channel the dimensions of the transmission zone and the receiver zone are the same. Therefore, our rectangular channel environment can be characterized using three parameters: the width of the channel, the length of the channel, and the separation distance between the transmission zone and the receiver zone.

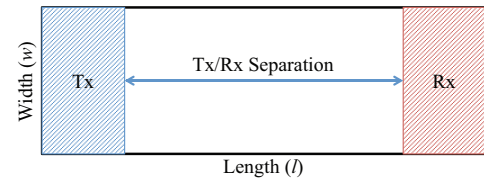


Fig. 2. The rectangular molecular communication environment with the transmission zone on the left and the receiver zone on the right.

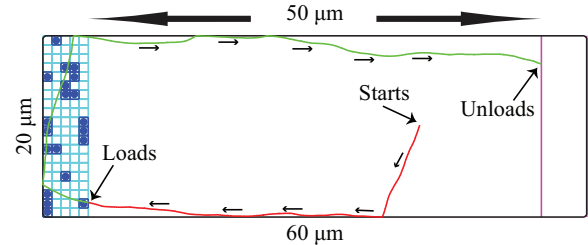


Fig. 3. A sample trajectory of active transport. The path of the MT is the line in the middle of the rectangular channel.

In order to capture the loading process in the simulations, we use the grid loading structure proposed in [14]. For loading an information particle, the MT filament must drive close to the anchored particle. Therefore, we divide the transmission zone into a square grid, where the length of each square in the grid is the same as the diameter of the particles. We then distribute particles randomly and uniformly among the squares in the grid. In general, we assume that the MTs can load multiple particles, which we know to be possible based on lab experiments; thus, if a MT enters a square which is occupied by a particle, and it has an empty loading slot available, we assume the MT loads that particle. For unloading, we assume all the loaded particles are unloaded as soon as a MT enters the receiver zone. Figure 3 shows a sample trajectory with the loading and unloading mechanism. Notice that the transmission and the receiver zones in this figure are rectangular strips across the width of the channel.

III. INFORMATION RATE AND TRANSPORT MODEL

In this paper, we consider information transmission as a mass transfer problem. In the simplest possible conception of this scheme, the particles themselves are not information-bearing, and a message is conveyed in the *number* of particles released by the transmitter. For example, if a maximum of three particles may be used, from a traditional communication system perspective, we may form messages two bits long (i.e., $\log_2 4$): “00” for 0 particles, “01” for 1 particle, “10” for 2 particles, and “11” for 3 particles. However, this message might not be perfectly conveyed to the receiver: given a time limit T for the communication session, it is possible that some of the particles will not arrive at the receiver after T has elapsed. We refer to time limit T as *time per channel use*. In other words, time per channel use is a predefined amount of time representing the time duration for a single message

transmission session, and it is one of the parameters of the molecular communication system.

Since we have assumed that messages are encoded in the number of particles, let $X \in \mathcal{X} = \{0, 1, 2, \dots, x_{max}\}$ represent the number of information particles released into the medium by the transmitter, $Y \in \mathcal{Y}$ represent the number that arrive at the destination after time per channel use T , and x_{max} be the maximum number of particles the transmitter can release per channel use.

The maximum rate at which any communication system can reliably transmit information over a noisy channel is bounded by a limit called *channel capacity*. The channel capacity can be calculated as,

$$C = \max_{f_X(x)} I(X; Y), \quad (1)$$

where $I(X; Y)$ is the mutual information between X and Y . Mutual information is defined as

$$I(X; Y) = E \left[\log_2 \frac{f_{Y|X}(y|x)}{\sum_x f_{Y|X}(y|x) f_X(x)} \right], \quad (2)$$

where, $f_{Y|X}(y|x)$ represents the probability of receiving symbol y at the destination, given that symbol x was transmitted by the source; $f_X(x)$ represents the probability of transmitting symbol x ; and $E[\cdot]$ represents expectation.

