Quantitative analysis of virus and plasmid trafficking in cells

Thibault Lagache and Emmanuel Dauty

Département de Mathématiques et de Biologie, Ecole Normale Supérieure, 46 rue d'Ulm 75005 Paris, France

David Holcman

Department of Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel
and Département de Mathématiques et de Biologie, Ecole Normale Supérieure, 46 rue d'Ulm 75005 Paris, France

Intracellular transport of DNA carriers is a fundamental step of gene delivery. By combining both theoretical and numerical approaches we study here single and several viruses and DNA particles trafficking in the cell cytoplasm to a small nuclear pore. We present a physical model to account for certain aspects of cellular organization, starting with the observation that a viral trajectory consists of epochs of pure diffusion and epochs of active transport along microtubules. We define a general degradation rate to describe the limitations of the delivery of plasmid or viral particles to a nuclear pore imposed by various types of direct and indirect hydrolysis activity inside the cytoplasm. By replacing the switching dynamics by a single steady state stochastic description, we obtain estimates for the probability and the mean time for the first one of many particles to go from the cell membrane to a small nuclear pore. Computational simulations confirm that our model can be used to analyze and interpret viral trajectories and estimate quantitatively the success of nuclear delivery.

I. INTRODUCTION

The study of the motion of many particles inside a biological cell is a problem with many degrees of freedom and a large parameter space. The latter may include the different diffusion constants of the different species, velocities along microtubules, their number, the geometry of cell and nucleus, the number and sizes of nuclear pores, the various degradation factors, and so on. The experimental and numerical exploration of this multidimensional parameter space is limited perforce to a small part thereof, due to the great complexity of the biological cell. A great reduction in complexity is often achieved by coarse graining the complex motion by means of effective equations and their explicit analytical solutions, which is the approach we adopt here. We are specifically concerned with finding a concise description of virus and plasmid trafficking in cell cytoplasm.

Recent studies of natural viruses [1–3] and synthetic (amphiphiles) DNA carriers [4] uncover details of the cellular pathways and the complexity of cellular infection. Viruses invade mammalian cells through multistep processes, which begin with the uptake of particles in endosomal compartments. After escape, the particle moves inside the cytoplasm, and the journey ends at a nuclear pore where its DNA is imported. We focus here only on the free cytoplasmic trafficking, a step that both natural and synthetic DNA carriers share. Cytoplasmic trafficking remains a major obstacle to gene delivery, because the cytosolic motion of large DNA molecules is limited by physical and chemical barriers of the crowded cytoplasm [5,6]. Whereas molecules smaller than 500 kDa can diffuse, larger cargos such as viruses or nonviral DNA particles, require an active transport system [7]. Viral infection is much more efficient than gene transfer using polymer-or lipid-based vectors, where a large amount of endocytozed DNA (typically over 100 000 copies of the gene) is required to produce a cellular response, while only a few copies seem to be necessary in the case of viruses.

A recent study [8] showed that microtubules shape the distribution of molecular motors and vesicle trafficking inside the cell cytoplasm by means of a combination of experiments and numerical simulations. The distribution of viral species was analyzed in Refs. [9,10] by means of the mass-action law and Brownian simulations, but not at a single particle level. In addition, the problem of a viral particle reaching a small nuclear pore was not considered there and this question is central here. In general, the mechanism of a single DNA and viral delivery to a small nuclear pore in the cytoplasm is still an open question. The mean time for a random particle to arrive to a small target has been studied in Ref. [11] and in the context biophysical questions and cell biology in Refs. [12–14]. We propose here a coarse-grained reduced description of viral trafficking in the cytoplasm and compare it to plasmid diffusion. Specifically, we are interested in the probability p_N and the mean time τ_N for a DNA carrier or a virus to arrive to a small nuclear pore. The evaluation of these quantities calls for a quantitative approach to the description of particle trajectories at an individual level and also, to quantify the role of the cell organization.

The paper is organized as follow. We start with the observations that a viral movement can be described as a combination of intermittent switches between pure Brownian diffusion and active transport along microtubules [15] (Fig. 1), while DNA motion can be characterized as pure Brownian. We also account for multiple factors involved in degradation, such as hydrolyzation, destruction by proteasomes, or any other factors that prevent irreversibly the particle from reaching a nuclear pore such as entanglements in the cytoskeleton that definitively trap plasmids. This degradation process is modeled as killing with a time-independent rate $k(\boldsymbol{x})$. We use the overdamped Langevin dynamics with a degradation rate to describe the viral and DNA motion. We first recall the Fokker-Planck-type equations [16–18] and run Brownian simulations to compare with the asymptotic approximations of p_N and τ_N derived analytically [16,18] and use these re-
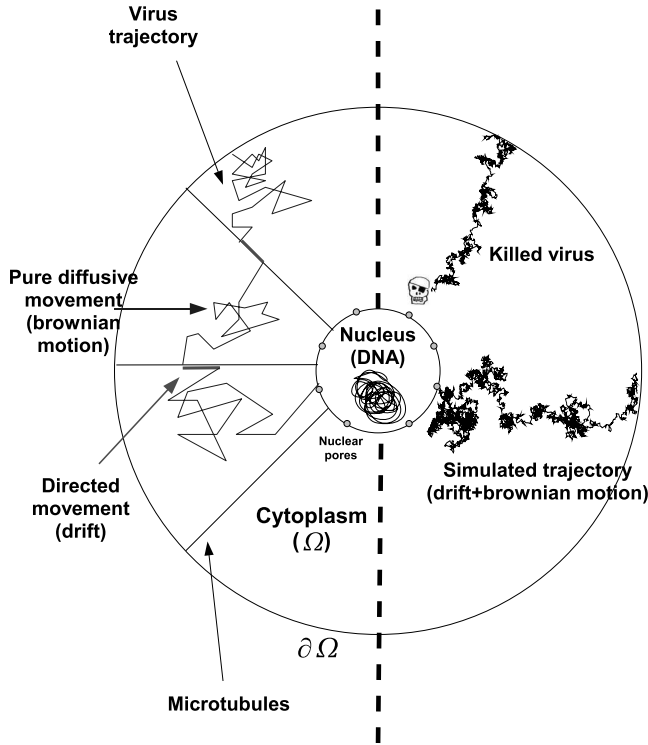
FIG. 1. Schematic representation of the viral trajectory approximation: on the left side of the idealized cell, a real trajectory consists of intermittent Brownian and drift epochs, whereas on the right side, we show two simulated trajectories obtained by Eq. (3). In one of them, the viral particle arrives alive to a nuclear pore, while in the other, it is killed inside the cytoplasm. The round dots on the nucleus surface represent nuclear pores.

sults to estimate the range of validity of our analytical formula. We further compare our numerical simulations and the new analytical formula for the distribution of killed viral particles. The second part of the paper is dedicated to study for many independent viral or DNA particles, the mean time for the first particle to reach a nuclear pore. This mean time is much faster than the time for a single particle to reach a nuclear pore and we obtain here an analytical expression which we then compare to Brownian simulations. In the last part, using an asymptotic analysis, we obtain estimates for p_N and τ_N in the large k limit. The present approach is an attempt to develop a theoretical tool for the analysis of virus and DNA particle dynamics at the single molecule level and, hopefully, for the study of trafficking of synthetic vectors, a necessary step toward gene delivery.

II. MODELING INTRACELLULAR VIRAL AND DNA TRAFFICKING

Modeling DNA carriers trajectories. We model viral trajectories as a collection of pieces, each of which is characterized either as directed movement along microtubules or pure Brownian motion [1–3]. In contrast, DNA motion in the cytoplasm can be adequately described as pure Brownian motion [6]. Particles moving inside the cell are reflected at impermeable surfaces and are absorbed at nuclear pores. A

virus travels on microtubules as long as it binds to a motor. The three- or two-dimensional position of a particle $\mathbf{X}(t)$ is described by the coarse-grained stochastic dynamics

$$\dot{\mathbf{X}} = \begin{cases} \sqrt{2D}\dot{\mathbf{w}} & \text{for a free particle,} \\ \mathbf{V}(t) & \text{for a bound particle,} \end{cases} \tag{1}$$

where $\mathbf{w}$ is a δ-correlated standard white noise and $\mathbf{V}(t)$ is a time-dependent velocity along a microtubule. The velocity $\mathbf{V}(t)$ can be either positive or negative, depending on whether a viral particle binds to a dynein or to a kinesin motor. However, it is not clear what regulatory mechanisms is involved in such a choice [19].

Mathematical description of a viral trajectory in the cytoplasm. We consider the trafficking of a viral particle from an endosome or the cell membrane to a small nuclear pore. The cell cytosol is a bounded spatial domain Ω, whose boundary $\partial\Omega$ is the external membrane $\partial\Omega_{\text{ext}}$ and the nuclear envelope (Fig. 1). Most of the nuclear membrane consists of a reflecting boundary ∂N_r, except for small nuclear pores ∂N_a, where a viral particle can enter the nucleus. We assume that a viral particle that reaches a pore is instantly absorbed, so that this boundary is purely absorbing for trajectories. The ratio of the surface areas is assumed small,

$$\varepsilon = \frac{|\partial N_a|}{|\partial\Omega|} \ll 1. \tag{2}$$

Homogenization of viral trajectory. To replace the intermittent dynamics between free diffusion and the drift motion along microtubules, described in Eq. (1), we use a calibration procedure described in Refs. [20,21]. In this homogenization procedure, the motion is described by the overdamped limit of the Langevin equation

$$dX = \boldsymbol{b}(X)dt + \sqrt{2D}d\mathbf{W}, \tag{3}$$

where D is the diffusion constant and $\boldsymbol{b}(X)$ represents the steady state drift that account for the microtubules density, the forward and backward binding rate and the velocity along the microtubules [20,21]. Because most of the microtubules starting from the cell surface converge to the centrosome, a specialized organelle located nearby the cell nucleus (Fig. 1), we choose in a first approximation a radially symmetric effective drift $\boldsymbol{b}(X)$ converging to the nucleus surface. We thus neglected the minor contribution of microtubules that are not oriented along the radial direction. This radial geometry approximation is actually common in biophysical modelings of *in vitro* experiments [8,22]. Thus, although viruses move bidirectionally on microtubules, the overall movement is directed toward the nucleus, and we only consider here this average component [19]. The drift component (3) can be written as

$$\boldsymbol{b}(\mathbf{X}) = -b(r)\frac{X}{|X|}, \tag{4}$$

with $r = |X|$ is the radial distance to the cell center. In first approximation, we approximate $b(r)$ as a constant $b(r) = B$, which depends on many parameters, such as the density of microtubules, the binding and unbinding rates and the averaged velocity of the directed motion along microtubules

[20,21]. Because the microtubule density increases near the nucleus, a radial dependent drift is more accurate [21], but we already show [20] that the constant approximation is good enough and it leads to more concise analytical expressions.

