Minimal framework for optimizing vaccination protocols targeting highly mutable pathogens

Saeed Mahdisoltani [1,2,3,*] Pranav Murugan,[3,4] Arup K. Chakraborty,[1,2,3,5,6,†] and Mehran Kardar [3,‡]

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[4]*Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[5]*Ragon Institute of Massachusetts General Hospital,*
Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts 02139, USA
[6]*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

A persistent public health challenge is identifying immunization schemes that are effective against highly mutable pathogens such as HIV and influenza viruses. To address this, we analyze a simplified model of affinity maturation, the Darwinian evolutionary process B cells undergo during immunization. The vaccination protocol determines the selection forces that steer affinity maturation to generate antibodies. We focus on identifying the optimal selection forces exerted by a generic time-dependent vaccination protocol to maximize the production of broadly neutralizing antibodies (bnAbs) that can protect against a broad spectrum of pathogen strains. The model utilizes a path integral representation and operator approximations within a mean-field limit and provides guiding principles for optimizing time-dependent vaccine-induced selection forces to enhance bnAb generation. We compare our analytical mean-field results with the outcomes of stochastic simulations, and we discuss their similarities and differences.

I. INTRODUCTION

The adaptive immune system has a remarkable ability to detect and combat a virtually unlimited number of previously unseen pathogens [1,2]. Immune responses are mediated by T lymphocytes (T cells) and B lymphocytes (B cells), each equipped with surface receptors known as T cell receptors (TCRs) and B cell receptors (BCRs), respectively. In a healthy human adult, there are approximately 10^{11} T cells and B cells distributed throughout the body [3]. With the staggering number of TCR and BCR sequences ($>10^{14}$) generated via the V(D)J recombination process [4], most T cells and B cells express a unique surface receptor distinct from others. This extensive receptor diversity enables the immune system to mount specific responses against each new pathogen encountered, including those that emerge after an individual's birth.

Upon exposure to a new pathogen or vaccine component (collectively referred to as antigens), B cells and antibodies with high binding affinity for the surface proteins of the pathogen are generated by a process called *affinity maturation* [1,5]. Affinity maturation entails an accelerated Darwinian evolution of naive B cells, and it occurs within specialized structures called germinal centers (GCs), which are transiently formed in secondary lymphoid organs such as lymph nodes

[6]. First, naive B cells undergo activation and become a candidate for entry into GCs if their BCRs can bind sufficiently strongly to specific residues known as *epitopes* on the surface proteins of the antigen [5]. Upon entering the GC, the activated B cells undergo rapid replication while concurrently accumulating somatic mutations into their BCRs at an accelerated rate. The GC B cells then interact with the infecting antigen, which is displayed on follicular dendritic cells (FDCs) located within the GC environment. B cells whose receptors exhibit stronger binding to the displayed antigen have a greater probability of internalizing the antigen and are consequently more likely to undergo *positive selection*. Positive selection involves multiple stages of productive interactions of the B cells with certain T cell types within the GC. Conversely, B cells with low-affinity BCRs are typically eliminated via apoptosis as they are less likely to successfully compete with high-affinity B cells for the limited selection signals [7].

The majority of positively selected GC B cells undergo multiple rounds of replication, somatic mutation, and selection processes. As a result of the competition among GC B cells during the affinity maturation process, an initially naive B cell population with low-affinity BCRs gradually evolves to exhibit strong binding to the antigen. During each round, a small number of the positively selected B cells exit the GCs and differentiate into either plasma cells or memory B cells [5,7]. Plasma cells secrete antibodies, which are soluble forms of the BCR that can bind the antigen's surface proteins with high affinity and neutralize the pathogen's ability to infect host cells. Thus, the affinity maturation process generates

---

*Contact author: saeedmah@mit.edu
†Contact author: arupc@mit.edu
‡Contact author: kardar@mit.edu

an effective immune response by producing antibodies that specifically neutralize a particular pathogen [5].