IV. CHANNEL DESIGN AND OPTIMIZATION

In the previous section we introduced the notion of time per channel use, which is an important parameter of the molecular communication system. In this section, we derive an estimated optimization formula which could be used to estimate the optimal channel dimensions based on parameters such as time per channel use. (Use of the formula, though approximate, is much faster than performing a high-fidelity simulation.)

A. Important Optimization Factors

One of the most important factors affecting the channel capacity is the number of MT trips (K) during a single time per channel use interval T . We define a single MT *trip* as the movement of the MT from anywhere in the channel to the transmission zone and then the receiver zone. For example, a single MT trip is shown in Figure 3. During a single trip, a MT starts its movement from anywhere in the channel, visits the transmission zone and then visits the receiver zone. During each trip the MT could potentially carry a number of information particles from the transmission zone to the receiver zone. Therefore, the total number of MT trips, K , during the time per channel use interval, T , limits the number of information particles that could be potentially delivered to the receiver.

In practice, there are typically more than one MTs inside the channel. Furthermore, the number of MTs inside the channel is not a constant, but is dependent on the volume of fluid in the channel: the DNA-covered MTs are prepared in chemical solutions, and therefore have a constant concentration inside the solution. This solution, with constant concentration of the MTs, is then poured into the microchannel. Hence, the larger

the volume of fluid in the channel the higher the number of MTs in the channel.

B. Mathematical Model

Let the random variable $K^{(T)}$ be the number of MT trips between the transmitter and the receiver in time duration T seconds, where T is time per channel use. Let $S^{(T)}$ be the number of trips for a single MT during T seconds, C be the concentration of MTs as number of MTs per unit volume, w be the width of the channel, l be the length of the channel, and h be the height of the channel. In [13], it was shown that a good estimate for the number of MT trips, when a single MT is employed is given by,

$$E[S^{(T)}] \approx \left\lfloor \frac{v_{avg} T}{P(w, l)} \right\rfloor, \quad (3)$$

where $\lfloor \cdot \rfloor$ is the floor function, v_{avg} is the average speed of the MTs, T is time per channel use, and $P(w, l)$ is the perimeter of the channel given by,

$$P(w, l) = 2w + 2l. \quad (4)$$

The floor function is used because the number of MT trips during T seconds is an integer.

Let M be the number of MTs inside the channel. Because we assume the concentration of the MTs is constant, given the channel dimensions, the number of MTs in the channel is given by,

$$M = V(w, l, h)C, \quad (5)$$

where C is the number of MTs per unit volume, and $V(w, l, h)$ is the volume of the channel given by,

$$A(w, l) = w \times l \times h. \quad (6)$$

Let the random variables $K_1^{(T)}, K_2^{(T)}, K_3^{(T)}, \dots, K_M^{(T)}$ be the number of trips for each of the M MTs during T seconds. The total number of MT trips during T seconds is therefore given by,

$$K^{(T)} = \sum_{i=1}^M K_i^{(T)}. \quad (7)$$

The average number of MT trips during T seconds is therefore calculated as,

$$E[K^{(T)}] = \sum_{i=1}^M E[K_i^{(T)}]. \quad (8)$$

We assume that the number of trips for individual MTs are independent and identically distributed, because they are chemically similar and don't interact with each other. Due to the identical distributions, the equation simplifies to

$$E[K^{(T)}] = \sum_{i=1}^M E[S^{(T)}] = M E[S^{(T)}]. \quad (9)$$

Using the approximation shown in Equation (3), and Equation (5), the total number of trips can be estimated as

$$E[K^{(T)}] \approx [V(w, l, h)C] \times \left\lfloor \frac{v_{avg} T}{P(w, l)} \right\rfloor. \quad (10)$$

C. Channel Optimization

Given a particular time per channel use value (T), a particular average speed for the MTs (v_{avg}), and a particular density of MTs per unit volume (C), the average number of MT trips between the transmission and the receiver zone is dependent of the size of the channel as shown in Equation (10). For simplicity, without loss of generality, in the rest of this work we assume that all channels have a constant height equal to $10\mu\text{m}$.