From trajectory description to the probability and mean arrival time. Viral killing or immobilization and naked DNA degradation by nucleases, are coarse grained into a steady state degradation or killing rate $k(\mathbf{X})$. We briefly recall (see Ref. [18] for details) how to derive the asymptotic expressions for the probability P_N, that a DNA carrier (single virus or DNA) arrives to a small nuclear pore alive and for the mean first passage time (MFPT) τ_N, using approximation (2). The asymptotic estimates depend on the diffusion constant D, the amplitude of the drift B, and k. These computations are based on the small hole theory [14], which describes a Brownian particle confined to a bounded domain by a reflecting boundary, except for a small absorbing window, through which it escapes. The domain Ω contains a spherical nucleus of radius δ (a disk in the two-dimensional case). The survival probability density function (SPDF) $p(\mathbf{x},t)$ to find the virus or naked DNA alive inside the volume element $\mathbf{x}+d\mathbf{x}$ at time t is given by [18]

$$p(\mathbf{x},t)d\mathbf{x} = \Pr\{X(t) \in \mathbf{x} + d\mathbf{x}, \tau^k > t, \tau^a > t | p_i\}, \quad (5)$$

where τ^a is the first passage time of a live DNA carrier to the absorbing boundary ∂N_a, τ^k is the first time it is hydrolyzed or immobilized, and p_i is the initial distribution. The SPDF $p(\mathbf{x},t)$ of the motion (3) is the solution of the mixed initial boundary value problem for the Fokker-Planck equation (FPE) [16]

$$\frac{\partial p}{\partial t}(\mathbf{x},t) = D\Delta p(\mathbf{x},t) - \nabla \cdot \mathbf{b}(\mathbf{x})p(\mathbf{x},t) - kp(\mathbf{x},t),$$

$$p(\mathbf{x},0) = p_i(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega \quad (6)$$

with the boundary conditions

$$p(\mathbf{x},t) = 0 \quad \text{for } \mathbf{x} \in \partial N_a$$

$$\mathbf{J}(\mathbf{x},t) \cdot \mathbf{n}_x = 0 \quad \mathbf{x} \in \partial N_r \cup \partial\Omega_{\text{ext}}, \quad (7)$$

where $\mathbf{n}_x$ is the unit outer normal at a boundary point $\mathbf{x}$. The flux density vector $\mathbf{J}(\mathbf{x},t)$ is defined as

$$\mathbf{J}(\mathbf{x},t) = -D\nabla p(\mathbf{x},t) + \mathbf{b}(\mathbf{x})p(\mathbf{x},t). \quad (8)$$

The survival probability P_N that a live DNA carrier arrives at the nucleus is $P_N = \Pr\{\tau^a < \tau^k\}$ [17]. This probability can be expressed in terms of the SPDF [17] by

$$P_N = 1 - \Pr\{\tau^a > \tau^k\} = 1 - \int_\Omega k(\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x}, \quad (9)$$

where $\tilde{p}(\mathbf{x}) = \int_0^\infty p(\mathbf{x},t)dt$ is the solution of equation

$$D\Delta\tilde{p}(\mathbf{x}) - \nabla \cdot \mathbf{b}(\mathbf{x})\tilde{p}(\mathbf{x}) - k(\mathbf{x})\tilde{p}(\mathbf{x}) = -p_i(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega$$

with the boundary conditions (7). Using the probability density function (PDF) of the time to absorption, conditioned on the event that the DNA carrier escapes alive

TABLE I. The theoretical mean time and probability for a plasmid to reach a small nuclear pore are tested against Brownian simulations. The parameters are $R = 20\mu$ m, $\delta = 20\mu$ m, $\eta = \frac{\pi}{12}\delta = 1.05$ μm, $D = 0.02$ μm^2 s^{-1}, $n = 1$.

Time and probability	τ_N	P_N
Theoretical values	3567 s	0.90%
Simulated values (2000 particles)	3564 s	0.97%

$\Pr\{\tau^a < t | \tau^a < \tau^k\}$, we define the conditional mean time to absorption as

$$\tau_N = E[\tau^a | \tau^a < \tau^k] = \int_0^\infty (1 - \Pr\{\tau^a < t | \tau^a < \tau^k\})dt.$$

Following the computations of Ref. [18], we get

$$\tau_N = \frac{\int_\Omega \tilde{p}(\mathbf{x})d\mathbf{x} - \int_\Omega k(\mathbf{x})q(\mathbf{x})d\mathbf{x}}{1 - \int_\Omega k(\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x}}, \quad (10)$$

where

$$q(\mathbf{x}) = \int_0^\infty s\tilde{p}(\mathbf{x},s)ds \quad (11)$$

satisfies [18]

$$-\tilde{p} = D\Delta q(\mathbf{x}) - [\nabla \cdot \mathbf{b}q] - kq \quad \text{for } \mathbf{x} \in \Omega \quad (12)$$
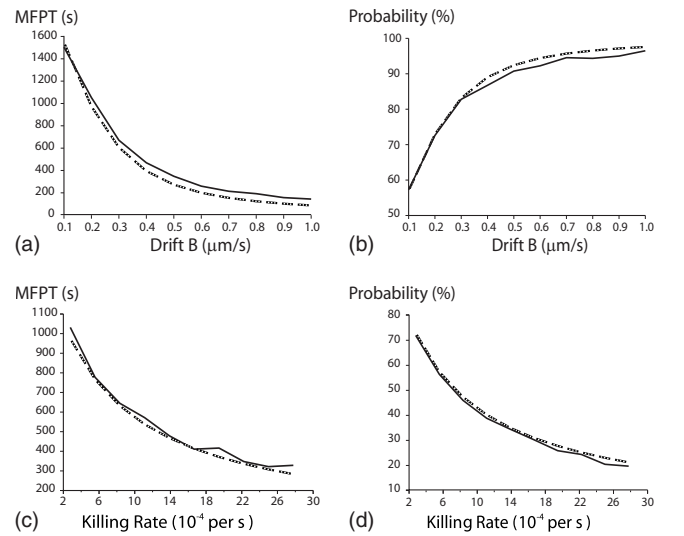
with boundary conditions (7).

FIG. 2. MFPT (top left) and the arrival probability (top right) for increasing values of the drift ($k = \frac{1}{3600}$ s^{-1}) and for increasing values of the steady state degradation rate ($B = 0.2$ μms^{-1}) (bottom). 2000 random trajectories are simulated, theoretical and simulated graphs are, respectively, drawn with dashed and solid lines. The parameters are $R = 20$ μm, $\delta = 4$ μm, $\eta = \frac{\pi}{12}\delta = 1.05$ μm, $D = 1.3$ μm^2 s^{-1} [3], and $n = 1$.

Comparison of the Brownian simulations with the asymptotic analytical formula: The plasmid case. The two extreme cases where the previous equations can be developed into analytical formula are a high and small degradation rate compared to the cytoplasm exploring rate defined as $\frac{D}{|\Omega|}$, with $|\Omega|$ the volume of cell cytoplasm. For small k, we obtained in Ref. [18] explicit expression for P_N and τ_N, for a nucleus containing n well separated small holes (nuclear pores) on its surface. In a three dimensional cell, the asymptotic analysis for naked DNA ($\boldsymbol{b}=\boldsymbol{0}$) leads to

$$P_N = \frac{1}{1 + \dfrac{|\Omega|\tilde{k}}{4nD\eta}} \quad \text{and} \quad \tau_N = \frac{\left(\dfrac{|\Omega|}{4D\eta n}\right)}{1 + \left(\dfrac{|\Omega|\tilde{k}}{4nD\eta}\right)}, \qquad (13)$$

where $\tilde{k}=\frac{1}{|\Omega|}\int_\Omega k(\boldsymbol{x})d\boldsymbol{x}$ and η is the radius of a small absorbing disk (a nuclear pore). Formula (13) does not depend on the specific shape of the degradation rate k, but rather on its integral.

We compare here this asymptotic formula with pure Brownian simulations (no drift) for Eq. (3) with the parameters $R=20$ μm, $\delta=\frac{R}{5}$, $\eta=\delta\frac{\pi}{12}=1.05$ μm, $k=\frac{1}{3600}$ s^{-1} [23], $D=0.02$ μm^2 s^{-1} [6], and $n=1$ (a single big hole), which corresponds to a cell with 2% of the nuclear surface occupied by a large nuclear pore (the $n=2000$ pores of radius 25 nm [24] observed experimentally occupy exactly 2% of the nuclear membrane). Numerical simulations using an effective big hole actually leads to an over estimation of the mean time compared with many holes: Formula (13) is only valid for few well-separated holes and a separate study should reveal the true formula for the narrow escape time with many

holes (the last section of Ref. [25] has started to address this issue). Finally, the small diffusion constant $D=0.02$ μm^2 s^{-1} accounts for electrostatics binding and entanglements that slow down processing of plasmids. The results are summarized in Table I, where we observe a nice agreement between the analytical formula and our Brownian simulations.