Memory B cells, on the other hand, play a crucial role in the immune system by retaining a memory of the antigen that activated their parent B cell during the initial exposure. This immunological memory enables a rapid immune response upon reencountering the same antigen in the future [7]. Upon reexposure to a pathogen, existing memory B cells are selected in an affinity-dependent manner and rapidly expanded outside GCs, a process akin to the mechanisms occurring within GCs but with little to no mutation [1]. A significant fraction of these expanded memory B cells differentiate into plasma cells, secreting a surge of antibodies that offer immediate protection. Meanwhile, new GCs are also formed, producing even higher affinity memory B cells and antibodies over more extended timescales, hence bolstering the immune response against the recurring pathogen.

Vaccines leverage immunological memory to confer protection against severe or life-threatening illnesses caused by specific pathogens they target [8,9]. By eliciting memory cells and antibodies, vaccines prime the immune system to mount a rapid and specific response to the targeted pathogen upon subsequent exposures [8]. While vaccines have significantly improved in terms of safety and efficacy over time, developing effective vaccines against highly mutable pathogens remains a continuing challenge. Pathogens such as HIV, influenza, and SARS-CoV-2 mutate rapidly, allowing them to evade the immune responses targeted at specific strains. Surface proteins of highly mutable viruses can exhibit significant variation across different strains, rendering vaccine-induced antibodies that neutralize a specific strain ineffective against other strains of the virus. This variability complicates vaccine development efforts, as vaccines must contend with the ever-evolving nature of these pathogens.

It is well-established that certain residues within the surface proteins of highly mutable viruses remain relatively conserved across mutant strains to preserve a virus's ability to infect host cells [10,11]. Antibodies targeting these conserved regions, or epitopes, hold the potential to neutralize a broad spectrum of virus strains [12–14]. These broadly neutralizing antibodies (bnAbs) can naturally evolve in HIV-infected patients, demonstrating the immune system's ability to elicit such cross-reactive antibodies [14,15]. However, bnAbs are generated only in rare individuals in response to natural infection or conventional vaccination, and their emergence typically occurs over extended periods and in limited quantities, thus offering limited protection [15–17]. Vaccination strategies aimed at eliciting bnAbs have emerged as a promising approach for diseases like HIV and influenza [14,16,18]. These vaccines have the potential to confer protection against variant strains by inducing a robust immune response targeted at conserved epitopes. Despite the advancements, designing universal variant-proof vaccines for mutating pathogens remains a formidable challenge, as the typical immune response to such pathogens is predominantly driven by strain-specific B cells and antibodies.

For modeling purposes, we can conceptualize the affinity maturation process as the evolutionary population dynamics of B cells "species" within GCs, characterized by inherent stochasticity stemming from randomness in various steps including B cell activation, GC entry, replication, antigen internalization, T cell selection, and more. In this framework, vaccine antigens serve as external interventions guiding this stochastic dynamics process towards desired goals. Designing vaccination protocols that confer broad protection against mutant viruses entails selecting appropriate external factors that steer GC processes to increase the likelihood of rare evolutionary trajectories leading to the development of bnAbs. For instance, external manipulations can involve immunization with variant antigens sharing conserved regions of a virus's spike while differing in variable parts. Experimental and computational investigations have focused on enhancing bnAb production through carefully designed vaccination protocols [14,19–23], but achieving success remains elusive. Albeit constrained, there are various routes to guiding affinity maturation, including use of variant antigens, number of vaccine doses, timing between shots, antigen composition in each shot, and dosage. The large space of variables for choosing vaccination protocols makes an unguided search unlikely to succeed.

Computational approaches have proven invaluable in systematically exploring the large space of vaccination protocols, providing insights into the outcomes of affinity maturation in response to diverse vaccine designs [24]. Several computational studies have investigated the efficacy of sequential and cocktail immunization procedures for eliciting bnAbs, highlighting an optimal difference between the variable parts of antigen variants included in vaccine shots that maximize the likelihood of bnAb production [22,23,25,26]. Signatures of such optimal antigen differences have been observed in HIV patients who naturally develop bnAbs [27,28] and in mice vaccinated by sequentially administered influenza antigens [19,29,30]. Mechanistic explanations for why an optimal difference between variant antigens promotes bnAb evolution have also been proposed [22,23,25,26]. It is worth noting that existing computational studies typically focus on immunization with a fixed cocktail of antigens [25,26] or sequential procedures where the variable parts of the antigens are sufficiently different to maximize the selection pressure in favor of the bnAb sequences [22,23]. However, Refs. [19,31] reported that sequential immunization protocols effectively progressed toward generating bnAb-like antibodies in engineered mice only when the structure and sequence of the stimulating epitopes changed gradually with subsequent boosts.