Since we assume that the information is encoded in the number of information particles, the information transmission rate is limited by the number of MT trips. Therefore, to maximize the information rate we would like to increase the average number of MT trips. Using Equation (10) we have

$$\max(E[K^{(T)}]) \approx \max_{w,l} \left([V(w, l, h)C] \times \left[\frac{v_{\text{avg}}T}{P(w, l)} \right] \right). \quad (11)$$

Because we assume the height of the channel is a constant, we maximize the channel parameters with respect to the width and the length of the channel only. The values of w and l that maximize the information rate, are the optimal channel dimensions. Because all the parameters except the dimensions of the channel are assumed to be constants, we observe that an important factor in optimization is the ratio of the cross sectional area to the perimeter of the channel. Therefore, we expect the square channels to have the largest average number of MT trips.

Figure 4 shows the average number of MT trips estimated using Equation (10) as a function of the width and the length of the channel. For this graph $C = 0.001$ MT/fL, $T = 200$ seconds, and $v_{\text{avg}} = 0.5 \mu\text{m/s}$. From the plot we observe that the square channel size of $31 \mu\text{m} \times 31 \mu\text{m}$ is estimated to have the largest number of MT trips between the transmitter and the receiver. Please note, the jagged features of the graph are due to the floor function used in the optimization estimation formula. Furthermore, for larger channel dimensions, the number of MT trips drops to zero since for the given time per channel use, no single MT can complete one trip.

V. RESULTS

In this section, we compare the optimal channel estimations obtained using Equation (10) with actual Monte Carlo simulations. For our Monte Carlo simulations we use the same simulation techniques presented in [12] with the following parameters: simulation time steps of $\Delta T = 0.1$ seconds, MT diffusion coefficient $D = 2.0 \cdot 10^{-3} \mu\text{m}^2/\text{s}$, average speed of the MT $v_{\text{avg}} = 0.5 \mu\text{m/s}$, and persistence length of the MT trajectory $L_p = 111 \mu\text{m}$. We also assume the size of the information particles is $1\mu\text{m}$, the average length of the MTs is $10\mu\text{m}$, and each MT can load up to 5 information particles in one trip from the transmission zone to the receiver zone. These parameters are all selected based on experimental observations of DNA covered MTs moving over a kinesin covered substrate.

We keep channel height constant at $h = 10 \mu\text{m}$. Of the other two dimensions, we define the width and length of the channel

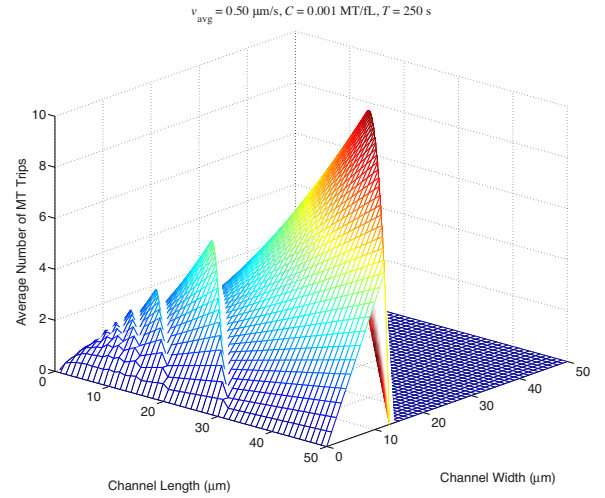


Fig. 4. Equation (10) plotted as a function of the width and the length of the channel.