Comparison of the Brownian simulations with the asymptotic analytical formula: The virus case. For a virus trajectory governed by Eq. (3) with a potential drift $\boldsymbol{b}=-\nabla\Phi$, when the degradation rate is small compared to the diffusion rate, the leading order term of the probability and the mean time are given by [18]

$$P_N = \frac{\dfrac{1}{|\partial\Omega|}\displaystyle\int_{\partial\Omega} e^{-\Phi(\boldsymbol{x})/D}dS_{\boldsymbol{x}}}{\dfrac{1}{|\partial\Omega|}\displaystyle\int_{\partial\Omega} e^{-\Phi(\boldsymbol{x})/D}dS_{\boldsymbol{x}} + \dfrac{1}{4nD\eta}\displaystyle\int_\Omega e^{-\Phi(\boldsymbol{x})/D}k(\boldsymbol{x})d\boldsymbol{x}},$$

$$(14)$$

$$\tau_N = \frac{\dfrac{1}{4nD\eta}\displaystyle\int_\Omega e^{-\Phi(\boldsymbol{x})/D}d\boldsymbol{x}}{\dfrac{1}{|\partial\Omega|}\displaystyle\int_{\partial\Omega} e^{-\Phi(\boldsymbol{x})/D}dS_{\boldsymbol{x}} + \dfrac{1}{4nD\eta}\displaystyle\int_\Omega e^{-\Phi(\boldsymbol{x})/D}k(\boldsymbol{x})d\boldsymbol{x}},$$

$$(15)$$

where $S_{\boldsymbol{x}}$ is the surface element corresponding to the boundary position $\boldsymbol{x}$. For a scalar drift B and degradation rate k that are both constant, in an idealized spherical cell (radius R), a direct estimation of Eqs. (14) and (15) gives

$$P_N = \frac{e^{-B\delta/D}}{\dfrac{\pi}{nD\eta}\left\{e^{-B\delta/D}\left[\dfrac{D}{B}\delta^2 + 2\left(\dfrac{D}{B}\right)^2\delta + 2\left(\dfrac{D}{B}\right)^3\right] - e^{-BR/D}\left[\dfrac{D}{B}R^2 + 2\left(\dfrac{D}{B}\right)^2 R + 2\left(\dfrac{D}{B}\right)^3\right]\right\}k + e^{-B\delta/D}}, \qquad (16)$$

$$\tau_N = \frac{\dfrac{\pi}{nD\eta}\left\{e^{-B\delta/D}\left[\dfrac{D}{B}\delta^2 + 2\left(\dfrac{D}{B}\right)^2\delta + 2\left(\dfrac{D}{B}\right)^3\right] - e^{-BR/D}\left[\dfrac{D}{B}R^2 + 2\left(\dfrac{D}{B}\right)^2 R + 2\left(\dfrac{D}{B}\right)^3\right]\right\}}{\dfrac{\pi}{nD\eta}\left\{e^{-B\delta/D}\left[\dfrac{D}{B}\delta^2 + 2\left(\dfrac{D}{B}\right)^2\delta + 2\left(\dfrac{D}{B}\right)^3\right] - e^{-BR/D}\left[\dfrac{D}{B}R^2 + 2\left(\dfrac{D}{B}\right)^2 R + 2\left(\dfrac{D}{B}\right)^3\right]\right\}k + e^{-B\delta/D}}. \qquad (17)$$

Contrary to the formula given in Ref. [18], to match the Brownian simulations, we have kept in those expression the dependency in R. In Fig. 2, we compare Eqs. (16) and (17) with Brownian simulations for several values of the drift and a constant degradation rate. Furthermore, Eqs. (16) and (17) show that the main contribution to the probability and the mean time comes from a boundary layer located near the nucleus surface.

To see the efficiency of formulas (16) and (17), we can now predict the effect of changing the effective drift $B=0.2$ by $\pm 30\%$. We recall that value of the drift come from the following rational: for a large number of microtubules, the drift B equals the apparent velocity (which is about 10% [27] of the minus end velocity, approximatively equal to 2 μm/s [3]). We found that increasing the drift leads to a probability $P_N^{+30\%}=0.80$ and a mean time $\tau_N^{+30\%}=731$ s, while reducing the drift gives $P_N^{-30\%}=0.64$ and $\tau_N^{-30\%}=1293$ s.

We conclude that decreasing the drift amplitude by 30% increases the time by 33% ($\tau_N = 974$ s) and decreases the probability by 12% ($P_N = 0.73$), while increasing the drift by 30%, reduces the time by 22% and increases the probability by 10%. These results show the nonlinear effect of the drift. In a biological context, decreasing the drift can be implemented by disrupting the microtubule network. Moreover, using formula (13) with the viral parameters given above, we obtain for zero drift ($B = 0$), a mean arrival time equal to $\tau_N = 2262$ s. We conclude that the drift due to active transport along microtubules decreases τ_N of a virus to a nuclear pore by a factor 2.5.

III. DISTRIBUTION OF DEGRADED DNA CARRIERS

Gene delivery using viral vectors, such as AAV, needs the most efficient virus *i.e.* the one reaching the nucleus alive with the highest efficiency. Following endocytosis, viruses can be destroyed either in lysosomes or in the cytoplasm (somewhere between its endosomal release and nuclear pore

binding). The distribution of killed viruses can give insights on the cytoplasm degradation activity. For a given steady state degradation rate $k(\mathbf{x})$, the probability $p_k(\mathbf{x})d\mathbf{x}$ that a DNA carrier is degraded in the ball $B(\mathbf{x}, d\mathbf{x})$ of center $\mathbf{x}$ and radius $d\mathbf{x}$ is given by

$$p_k(\mathbf{x})d\mathbf{x} = \tilde{p}(\mathbf{x})k(\mathbf{x})d\mathbf{x}. \tag{18}$$

where for sufficiently small nuclear pores and degradation rate, the leading order term of $\tilde{p}(\mathbf{x})$ is given by [18]

$$\tilde{p}(\mathbf{x}) \approx \frac{\dfrac{e^{-\Phi(\mathbf{x})/D}}{4Dn\eta}}{\dfrac{1}{|\partial\Omega|}\displaystyle\int_{\partial\Omega} e^{-\Phi(\mathbf{x})/D}dS_x + \dfrac{\displaystyle\int_{\Omega} e^{-\Phi(\mathbf{x})/D}k(\mathbf{x})d\mathbf{x}}{4Dn\eta}}. \tag{19}$$

In a spherical geometry with a constant degradation rate k and a constant radial drift $\mathbf{b}(\mathbf{x}) = -B\frac{\mathbf{r}}{|\mathbf{r}|} \neq 0$ [i.e., for a potential $\Phi(r) = Br$], we get

$$p_k(r) = \frac{ke^{-Br/D}}{4Dn\eta e^{-B\delta/D} + 4\pi k\left[e^{-B\delta/D}\left(\delta^2\dfrac{D}{B} + 2\delta\dfrac{D^2}{B^2} + \dfrac{D^3}{B^3}\right) - e^{-BR/D}\left(R^2\dfrac{D}{B} + 2R\dfrac{D^2}{B^2} + \dfrac{D^3}{B^3}\right)\right]}. \tag{20}$$

For $n\eta \ll 1$, we obtain

$$p_k(r) = \frac{e^{-Br/D}}{4\pi\left[e^{-B\delta/D}\left(\delta^2\dfrac{D}{B} + 2\delta\dfrac{D^2}{B^2} + \dfrac{D^3}{B^3}\right) - e^{-BR/D}\left(R^2\dfrac{D}{B} + 2R\dfrac{D^2}{B^2} + \dfrac{D^3}{B^3}\right)\right]}. \tag{21}$$

In Fig. 3, we compare the theoretical distribution (21) with the killed viruses distribution obtained with Brownian simulations (in spherical geometry). The simulations and the analytical formula agree nicely and the maximum of the degra-
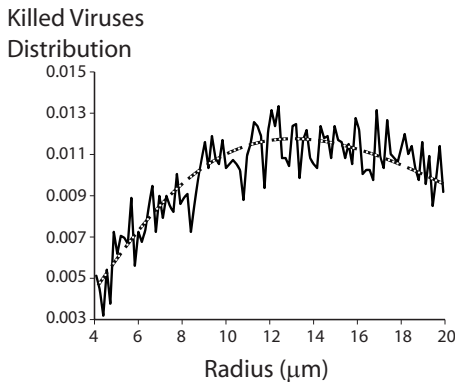

Killed Viruses
Distribution

FIG. 3. Distribution of killed virus $4\pi r^2 p_k(r)$. The simulation is obtained for 30 000 Brownian trajectories (solid line). Parameters: $R = 20$ μm, $\delta = 4$ μm, $\eta = \frac{\pi}{12}\delta = 1.05$ μm, $D = 1.3$ μm^2 s^{-1}, $n = 1$, and $B = 0.2$ μm s^{-1}.

dation density probability [equal to $p_k(r)4\pi r^2 dr$] is obtained by a direct computation using formula (21). We found that it is achieved for a radius $r = 2\frac{D}{B} = 13$ μm.