In this paper, we use analytical and computational methods to model the complex task of designing vaccines capable of eliciting immune responses that confer protection against diverse mutant strains within a virus family. To unveil guiding principles for optimizing bnAb production, we investigate generic time-dependent vaccination schemes that involve administering mixtures of relatively similar antigens over time. Our model significantly expands upon previous works in Refs. [23] and [22] by allowing variations in both the location and spread of the fitness landscape induced by the vaccination protocol. To describe BCR-antigen interactions, we adopt a shape-space representation [32,33], which enables us to capture the essence of the evolutionary dynamics of B cell populations in response to vaccination in a minimal framework.

The rest of the paper is organized as follows: In Sec. II, we describe a simplified model that captures the essential steps of affinity maturation of the B cell population in GCs, namely replication, apoptosis, and mutations. After developing a mean-field approximation to the stochastic dynamics, Secs. III A and III B provide optimization results based on a path integral representation, and operator approximation methods. In particular, for a Gaussian initial germline B cell population that is away from the bnAb state, we derive an exact identity [Eq. (10)] that directly provides the mean-field prediction for the optimal location for the focus of the fitness landscape imposed by the vaccine [Eq. (9)]. Using this identity, along with operator methods for approximating the solution to the population dynamics, we obtain a result [Eq. (16)] that enables us to investigate the optimal spread of the vaccine fitness. We next compare the mean-field predictions with stochastic simulations in Sec. IV. Finally, we present concluding remarks and outline future research directions in Sec. V. There are four Appendixes that contain further information regarding the derivation of the continuum mean-field model (Appendix A), the details of the path integral calculation (Appendix B), the operator approximation for the discrete-bin dynamics (Appendix C), and different aspects of the B cell dynamics under different vaccine protocols (Appendix D).

II. THEORETICAL MODEL

A. Minimal model for BCR evolution

To construct a mechanistic model of affinity maturation, we frame it as the population dynamics governing the evolution of B cells—or equivalently, their BCRs—where replication rates are determined by their binding affinity to the stimulating antigen. The interaction between BCRs and antigens is typically captured through phenomenological approaches such as the string model [22,25,34]. Here we adopt a more coarse-grained approach based on the concept of *shape-space* representation [26,32,33], where BCRs and antigens are represented by points in a d-dimensional Euclidean space (Fig. 1). Each dimension in this space corresponds to parameters pertinent to calculating BCR-antigen interactions, including factors such as amino acid sequences, spatial conformations, hydrophobicity, and the like. As these coordinates can assume distinct values, we initially discretize the shape space into a number of *similarity* bins indexed by their location $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$.

BCRs and antigens positioned closely in shape space are presumed to have complementary characteristics that allow them to bind strongly, while those situated farther apart exhibit weaker binding [32,35]. Following the approach outlined in Ref. [26], we characterize the binding free energy between a BCR and an antigen located in similarity bins $\boldsymbol{R}_{\mathrm{BCR}}$ and $\boldsymbol{R}_{\mathrm{Ag}}$ by their distance as

$$E_{\mathrm{bind}} \propto k_{\mathrm{B}} T \, ||\boldsymbol{R}_{\mathrm{BCR}} - \boldsymbol{R}_{\mathrm{Ag}}||^2.$$