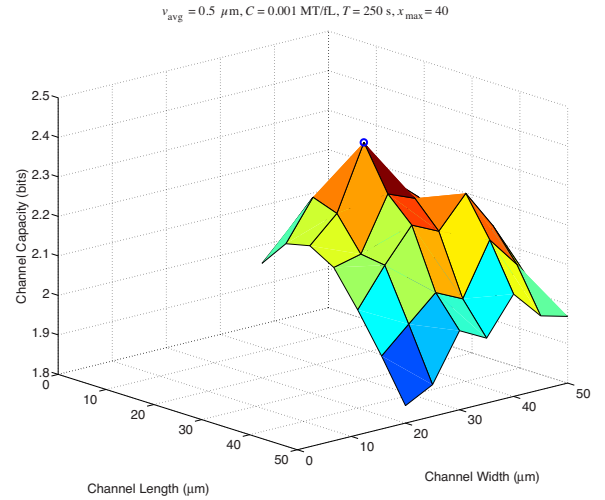


Fig. 5. Channel capacity in bits versus the width and the length of the channel (obtained using simulations).

as the smaller dimension and larger dimension of the channel, respectively. We assume the concentration of MTs is $C = 0.001$ MT/fL. We assume that the transmission and receiver zones are always along the width of the channel as shown in Figure 2. For the width and the length of the channel, we consider the range from $20 \mu\text{m}$ to $50 \mu\text{m}$ in steps of $5 \mu\text{m}$. We assume the area of the transmission and the receiver zones are always between 80 - $120 \mu\text{m}^2$. Based on previous observation, this slight difference in the areas of the transmission and the receiver zones across different channel dimensions does not effect the channel capacity significantly.

A. Square versus Rectangular Channels

From our Monte Carlo simulations we observe that square, or nearly square channels achieve the highest channel capacity. This result is in agreement with the estimates of our optimization formula. As shown in Figure 4, optimal channels

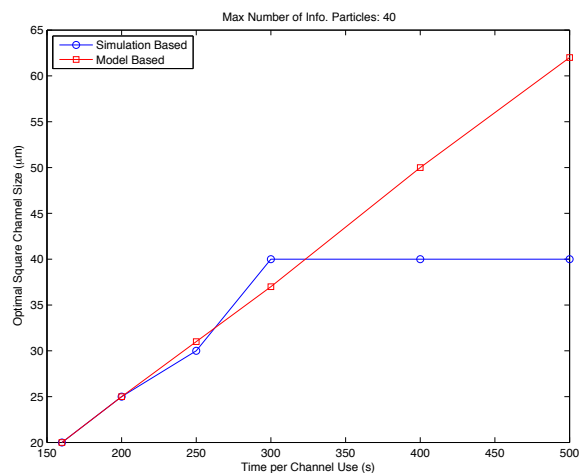


Fig. 6. The size of the optimal square channel in μm versus the time per channel use in seconds.

are estimated to be squares when our optimization formula is employed. Similarly, Figure 5 shows the channel capacity in bits versus the width and the length of the channel for time per channel use of 250 seconds, and $x_{max} = 40$ (i.e. maximum number of information particles that could be released by the transmitter is 40). In this plot, the maximum is shown by the blue circle. For this particular value of time per channel use, the maximum channel capacity is achieved when the channel dimensions are $30 \mu\text{m} \times 30 \mu\text{m}$, which is in agreement with the optimal channel dimensions obtained from Figure 4.

B. Optimal Channel Results

Because in the previous section we showed that the optimal channels are generally square-shaped, we only consider square channels here. In Figure 6, we compare the optimal channel dimensions obtained using our approximate channel optimization formula and Monte Carlo simulations. This figure shows both the excellent performance of our model for small time per channel use, but also the limits of our model for large time per channel use. The red plot is the estimation of the optimal channel size obtained from our proposed optimization formula, while the blue plot represents the optimal channel size obtained through simulations. From the figure we can see that for time per channel use values smaller than 300 seconds, our approximate formula can estimate the optimal channel dimensions accurately. However, as the value of time per channel use increases, the error in the estimation increases because of two effects. First, the variance of number of MT trips increases because the channels can be larger and include more MTs. Our approximate model does not capture the effects of variance, and only maximizes the average number of trips. Second, because our approximate model estimates the number of trips for a single MT using Equation (3), as the size of the channel increases and thereby number of MTs increase the error in the estimation also increases.