IV. IMPACT OF THE DEGRADATION DENSITY DISTRIBUTION

In the plasmid case, because P_N depends only on the integral of $k(\mathbf{x})$ over the cytoplasmic domain, the degradation distribution does not impact the arrival probability. However, because viral particles spend most of their time in the nuclear neighborhood, a large concentration of killing factors such as proteasomes in that area could substantially decrease the arrival probability P_N. To study the impact of the degradation distribution, we compare the virus arrival probability $P_N = 73\%$ obtained with a constant degradation rate k with the one obtained with an exponentially distributed in a nuclear neighborhood. We chose $k(r) = \alpha e^{-\lambda r}$ where $\alpha = \frac{k|\Omega|}{\int_{\delta}^{R} e^{-\lambda r}4\pi r^2 dr}$ is a normalization factor and λ a constant. A direct computation gives $\alpha = \frac{k|\Omega|}{h(\lambda)}$ with $h(\lambda) = 4\pi[e^{-\lambda\delta}(\frac{\delta^2}{\lambda} + \frac{2\delta}{\lambda^2} + \frac{2}{\lambda^3}) - e^{-\lambda R}(\frac{R^2}{\lambda} + \frac{2R}{\lambda^2} + \frac{2}{\lambda^3})]$. In that case, we obtain
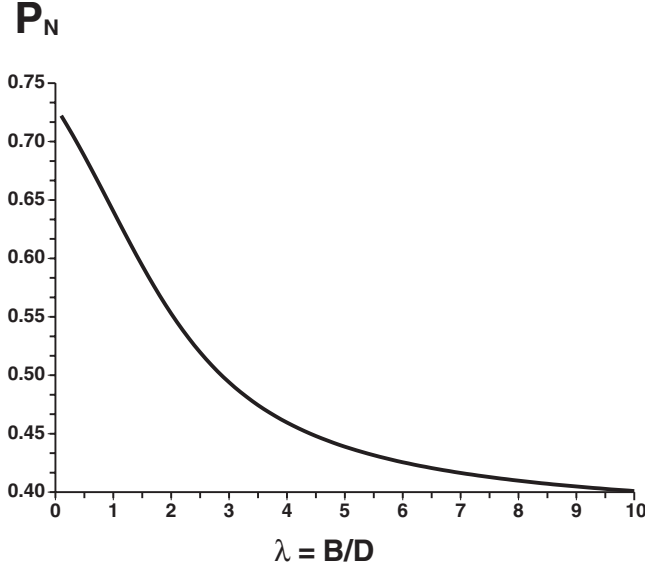
P_N

FIG. 4. The arrival probability P_N is plotted as a function of the characteristic length $\lambda = \frac{B}{D}$. We choose an exponential distribution for the degradation rate $k(r) = \alpha e^{-\lambda r}$, concentrated in the neighborhood of the nucleus where viruses accumulate.

$$P_N = \frac{e^{-B\delta/D}}{\dfrac{k|\Omega|}{4nD\eta} \dfrac{h\left(\lambda + \dfrac{B}{D}\right)}{h(\lambda)} + e^{-B\delta/D}}. \quad (22)$$

In Fig. 4, we plotted P_N as a function of λ: When degradation factors and virions colocalize, which happens for $\lambda \approx \frac{B}{D}$, we obtain that $P_N = 64\%$, which gives a 9% decay compared to the constant killing field case ($P_N = 73\%$). We conclude that the degradation factor distribution does not impact drastically the virions arrival probability.

V. MEAN FIRST PASSAGE TIME OF THE FIRST DNA CARRIER TO A NUCLEAR PORE

Hereafter, we compute the conditioned MFPT $\tau_{\text{first}}(M)$ for the first DNA carrier to attain a nuclear pore. The M-DNA carriers trajectories are independent and we shall use the conditioned MFPT τ_N^j of the jth carrier to a nuclear pore. As in Ref. [18], we consider the absorbing time $\tau_{\text{first}}^a(M)$ of the first DNA carrier to the absorbing boundary ∂N_a and the first time $\tau_{\text{first}}^k(M)$ it is degraded. The probability the first DNA carrier arrives to the absorbing boundary before time t conditioned on not been killed is then given by

$$P(t) = \Pr\{\tau_{\text{first}}^a(M) < t \,|\, \tau_{\text{first}}^a(M) < \tau_{\text{first}}^k(M), p_i\}. \quad (23)$$

The conditional MFPT $\tau_{\text{first}}(M)$ is defined by

$$\tau_{\text{first}}(M) = \int_0^\infty t \frac{dP(t)}{dt} dt = \int_0^\infty [P(\infty) - P(t)] dt. \quad (24)$$

To derive an expression for $\tau_{\text{first}}(M)$, we shall compute $P(t)$ by using Bayes law

$$P(t) = \frac{\Pr\{\tau_{\text{first}}^a(M) < t, \tau_{\text{first}}^a(M) < \tau_{\text{first}}^k(M), p_i\}}{\Pr\{\tau_{\text{first}}^a(M) < \tau_{\text{first}}^k(M), p_i\}}. \quad (25)$$

To estimate the numerator $N(t) = \Pr\{\tau_{\text{first}}^a(M) < t, \tau_{\text{first}}^a(M) < \tau_{\text{first}}^k(M), p_i\}$, we use that

$$\Pr\{\tau_{\text{first}}^a(M) < t, \tau_{\text{first}}^a(M) < \tau_{\text{first}}^k(M), p_i\}$$
$$= 1 - \Pr\{\tau_{\text{first}}^a(M) > t \text{ or } \tau_{\text{first}}^a(M) > \tau_{\text{first}}^k(M), p_i\}. \quad (26)$$

The event $\{\tau_{\text{first}}^a(M) > t \text{ or } \tau_{\text{first}}^a(M) > \tau_{\text{first}}^k(M)\}$ means that, at time t, none of the M-DNA carriers have reached alive a small nuclear pore. Since the particles are independent, we obtain

$$\Pr\{\tau_{\text{first}}^a(M) > t \text{ or } \tau_{\text{first}}^a(M) > \tau_{\text{first}}^k(M), p_i\}$$
$$= \prod_{j=1}^{j=M} (1 - \Pr\{\tau_j^a < t, \tau_j^a < \tau_j^k, p_i\}), \quad (27)$$

where τ_j^a (τ_j^k) is the first time the jth particle is absorbed (killed). Using the interpretation of the flux [18], we get that for any of the particles

$$\Pr\{\tau_j^a < t, \tau_j^a < \tau_j^k, p_i\} = \int_0^t \oint_{\partial\Omega} \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n_x} dS_{\mathbf{x}} = \int_0^t J(s) ds, \quad (28)$$

where $\mathbf{n_x}$ denotes the normal derivative at the boundary point $\mathbf{x}$ and the flux is defined in Eq. (8). Finally, we obtain the following expression for the numerator:

$$N(t) = \Pr\{\tau_{\text{first}}^a(M) < t, \tau_{\text{first}}^a(M) < \tau_{\text{first}}^k(M), p_i\}$$
$$= 1 - \left(1 - \int_0^t J(s) ds\right)^M. \quad (29)$$

Similarly the denominator $D(t)$ of $P(t)$ is given by

$$D(t) = \Pr\{\tau_{\text{first}}^a(M) < \tau_{\text{first}}^k(M), p_i\}$$
$$= 1 - \Pr\{\tau_{\text{first}}^a(M) > \tau_{\text{first}}^k(M), p_i\}, \quad (30)$$

and because the particles are independent

$$D(t) = 1 - \prod_{j=1}^{j=M} \Pr\{\tau_j^a > \tau_j^k, p_i\}. \quad (31)$$

Using the definition of the probability P_N that a particle is killed before reaching the nucleus [18], we get

$$D(t) = 1 - (1 - P_N)^M. \quad (32)$$

Finally, the probability density function is given by

$$P(t) = \frac{N(t)}{D(t)} = \frac{1 - [1 - \int_0^t J(s) ds]^M}{1 - (1 - P_N)^M} \quad (33)$$

and the conditional MFPT $\tau_{\text{first}}(M)$ of the first particle is equal to Eq. (24):

$$\tau_{\text{first}}(M) = \int_0^\infty \frac{[1 - \int_0^t J(s)ds]^M - [1 - \int_0^\infty J(s)ds]^M}{1 - (1 - P_N)^M} dt.$$

(34)

Hereafter, we shall estimate the leading order term for $\tau_{\text{first}}(M)$. In the long time asymptotic, we approximate the PDF by its first exponential term. The leading order term of $p(\mathbf{x}, t)$ is given by

$$p(\mathbf{x}, t) \approx p(\mathbf{x}, 0)e^{-\lambda_0 t}, \quad \text{with} \int_\Omega p(\mathbf{x}, 0)dx = 1, \quad (35)$$

where $\lambda_0 = \frac{1}{\tau_N}$ (Ref. [16], p. 175), is the first eigenvalue (this implies that there is no contribution of the initial condition on the other eigenfunctions, see also Ref. [13]). Replacing $p(\mathbf{x}, t)$ by its long time approximation in Eq. (6), we obtain the following equation for $p(\mathbf{x}, 0)$:

$$-\frac{1}{\tau_N}p = D\Delta p - \nabla[bp] - kp. \quad (36)$$

Using Eq. (28), we obtain an explicit expression for the flux $J(t)$ by integrating Eq. (36) over the domain Ω, with $\int_\Omega p(\mathbf{x}, 0)dx = 1$, we obtain

$$J(t) = \frac{e^{-t/\tau_N}}{\tau_N}\left(\int_\Omega p(\mathbf{x}, 0)d\mathbf{x} - \tau_N \int_\Omega k(\mathbf{x})p(\mathbf{x}, 0)d\mathbf{x} \right)$$

$$= \frac{e^{-t/\tau_N}}{\tau_N}\left(1 - \tau_N \int_\Omega k(\mathbf{x})p(\mathbf{x}, 0)d\mathbf{x} \right). \quad (37)$$

Using the probability P_N (9) and $\tilde{p}(\mathbf{x}) = \int_0^\infty p(\mathbf{x}, t)dt = \frac{\tilde{p}(\mathbf{x})}{\tau_N}$, we get an expression for the flux

$$J(t) = \frac{e^{-t/\tau_N}}{\tau_N}\left(1 - \int_\Omega k(\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x} \right) = \frac{P_N}{\tau_N}e^{-t/\tau_N}. \quad (38)$$

Replacing $\int_0^t J(s)ds$ by its approximation (38) in relation (34) we get

$$\tau_{\text{first}}(M) = \int_0^\infty \frac{[1 - P_N(1 - e^{-t/\tau_N})]^M - (1 - P_N)^M}{1 - (1 - P_N)^M} dt. \quad (39)$$