Consequently, the associated binding affinity is expressed as $(E_{\mathrm{thr}} - E_{\mathrm{bind}})$, with E_{thr} denoting the threshold binding free energy for B cell activation. B cells bind to the antigen presented on the FDCs with an affinity determined by the equilibrium constant of the BCR-antigen binding, defined as $K_a = \exp[(E_{\mathrm{thr}} - E_{\mathrm{bind}})/k_{\mathrm{B}} T]$. Increased affinities correspond to stronger BCR-antigen binding with higher equilibrium con-
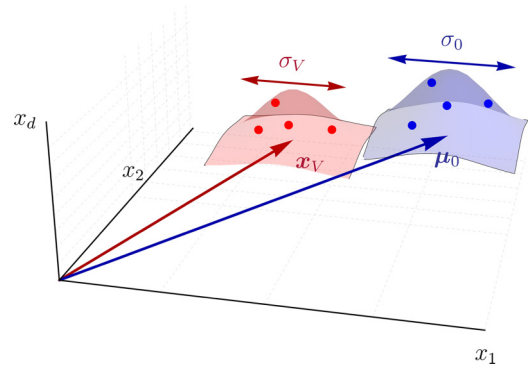

FIG. 1. Schematic depiction of a d-dimensional shape space. The germline B cell population $n_0(\boldsymbol{x})$ (in blue) is centered at $\boldsymbol{\mu}_0$ with a spread of σ_0, while the vaccine-induced fitness profile (in red) is parametrized by the its center $\boldsymbol{x}_V$ and spread σ_V. The blue and red dots denote specific antibodies and antigens, respectively, and the bell-like surfaces are the coarse-grained approximations of the corresponding density profiles. Given a constrained total vaccine dose [Eq. (4)], the objective is to determine a time-dependent vaccination profile that optimally shifts the B cell population in shape space, ultimately maximizing the final bnAb count corresponding to BCRs located at $\boldsymbol{x} = 0$.

stants. Affinity maturation can amplify the binding affinity of naive BCRs tenfold, and its associated K_a a thousandfold [36]. Furthermore, we assume that BCRs located closer to the origin ($\boldsymbol{x} = 0$) in the shape space have greater affinity toward the conserved residues of the antigen surface proteins. These BCRs exhibit increased tolerance to mutations in the surrounding variable residues of the antigen, facilitating the elicitation of bnAbs. Our objective, therefore, is to devise a vaccination procedure conducive to directing affinity maturation to optimally guide the initially naive BCRs toward the $\boldsymbol{x} = 0$ bin.

Since high-affinity B cells are more likely to undergo positive selection and further replication, the binding affinity to the vaccine antigen(s) effectively imposes a *fitness function* on GC B cells. When a single antigen type is included in a vaccine shot, the induced fitness function exhibits a sharply peaked distribution centered around the complementary BCR bin within the shape space. This distribution features a minimal spread, denoted as σ_{min}, to encompass nearby bins whose BCRs may also bind the given antigen, albeit with reduced affinity. Conversely, in the case of a cocktail shot containing mixtures of different antigen types, the overall fitness function results from the summation of the individual fitness functions induced by each antigen type.[1] This typically yields a nonconvex fitness landscape with multiple peaks and valleys. However, if the antigens included in a cocktail shot are sufficiently similar (e.g., sharing a common epitope containing conserved residues and surrounding variable portions that are different), the individual fitness peaks would cluster closely in shape space. Henceforth, we assume the administered anti-

---

[1] We are assuming a "see all" scenario wherein the antigens are well-mixed and presented homogeneously on FDCs, such that B cells encounter all the vaccine antigens at every round [26].

gens at each time point are sufficiently similar to allow for approximating the overall vaccine-induce fitness landscape with a smooth convex function encompassing these peaks (Fig. 1). The resultant coarse-grained fitness function is then roughly centered in the midst of the individual antigens, with a spread generally narrower for more similar antigens. We represent this coarse-grained fitness function as $V(\boldsymbol{x})$, with its peak located at $\boldsymbol{x}_V$ (referred to as the vaccine center or focus) and its spread denoted as σ_V. It is worth noting that under the assumption of similar antigens, $\boldsymbol{x}_V$ may generally be distant from the BCR bin associated with the bnAb sequences.