Although, channel optimization can be performed using computer simulations, these simulations are computationally

intensive and time consuming. For example, the high-fidelity Monte Carlo simulation results presented in this work required about a month on 40 different CPU cores to complete. Using the presented optimization formula however, the optimal channel dimensions can be estimated in seconds on a single computer. Furthermore, despite its limitations, since many potential applications of molecular communication such as diagnostic chips for healthcare require quick communication with low values of time per channel use, our optimization formula can be applied in design and optimization of these channels.

REFERENCES

- [1] I. F. Akyildiz, F. Brunetti, and C. Blazquez, "Nanonetworks: A new communication paradigm," *Computer Networks*, vol. 52, no. 12, pp. 2260–2279, 2008.
- [2] S. Hiyama and Y. Moritani, "Molecular communication: harnessing biochemical materials to engineer biomimetic communication systems," *Nano Communication Networks*, vol. 1, no. 1, pp. 20–30, 2010.
- [3] A. W. Eckford, "Nanoscale communication with brownian motion," in *Proc. of the 41st Annual Conference on Information Sciences and Systems*, (Baltimore, MD), pp. 160–165, 2007.
- [4] B. Atakan and O. Akan, "An information theoretical approach for molecular communication," in *Proc. of the 2nd International Conference on Bio-Inspired Models of Network, Information and Computing Systems*, (Budapest, Hungary), pp. 33–40, 2007.
- [5] M. J. Moore, T. Suda, and K. Oiwa, "Molecular communication: modeling noise effects on information rate," *IEEE Transactions on NanoBioscience*, vol. 8, no. 2, pp. 169–180, 2009.
- [6] A. W. Eckford, "Timing information rates for active transport molecular communication," in *Nano-Net*, vol. 20 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 24–28, Springer Berlin Heidelberg, 2009.
- [7] M. Pierobon and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 4, pp. 602–611, 2010.
- [8] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "On the characterization of binary concentration-encoded molecular communication in nanonetworks," *Nano Communication Networks*, vol. 1, no. 4, pp. 289–300, 2010.
- [9] T. Nitta, A. Tanahashi, M. Hirano, and H. Hess, "Simulating molecular shuttle movements: Towards computer-aided design of nanoscale transport systems," *Lab on a Chip*, vol. 6, no. 7, pp. 881–885, 2006.
- [10] T. Nitta, A. Tanahashi, and M. Hirano, "In silico design and testing of guiding tracks for molecular shuttles powered by kinesin motors," *Lab on a Chip*, vol. 10, no. 11, pp. 1447–1453, 2010.
- [11] N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "A simple mathematical model for information rate of active transport molecular communication," in *Proc. IEEE INFOCOM Workshops*, (Shanghai, P. R. China), pp. 473–478, 2011.
- [12] N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "Quick system design of vesicle-based active transport molecular communication by using a simple transport model," *Nano Communication Networks*, vol. 2, no. 4, pp. 175–188, 2011.
- [13] N. Farsad, A. W. Eckford, and S. Hiyama, "Channel design and optimization of active transport molecular communication," in *Proc. 6th International ICST Conference on Bio-Inspired Models of Network, Information, and Computing Systems*, (York, England), 2011.
- [14] N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "Information rates of active propagation in microchannel molecular communication," in *Proc. of the 5th International ICST Conference on Bio-Inspired Models of Network, Information, and Computing Systems*, (Boston, MA), p. 7, 2010.
- [15] A. W. Eckford, N. Farsad, S. Hiyama, and Y. Moritani, "Microchannel molecular communication with nanoscale carriers: Brownian motion versus active transport," in *Proc. of the IEEE International Conference on Nanotechnology*, (Seoul, South Korea), pp. 854 – 858, 2010.
- [16] S. Hiyama, R. Gojo, T. Shima, S. Takeuchi, and K. Sutoh, "Biomolecular-motor-based nano- or microscale particle translocations on DNA microarrays," *Nano Letters*, vol. 9, no. 6, pp. 2407–2413, 2009.