With the notation $\xi = 1 - P_N$ ($0 \le \xi \le 1$) we have

$$\tau_{\text{first}}(M) = \frac{1}{1 - \xi^M}\int_0^\infty \{[e^{-t/\tau_N} + \xi(1 - e^{-t/\tau_N})]^M - \xi^M\}dt. \quad (40)$$

Thus,

$$\tau_{\text{first}}(M) = \frac{1}{1 - \xi^M}\sum_{k=0}^{M-1}\binom{M}{k}\xi^k\int_0^\infty (1 - e^{-t/\tau_N})^k(e^{-t/\tau_N})^{M-k}dt$$

$$+ \frac{\xi^M}{1 - \xi^M}\int_0^\infty [(1 - e^{-t/\tau_N})^M - 1]dt.$$

An iterative integration by parts yields for $0 \le k \le M-1$

$$\int_0^\infty (1 - e^{-t/\tau_N})^k(e^{-t/\tau_N})^{M-k}dt = \frac{\tau_N}{(M-k)\binom{M}{k}}. \quad (41)$$

Consequently, we have

$$\tau_{\text{first}}(M) = \frac{\tau_N}{1 - \xi^M}\sum_{k=0}^{M-1}\frac{\xi^k}{M-k} + \frac{\xi^M}{1 - \xi^M}\int_0^\infty [(1 - e^{-t/\tau_N})^M$$

$$- 1]dt.$$

Concerning the right-hand side of the equation above, polynomial identity: $X^M - 1 = (X - 1)\Sigma_{k=0}^{M-1}X^k$ leads to

$$\int_0^\infty [(1 - e^{-t/\tau_N})^M - 1]dt = -\sum_{k=0}^{M-1}\int_0^\infty e^{-t/\tau_N}(1 - e^{-t/\tau_N})^k dt. \quad (42)$$

Replacing M by $k+1$ in Eq. (41), we get

$$\int_0^\infty [(1 - e^{-t/\tau_N})^M - 1]dt = -\sum_{k=0}^{M-1}\frac{\tau_N}{k+1} = -\sum_{l=0}^{M-1}\frac{\tau_N}{M-l}. \quad (43)$$

Finally we have the concise expression (note that τ_N is a function of ξ):

$$\tau_{\text{first}}(M) = \frac{\tau_N(\xi)}{1 - \xi^M}\left(\sum_{k=0}^{M-1}(\xi^k - \xi^M)\frac{1}{M-k} \right). \quad (44)$$

We compare in Fig. 5 (left) the analytical curves with the Brownian simulations. Both curves match very nicely, which confirms the validity of the long time asymptotic approximation.

In Fig. 5 (right), we plotted $\tau_{\text{first}}(M)/\tau_N$ as a function of ξ, which is an increasing function of ξ: when the number of DNA carriers reaching alive a nuclear pore decreases, the MFPT of the first survivor increases. Moreover, the curves confirm that for small ξ, the leading order term of $\tau_{\text{first}}(M)$ is

$$\frac{\tau_{\text{first}}(M)}{\tau_N(\xi)} \approx \frac{1}{M}, \quad (45)$$

whereas when ξ tends to 1 (i.e., almost all DNA carriers are killed before reaching nuclear pores) we get the approximation

$$\tau_{\text{first}}(M) \approx \frac{\tau_N}{M(1 - \xi)}\left(\sum_{k=0}^{M-1}(M-k)(1 - \xi)\frac{1}{M-k} \right) = \tau_N(1). \quad (46)$$

It would be interesting to find the general expression for $\tau_{\text{first}}(M)$ as a function of ξ.

VI. THE LARGE DEGRADATION RATE LIMIT

Because the previous analysis [18] does not give any range of validity of the asymptotic formula for the probability and the mean time to reach a nuclear pore, we decided to
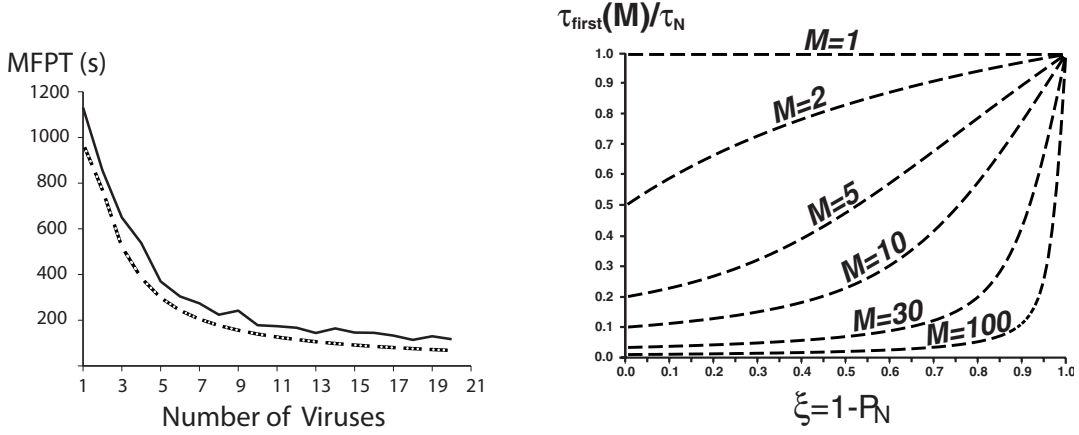
FIG. 5. Left: MFPT of the first virus to a nuclear pore. We generate 300 Brownian trajectories (solid line). Parameters are $R = 20$ μm, $\delta = 4$ μm, $\eta = \frac{\pi}{12}\delta = 1.05$ μm, $D = 1.3$ μm^2 s^{-1}, $n = 1$, and $B = 0.2$ μm s^{-1}. Right: Normalized MFPT of the first virus to the MFPT of a single virus, as a function of the probability $\xi = 1 - P_N$ to be killed before arriving to the nucleus. As ξ tends to 0, τ_{first} tends to $\frac{\tau_N}{M}$ ($\tau_N = 974$ s here); whereas τ_{first} tends to τ_N when almost all DNA carriers are degraded.

investigate more carefully the case where the degradation rate is large $k \gg 1$. We computed in Ref. [18] P_N and τ_N in the limit of a small degradation rate limit $k(\mathbf{x}) \ll 1$, however, in the plasmid case, the killing activity due to the protease could be much larger than the diffusion time scale. Thus, we derive hereafter new asymptotics in the large degradation rate limit. The analysis is quite different from Ref. [18]. We start with a constant degradation rate $k(\mathbf{x}) = k$ (the computations for a general radial degradation rate are given in the Appendix). We consider a uniform initial plasmid distribution over the cytoplasm $p_i(\mathbf{x}) = p_0 = \frac{1}{|\Omega|}$. To compute the probability P_N, we shall solve equation (10)

$$D\Delta\widetilde{p}(\mathbf{x}) - k(\mathbf{x})\widetilde{p}(\mathbf{x}) = -p_0 = -\frac{1}{|\Omega|},\qquad(47)$$

with the boundary conditions (7). When $\frac{D}{|\Omega|}$ is much smaller compared to k and for a particle starting far from nuclear pores, we approximated the solution of Eq. (47) by

$$p_{\text{outer}}(\mathbf{x}) = \frac{1}{k|\Omega|} + O(D).\qquad(48)$$

However, this outer solution does not match the absorbing conditions. We now construct an inner solution $p_{\text{inner}}(\mathbf{x})$ near the nuclear pores that will satisfy the absorbing conditions and match the outer solution. In a local coordinates (ρ, s) near ∂N_a, where ρ measures distance from ∂N_a, measured positively into Ω, and s are tangential variables in the plane $\rho = 0$ (see, for example, Ref. [26] and Fig. 6, where the local coordinate system is represented in a two-dimensional geometry with a single nuclear pore). Projecting Eq. (47) on the ρ coordinate (the variations of $\widetilde{p}$ with respect to s are small compared to the variation in ρ), we obtain for the leading order term p_{inner}:

$$\frac{d^2 p_{\text{inner}}(\rho)}{d\rho^2} - \frac{k}{D} p_{\text{inner}}(\rho) = -\frac{1}{D|\Omega|},\qquad(49)$$

satisfying the absorbing condition on the nuclear pore

$$p_{\text{inner}}(0) = 0.\qquad(50)$$

Far from the boundary layer [26], the matching condition is

$$\lim_{\rho/\sqrt{D}\to\infty} p_{\text{inner}}(\rho) = p_{\text{outer}} = \frac{1}{|\Omega|k}.\qquad(51)$$

Consequently, near the boundary we get

$$p_{\text{inner}}(\rho, s) = \frac{1}{|\Omega|k}(1 - e^{-(\sqrt{k/D})\rho}).\qquad(52)$$

To compute P_N, we use formula (9)

$$P_N = 1 - \int_\Omega k(\mathbf{x})\widetilde{p}(\mathbf{x})d\mathbf{x},\qquad(53)$$

which can be rewritten as

$$P_N = 1 - \left(\int_{\Omega\backslash\text{BL}} k p_{\text{outer}} d\mathbf{x} + \int_{\text{BL}} k p_{\text{inner}}(\rho) d\rho\right),\quad(54)$$

where BL is the boundary layer. Using expression (48) for p_{outer}, we get

$$\int_{\Omega\backslash\text{BL}} k p_{\text{outer}}(\mathbf{x})d\mathbf{x} = \frac{|\Omega \backslash \text{BL}|}{|\Omega|}\qquad(55)$$
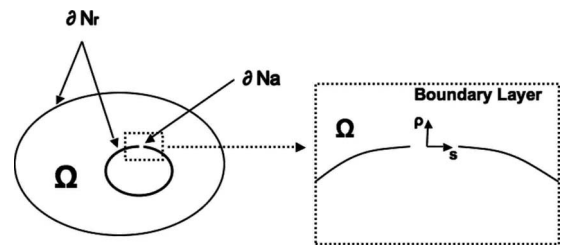
and finally

FIG. 6. Schematic representation of the boundary layer in a local coordinate system (ρ, s) near the boundary ∂N_a, where ρ is the distance from ∂N_a, measured positively, and s is the arc length.
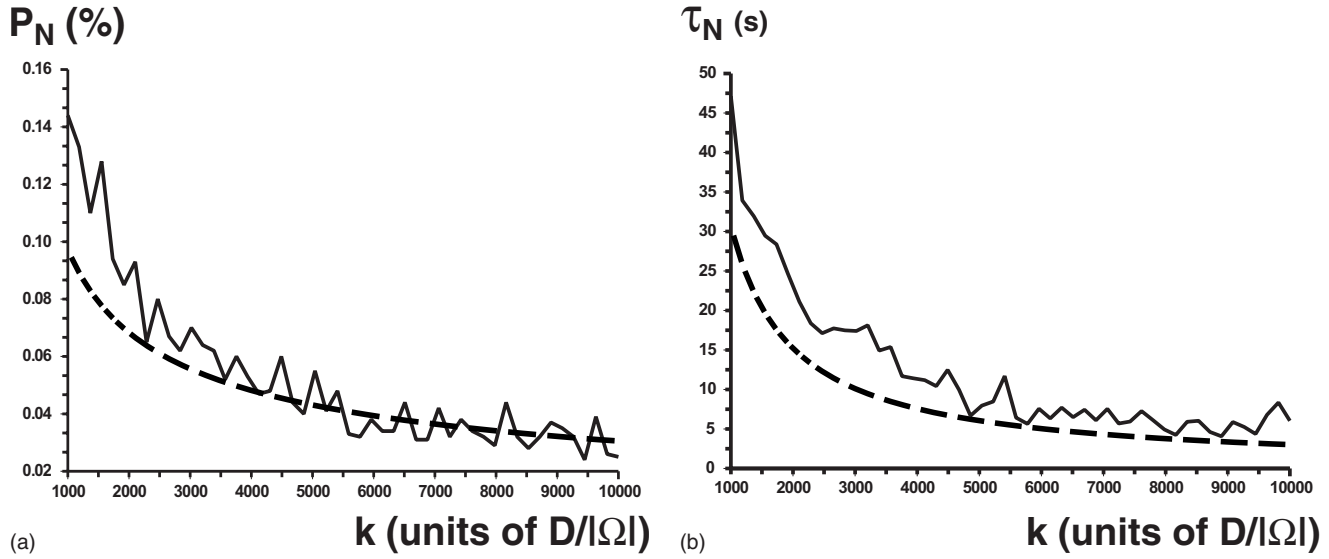
FIG. 7. The probability and mean time for a plasmid to reach a small nuclear pore plotted as a function of the constant degradation rate for a two dimensional flat cell. The Brownian simulations match the analytic solutions (57) and (67) only after a rate of $3000 \frac{D}{|\Omega|} = 3000 \frac{D}{\pi(R^2 - \delta^2)} \approx 0.05$ s^{-1}, around 200 higher than the normal rate $1/3600 \approx 2.8 \times 10^{-4}$ s^{-1}.

$$\int_{BL} k p_{inner}(\rho) d\rho = \frac{1}{|\Omega|} \left(|BL| + |\partial N_a| \int_0^{\rho_0} - e^{-(\sqrt{k/D})\rho} d\rho \right)$$
$$= \frac{1}{|\Omega|} \left(|BL| - |\partial N_a| \sqrt{\frac{D}{k}} (1 - e^{-(\sqrt{k/D})\rho_0}) \right), \tag{56}$$

where $\rho_0 \gg \sqrt{\frac{D}{k}}$ is the thickness of the boundary layer. Finally,

$$P_N = \frac{|\partial N_a|}{|\Omega|} \sqrt{\frac{D}{k}} + O(e^{-(\sqrt{k/D})\rho_0}). \tag{57}$$

In a three-dimensional cell, when the boundary consists of n well separated small holes of radius η, we obtain that

$$P_N = \frac{n\pi\eta^2}{|\Omega|} \sqrt{\frac{D}{k}} + O(e^{-(\sqrt{k/D})\rho_0}). \tag{58}$$

Because our analysis is local, it can be extended to any degradation rate, large compared to the exploring rate. In that case, when for n well separated narrow pores of size η_q, $1 \leq q \leq n$, located at position $x_1, \dots, x_n$, the asymptotic formula is

$$P_N \approx \sum_{q=1}^{n} \frac{\pi \eta_q^2}{|\Omega|} \sqrt{\frac{D}{k(x_q)}} + O(e^{-(\sqrt{k_0/D})\rho_0}), \tag{59}$$

where k_0 is the minimum value of $k(\mathbf{x})$ among the pores. More detailed computations are given in the appendix. From the fitting in Fig. 7 of the Brownian simulations with the analytical formula (57), we conclude that the matching occurs for a very large degradation rate (more than 200 times the normal rate [23]) and thus the large case limit might only be useful to characterize gene delivery for abnormal cells, where the degradation rate is large.

The MFPT τ_N to a small pore for a live virus is [18]

$$\tau_N = \frac{\int_\Omega \tilde{p}(\mathbf{x})d\mathbf{x} - \int_\Omega kq(\mathbf{x})d\mathbf{x}}{P_N}, \tag{60}$$

where $q(\mathbf{x}) = \int_0^\infty s\tilde{p}(\mathbf{x}, s)ds$ satisfies Eq. (12) with boundary conditions (7). To estimate τ_N, we consider for a small diffusion, an outer approximation of q given by

$$q_{outer} = \frac{p_{outer}}{k} = \frac{1}{|\Omega|k^2}. \tag{61}$$

The leading term of the inner solution q_{inner} in the boundary layer expansion of q satisfies

$$\frac{d^2 q_{inner}(\rho)}{d\rho^2} - \frac{k}{D} q_{inner}(\rho) = -\frac{p_{inner}}{D} = -\frac{1}{D|\Omega|k} (1 - e^{-(\rho\sqrt{k/D})}), \tag{62}$$

$$q_{inner}(0) = 0, \tag{63}$$

$$\lim_{\rho/\sqrt{D} \to \infty} q_{inner}(\rho) = q_{outer} = \frac{1}{|\Omega|k^2}. \tag{64}$$

Consequently, we get

$$q_{inner} = \frac{1}{|\Omega|k^2} (1 - e^{-(\sqrt{k/D})\rho}) - \frac{\rho}{2\sqrt{D}k^{3/2}|\Omega|} e^{-(\sqrt{k/D})\rho}. \tag{65}$$

We get then that

$$\int_\Omega \tilde{p}(\mathbf{x})d\mathbf{x} - \int_\Omega kq(\mathbf{x})d\mathbf{x} = |\partial N_a| \frac{\sqrt{D}}{2|\Omega|k^{3/2}} + O(e^{-(\sqrt{k/D})\rho_0}). \tag{66}$$

Finally,

$$\tau_N = |\partial N_a| \frac{\sqrt{D}}{2|\Omega|k^{3/2}P_N} + O(e^{-(\sqrt{k/D})\rho_0}) = \frac{1}{2k} + O(e^{-(\sqrt{k/D})\rho_0}).$$

(67)

For a large degradation rate, our analytical results match the Brownian simulations (see Fig. 7). Moreover, the present local analysis can be extended to any degradation rate and for n well separated narrow pores, located at position $x_1, \ldots, x_n$. We anticipate the following asymptotic formula;

$$\tau_N \approx \frac{1}{n}\sum_{q=1}^{n}\frac{1}{2k(x_q)} + O(e^{-(\sqrt{k_0/D})\rho_0}),$$

(68)

where $k_0 = \min_q k(x_q)$ is the minimum value of $k(\mathbf{x})$ among the pores. k_0 is the minimum concentration of killing factors among nuclear pores.

VII. CONCLUSION

By describing the intermittent dynamics of a DNA carrier inside the cytoplasm with an effective stochastic description (3), we derived a quantitative analysis of the nuclear DNA carrying at the single unit level. By modeling the DNA degradation, as protease activity, that occurs in the cell cytoplasm with a steady state degradation rate $k(\mathbf{x})$, we also derived expressions for the probability a DNA carrier hits a small nuclear pore and the mean time it takes (in both cases of small and large degradation rate). We also provided here the distribution of degraded particles.

When many independent viruses are involved, we computed the mean time to a nuclear pore for the first one. We tested our analytical results against Brownian simulations and we obtained that our curves match nicely. Our analysis provides a tool to explore the multidimensional parameter space of nuclear DNA carrying. Cytoplasmic trafficking is a limiting step of gene delivery and elucidating viral motion in the cytoplasm may provide a quantitative tool for the improvement and optimization of delivery of synthetic vectors.

ACKNOWLEDGMENTS

APPENDIX

We compute hereafter the probability P_N for a carrier moving by random motion to hit a small nuclear pore for a large (compared to the exploring rate) degradation rate $k(\mathbf{x})$. Our method is based on a boundary layer analysis and we generalize it here to a nonconstant degradation rate, similar to the one produced in this manuscript for a constant k: far from the nuclear pore, the leading order term of the outer solution is no longer constant and it is given by

$$p_{\text{outer}}(\mathbf{x}) = \frac{p_0}{k(\mathbf{x})} + O(D).$$

(A1)

The initial uniform distribution of DNA carriers is $p_0 = \frac{1}{|\Omega|}$. To compute the inner solution near the nuclear surface, we ex-

pand the steady state radial killing measure along the radial ρ coordinate,

$$k(\rho, s) = k_0(s) + k_1(s)\rho + O(\rho^2),$$

(A2)

where $k(\rho = 0, s) = k_0(s)$ and $\frac{dk}{d\rho}(\rho = 0, s) = k_1(s)$. Because p_{outer} does not necessary satisfy the reflecting boundary condition anymore, we construct two inner solutions: the first one p_{inner}^1 near ∂N_a and the second p_{inner}^2 near ∂N_r. Projecting Eq. (47) on the ρ coordinate (the variations of $\tilde{p}$ with respect to s are small compared to the variation in ρ), we obtain that the leading order terms of $p_{\text{inner}}^i(\rho, s)$ for $i = 1, 2$ satisfy

$$\frac{\partial^2 p_{\text{inner}}^i}{\partial \rho^2} - \frac{k_0(s) + k_1(s)\rho}{D}p_{\text{inner}}^i = -\frac{1}{|\Omega|D} \quad \text{for } 0 \leq \rho \leq \rho_0(s),$$

$$p_{\text{inner}}^1(\rho = 0, s) = 0 \quad \text{on } \partial N_a,$$

$$\frac{\partial}{\partial \rho}p_{\text{inner}}^2(\rho = 0, s) = 0 \quad \text{on } \partial N_r$$

for $i = 1, 2 \lim_{\rho/\sqrt{D} \to \infty} p_{\text{inner}}^i(\rho, s) = p_{\text{outer}}(\rho = 0, s) = \frac{1}{|\Omega|k_0(s)},$

(A3)

where $\rho_0(s) \gg \sqrt{\frac{D}{k_0(s)}}$ is the local thickness of the boundary layer. To solve the homogeneous equation

$$\frac{\partial^2 p_{\text{inner}}^i}{\partial \rho^2} - \frac{k_0(s) + k_1(s)\rho}{D}p_{\text{inner}}^i = 0,$$

(A4)

we use the change of variable

$$u = u(\rho, s) = \frac{k_0(s) + k_1(s)\rho}{\beta(s)D}, \quad \text{where } \beta(s) = \left(\frac{k_1(s)}{D}\right)^{2/3}.$$

(A5)

By this substitution in Eq. (A4), we get

$$\frac{\partial^2 p_{\text{inner}}^i}{\partial u^2} - up_{\text{inner}}^i = 0$$

(A6)

and the solution is

$$p_{\text{inner}}^i = C_0^i(s)\text{Ai}(u) + C_1^i(s)\text{Bi}(u),$$

(A7)

where $C_0^i(s)$ and $C_1^i(s)$ are real functions of s and Ai and Bi are the Airy functions (Ref. [28], p. 446). In the small diffusion limit $D \ll 1$, $u \geq \frac{k_0(s)}{\beta(s)D} = \frac{k_0(s)}{[k_1(s)]^{2/3}D^{1/3}} \gg 1$. Because either solutions p_{inner}^i are bounded, but $\lim_{u \to +\infty}\text{Bi}(u) = +\infty$, we get that $C_1^i = 0$ and, consequently,

$$p_{\text{inner}}^i = C_0^i(s)\text{Ai}(u).$$

(A8)

To obtain a particular solution $\bar{p}_{\text{inner}}^i$ of Eq. (A3), we write is as

$$\frac{\partial^2 \bar{p}_{\text{inner}}^i}{\partial u^2} - u\bar{p}_{\text{inner}}^i = -\frac{1}{|\Omega|\beta(s)D}.$$

(A9)

Using the Scorer's functions (Ref. [28], p. 448) and because

$$\lim_{u \to +\infty} \text{Hi}(u) = +\infty, \tag{A10}$$

we obtain that

$$\bar{p}_{\text{inner}}^{i} = \frac{\pi}{|\Omega|\beta(s)D} \text{Gi}(u). \tag{A11}$$

Collecting the results, we obtain that

$$p_{\text{inner}}^{i}(u,s) = C_0^i(s)\text{Ai}(u) + \frac{\pi}{|\Omega|\beta(s)D}\text{Gi}(u). \tag{A12}$$

Using the matching boundary conditions, we get

$$p_{\text{inner}}^{1}(\rho = 0,s) = p_{\text{inner}}^{1}\left(u = \frac{k_0(s)}{\beta(s)D}, s\right) = 0,$$

$$\frac{\partial p_{\text{inner}}^{2}}{\partial \rho}(\rho = 0,s) = \frac{k_1(s)}{\beta(s)D}\frac{\partial p_{\text{inner}}^{2}}{\partial u}\left(u = \frac{k_0(s)}{\beta(s)D}, s\right) = 0.$$

Using expression (A12), we obtain the equations

$$C_0^1(s)\text{Ai}\left(\frac{k_0(s)}{\beta(s)D}\right) + \frac{\pi}{|\Omega|\beta(s)D}\text{Gi}\left(\frac{k_0(s)}{\beta(s)D}\right) = 0, \tag{A13}$$

$$C_0^2(s)\text{Ai}'\left(\frac{k_0(s)}{\beta(s)D}\right) + \frac{\pi}{|\Omega|\beta(s)D}\text{Gi}'\left(\frac{k_0(s)}{\beta(s)D}\right) = 0, \tag{A14}$$

and thus

$$C_0^1(s) = -\frac{\pi}{|\Omega|\beta(s)D}\frac{\text{Gi}\left(\frac{k_0(s)}{\beta D}\right)}{\text{Ai}\left(\frac{k_0(s)}{\beta(s)D}\right)}, \tag{A15}$$

$$C_0^2(s) = -\frac{\pi}{|\Omega|\beta(s)D}\frac{\text{Gi}'\left(\frac{k_0(s)}{\beta(s)D}\right)}{\text{Ai}'\left(\frac{k_0(s)}{\beta(s)D}\right)}. \tag{A16}$$

Finally the inner solutions p_{inner}^{i} are given by

$$p_{\text{inner}}^{1}(u,s) = \frac{\pi}{|\Omega|\beta(s)D}\left(\text{Gi}(u) - \frac{\text{Gi}\left(\frac{k_0(s)}{\beta(s)D}\right)}{\text{Ai}\left(\frac{k_0(s)}{\beta(s)D}\right)}\text{Ai}(u)\right), \tag{A17}$$

$$p_{\text{inner}}^{2}(u,s) = \frac{\pi}{|\Omega|\beta(s)D}\left(\text{Gi}(u) - \frac{\text{Gi}'\left(\frac{k_0(s)}{\beta(s)D}\right)}{\text{Ai}'\left(\frac{k_0(s)}{\beta(s)D}\right)}\text{Ai}(u)\right). \tag{A18}$$

With the outer solution, we use the large u asymptotic of Gi(u), Gi$'(u)$, Ai(u), and Ai$'(u)$ (Ref. [28], pp. 448–450):

$$\text{Gi}(u) \approx \frac{1}{\pi u}, \quad \text{Gi}'(u) \approx \frac{7}{96\pi u^2},$$

$$\text{Ai}(u) \approx \frac{e^{-2/3 u^{3/2}}}{2\sqrt{\pi}u^{1/4}}, \quad \text{Ai}'(u) \approx -\frac{u^{1/4}e^{-2/3 u^{3/2}}}{2\sqrt{\pi}}.$$

For $\rho = \rho_0(s)$ and thus $u = \frac{k_0(s)+k_1(s)\rho_0(s)}{\beta(s)D} \gg 1$, using the asymptotic behavior for Ai and a Taylor expansion at order 1 $[k_1(s)\rho_0(s) \ll k_0(s)]$, we get

$$\text{Ai}\left(\frac{k_0(s) + k_1(s)\rho_0(s)}{\beta(s)D}\right) \approx \text{Ai}\left(\frac{k_0(s)}{\beta(s)D}\right)e^{-(\sqrt{k_0(s)/D})\rho_0(s)}. \tag{A19}$$

Consequently, using expressions (A17) and (A19) and the asymptotics above, we get

$$p_{\text{inner}}^{1}\left(\frac{k_0(s) + k_1(s)\rho_0(s)}{\beta(s)D}, s\right) \approx \frac{1}{|\Omega|[k_0(s) + k_1(s)\rho_0(s)]} - \frac{1}{|\Omega|k_0(s)}e^{-(\sqrt{k_0(s)/D})\rho_0(s)}, \tag{A20}$$

which matches well the outer solution (A1):

$$p_{\text{inner}}^{1}[\rho = \rho_0(s),s] = p_{\text{outer}}[\rho = \rho_0(s),s] + O(e^{-(\sqrt{k_0(s)/D})\rho_0(s)}). \tag{A21}$$

Similarly, p_{inner}^{2} also matches very well:

$$p_{\text{inner}}^{2}[\rho = \rho_0(s),s] = p_{\text{outer}}[\rho = \rho_0(s),s] + O((e^{-(\sqrt{k_0(s)/D})\rho_0(s)}). \tag{A22}$$

We will now use the previous asymptotic analysis for the probability density function to estimate the overall probability P_N that a virus hits a small nuclear pore. Using formula (9), we get

$$P_N = 1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x}$$

$$= 1 - \left(\int_{\Omega\backslash \text{BL}} p_{\text{outer}}(\mathbf{x})k(\mathbf{x})d\mathbf{x} + \int_{\text{BL}} p_{\text{inner}}^{i}(\mathbf{x})k(\mathbf{x})d\mathbf{x}\right). \tag{A23}$$

Using the outer solution expression and that $k(u,s) = \beta(s)Du$ [see Eq. (A5)] in the boundary layer, we have

$$P_N = 1 - \int_{\Omega\backslash \text{BL}}\frac{1}{|\Omega|}d\mathbf{x} - \int_{\text{BL}^1}\beta(s)Dup_{\text{inner}}^{1}(u,s)duds$$

$$- \int_{\text{BL}^2}\beta(s)Dup_{\text{inner}}^{2}(u,s)duds, \tag{A24}$$

where BL^1 and BL^2 are the boundary layers at resp. the absorbing and reflecting boundaries ($\text{BL} = \text{BL}^1 \cup \text{BL}^2$). Using expressions (A17) and (A18) for p_{inner}^{i} in (A24), we obtain that

$$P_N = 1 - \frac{|\Omega \setminus \mathrm{BL}|}{|\Omega|}$$

$$- \int_{\mathrm{BL}^1} \frac{\pi u}{|\Omega|} \left(\mathrm{Gi}(u) - \frac{\mathrm{Gi}\left(\dfrac{k_0(s)}{\beta(s)D} \right)}{\mathrm{Ai}\left(\dfrac{k_0(s)}{\beta(s)D} \right)} \mathrm{Ai}(u) \right) du ds$$

$$- \int_{\mathrm{BL}^2} \frac{\pi u}{|\Omega|} \left(\mathrm{Gi}(u) - \frac{\mathrm{Gi}'\left(\dfrac{k_0(s)}{\beta(s)D} \right)}{\mathrm{Ai}'\left(\dfrac{k_0(s)}{\beta(s)D} \right)} \mathrm{Ai}(u) \right) du ds.$$

Equivalently,

$$P_N = 1 - \frac{|\Omega \setminus \mathrm{BL}|}{|\Omega|} + \int_{\mathrm{BL}^1} \frac{\pi}{|\Omega|} \frac{\mathrm{Gi}\left(\dfrac{k_0(s)}{\beta(s)D} \right)}{\mathrm{Ai}\left(\dfrac{k_0(s)}{\beta(s)D} \right)} u \mathrm{Ai}(u) du ds$$

$$+ \int_{\mathrm{BL}^2} \frac{\pi}{|\Omega|} \frac{\mathrm{Gi}'\left(\dfrac{k_0(s)}{\beta(s)D} \right)}{\mathrm{Ai}'\left(\dfrac{k_0(s)}{\beta(s)D} \right)} u \mathrm{Ai}(u) du ds$$

$$- \int_{\mathrm{BL}} \frac{\pi}{|\Omega|} u \mathrm{Gi}(u) du ds.$$

For large u, using the asymptotic expansion for $\mathrm{Gi}(u)$ (A19), we obtain that

$$\int_{\mathrm{BL}} \frac{\pi}{|\Omega|} u \mathrm{Gi}(u) du ds \approx \frac{|\mathrm{BL}|}{|\Omega|}. \tag{A25}$$

Thus,

$$P_N = \int_{\mathrm{BL}^1} \frac{\pi}{|\Omega|} \frac{\mathrm{Gi}\left(\dfrac{k_0(s)}{\beta(s)D} \right)}{\mathrm{Ai}\left(\dfrac{k_0(s)}{\beta(s)D} \right)} u \mathrm{Ai}(u) du ds$$

$$+ \int_{\mathrm{BL}^2} \frac{\pi}{|\Omega|} \frac{\mathrm{Gi}'\left(\dfrac{k_0(s)}{\beta(s)D} \right)}{\mathrm{Ai}'\left(\dfrac{k_0(s)}{\beta(s)D} \right)} u \mathrm{Ai}(u) du ds. \tag{A26}$$

Using expression (A19), we obtain

$$\frac{\dfrac{\mathrm{Gi}\left(\dfrac{k_0}{\beta D} \right)}{\mathrm{Ai}\left(\dfrac{k_0}{\beta D} \right)}}{\dfrac{\mathrm{Gi}'\left(\dfrac{k_0}{\beta D} \right)}{\mathrm{Ai}'\left(\dfrac{k_0}{\beta D} \right)}} = \frac{\mathrm{Gi}\left(\dfrac{k_0}{\beta D} \right) \mathrm{Ai}'\left(\dfrac{k_0}{\beta D} \right)}{\mathrm{Gi}'\left(\dfrac{k_0}{\beta D} \right) \mathrm{Ai}\left(\dfrac{k_0}{\beta D} \right)} = O[(\beta D)^{-3/2}] = O\left(\frac{1}{\sqrt{D}} \right). \tag{A27}$$

In addition, in the small diffusion approximation $D \ll 1$, we have

$$P_N \approx \int_{\mathrm{BL}^1} \frac{\pi}{|\Omega|} \frac{\mathrm{Gi}\left(\dfrac{k_0(s)}{\beta(s)D} \right)}{\mathrm{Ai}\left(\dfrac{k_0(s)}{\beta(s)D} \right)} u \mathrm{Ai}(u) du ds. \tag{A28}$$

Using that $u = \frac{k_0(s) + k_1(s)\rho}{\beta(s)D} > \frac{k_0(s)}{\beta(s)D} \gg 1$ and the asymptotic expansions (A19), we obtain that

$$P_N \approx \frac{1}{|\Omega|} \int_{\mathrm{BL}^1}$$

$$\times \frac{\left(\dfrac{k_0(s) + k_1(s)\rho}{\beta(s)D} \right)^{3/4} e^{-2/3(k_0(s) + k_1(s)\rho/\beta(s)D)^{3/2}}}{\left(\dfrac{k_0(s)}{\beta(s)D} \right)^{3/4} e^{-2/3(k_0(s)/\beta(s)D)^{3/2}}} d\rho ds, \tag{A29}$$

that is,

$$P_N \approx \frac{1}{|\Omega|} \int_{\mathrm{BL}^1} \left(1 + \frac{k_1(s)}{k_0(s)}\rho \right)^{3/4}$$

$$\times e^{-2/3[k_0(s)/\beta(s)D]^{3/2}(\{1 + [k_1(s)/k_0(s)\rho]\}^{3/2} - 1)} d\rho ds. \tag{A30}$$

Because $0 \leq \rho \leq \rho_0(s)$ with $\rho_0(s) = O(\sqrt{\frac{D}{k_0(s)}}) \ll 1$, we use a Taylor expansion to obtain

$$P_N \approx \frac{1}{|\Omega|} \int_{\mathrm{BL}^1} e^{-[k_0(s)/\beta(s)D]^{3/2}[k_1(s)/k_0(s)]\rho} d\rho ds. \tag{A31}$$

Finally, by replacing $\beta(s)$ by its expression (A5),

$$P_N \approx \frac{1}{|\Omega|} \int_{\mathrm{BL}^1} e^{-[\sqrt{k_0(s)/D}]\rho} d\rho \, ds. \qquad (A32)$$

By integrating (A32) over ρ, we have

$$P_N \approx \frac{1}{|\Omega|} \int_{\partial N_a} \sqrt{\frac{D}{k_0(s)}} (1 - e^{-[\sqrt{k_0(s)/D}]\rho_0(s)}) ds. \qquad (A33)$$

For a sufficiently smooth killing field, when ∂N_a consists of n well separated small absorbing nuclear pore located at the points $(x_q)_{1 \leq q \leq n}$ on $\partial \Omega$, we finally obtain

$$P_N \approx \frac{|\partial N_a|}{|\Omega|} \sum_{q=1}^{n} \sqrt{\frac{D}{k(x_q)}} + O(e^{-(\sqrt{k_0/D})\rho_0}) \qquad (A34)$$

with $k_0 = \inf_q k(x_q)$ and $\rho_0 = \inf_{s \in \partial N_a} \rho_0(s)$. In a three-dimensional cell with narrow pores of radius η_q, $1 \leq q \leq n$, we obtain

$$P_N \approx \sum_{q=1}^{n} \frac{\pi \eta_q^2}{|\Omega|} \sqrt{\frac{D}{k(x_q)}} + O(e^{-(\sqrt{k_0/D})\rho_0}). \qquad (A35)$$

[1] U. F. Greber, Cell **124**, 741 (2006).
[2] N. Arhel *et al.*, Nat. Methods **3**, 817 (2006).
[3] G. Seisengerger *et al.*, Science **294**, 1929 (2001).
[4] G. Zuber *et al.*, Adv. Drug Delivery Rev. **52**, 245 (2001).
[5] A. S. Verkman, Trends Biochem. Sci. **27**, 27 (2002).
[6] E. Dauty and A. S. Verkman, J. Biol. Chem. **280**, 7823 (2005).
[7] B. Sodeik, Trends Microbiol. **8**, 465 (2000).
[8] F. Nedelec, T. Surrey, and A. C. Maggs, Phys. Rev. Lett. **86**, 3192 (2001).
[9] A. T. Dinh *et al.*, Biophys. J. **92**, 831 (2007).
[10] A. T. Dinh *et al.*, Biophys. J. **89**, 1574 (2005).
[11] M. J. Ward and J. B. Keller, SIAM J. Appl. Math. **53**, 770 (1993).
[12] I. V. Grigoriev *et al.*, J. Chem. Phys. **116**, 9574 (2002).
[13] Z. Schuss *et al.*, Proc. Natl. Acad. Sci. U.S.A. **104**, 16098 (2007).
[14] D. Holcman and Z. Schuss, J. Stat. Phys. **117**, 975 (2004).
[15] A. Ajdari, Europhys. Lett. **31**, 69 (1995).
[16] Z. Schuss, *Theory and Applications of Stochastic Differential Equations* (John Wiley, New York, 1980).
[17] D. Holcman *et al.*, Phys. Rev. E **72**, 031910 (2005).
[18] D. Holcman, J. Stat. Phys. **127**, 471 (2007).
[19] M. A. Welte, Curr. Biol. **14**, R525 (2004).
[20] T. Lagache and D. Holcman, SIAM J. Appl. Math. **68**, 1146 (2008).
[21] T. Lagache and D. Holcman, Phys. Rev. E **77**, 030901 (2008).
[22] R. Lipowsky, S. Klumpp, and T. M. Nieuwenhuizen, Phys. Rev. Lett. **87**, 108101 (2001).
[23] D. Lechardeur *et al.*, Gene Ther. **6**, 482 (1999).
[24] G. G. Maul *et al.*, J. Cell Biol. **73**, 748 (1977).
[25] D. Holcman and Z. Schuss, J. Phys. A: Math. Theor. **41**, 155001 (2008).
[26] B. J. Matkowsky and Z. Schuss, SIAM J. Appl. Math. **42**, 822 (1982).
[27] M. Suomalainen *et al.*, J. Cell Biol. **144**, 657 (1999).
[28] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).