Now let us describe a minimal model for the affinity maturation process within the representation outlined above. The germline distribution that initially seeds the GCs is denoted by $n(\boldsymbol{x}, t = 0) \equiv n_0(\boldsymbol{x})$. As affinity maturation progresses, this distribution evolves over time to $n(\boldsymbol{x}, t)$. We model the evolution of BCRs during affinity maturation through the following stochastic processes:

(1) *Replication:* the rate of B cell replication is determined by the stimulating antigen(s) included in the vaccine dose. By altering the composition of the presented antigens, we assume it is possible to modulate the vaccine-induced fitness, $V(\boldsymbol{x}, t)$, across different bins in shape space, with corresponding time-dependent vaccine center and spread, $\boldsymbol{x}_V(t)$ and $\sigma_V(t)$. Note that we introduce a time dependence in V to accommodate vaccination protocols where different antigens or cocktails of antigens are introduced at different times.

(2) *Mutation:* BCR mutations occur with rate $\gamma_{\boldsymbol{x},\boldsymbol{y}}$, corresponding to jumps between shape-space bins $\boldsymbol{x}$ and $\boldsymbol{y}$. While nucleotide mutations can theoretically move a BCR between any two similarity bins, the coarse-grained nature of shape-space coordinates—reflecting the effect of many amino acids on the binding affinity—suggests that large jumps are exceedingly rare. For the simplified mean-field model in the next section, we will assume mutational jumps are local and only occur between nearby bins. This assumption allows us to model them by the diffusion of BCRs in the appropriate continuum limit.

(3) *Apoptosis:* B cells die with rate $\lambda_{\boldsymbol{x}}$, representing their baseline apoptotic tendency within GCs when not receiving positive selection signals [5,37].

The rules outlined above describe a generalized linear birth-death-mutation process, and they are encapsulated in the following Fock space master equation [38,39]:

$$
\begin{aligned}
\partial_t &P(\{n\}, t) \\
&= \sum_{\boldsymbol{x}} V_{\boldsymbol{x}}(t) \left[(n_{\boldsymbol{x}} - 1) P(\{n\} - \mathbb{I}_{\boldsymbol{x}}, t) - n_{\boldsymbol{x}} P(\{n\}, t)) \right] \\
&\quad + \sum_{\boldsymbol{x},\boldsymbol{y}} \gamma_{\boldsymbol{x},\boldsymbol{y}} [(n_{\boldsymbol{x}} + 1) P(\{n\} + \mathbb{I}_{\boldsymbol{x}} - \mathbb{I}_{\boldsymbol{y}}, t) - n_{\boldsymbol{x}} P(\{n\}, t)] \\
&\quad + \sum_{\boldsymbol{x}} \lambda_{\boldsymbol{x}} [(n_{\boldsymbol{x}} + 1) P(\{n\} + \mathbb{I}_{\boldsymbol{x}}, t) - n_{\boldsymbol{x}} P(\{n\}, t)].
\end{aligned} \tag{1}
$$

Here, the BCR population is denoted by $\{n\} = \{n_1, n_2, \ldots\}$, where n_i is a non-negative integer representing the occupation number of the ith bin, and $P(\{n\}, t)$ is the probability of having that specific population at time t. On the right-hand side of the above equation, we use $\{n\} \pm \mathbb{I}_{\boldsymbol{x}}$ to denote a population that

differs from $\{n\}$ by one more or one less B cell residing in bin $\boldsymbol{x}$.

Each line on the right-hand side of Eq. (1) represents the total rate of change of population probabilities (i.e., gain minus loss) caused by replication, mutation, and apoptosis events, respectively. It is worth noting that by describing the occupancy numbers of the shape-space bins, the solution to Eq. (1) would provide a probabilistic description of the entire BCR population in the high-dimensional shape space.

B. Mean-field dynamics and the continuum limit

The abstract model described above offers a simplified perspective on affinity maturation. However, deriving insights from the resulting master equation [Eq. (1)] is challenging due to the exponentially large Fock space and the interconnectedness of the time evolution for all configurations $\{n\}$. To address this, and obtain an analytically tractable formulation, we consider the dynamics of population averages, denoted by $\langle n_{\boldsymbol{x}}(t) \rangle = \sum_{\{n\}} n_{\boldsymbol{x}} P(\{n\}, t)$, and, subsequently, taking the continuum limit of the shape space. This leads to the following continuum mean-field equation (see Appendix A for derivation):

$$
\partial_t n(\boldsymbol{x}, t) = D \nabla^2 n(\boldsymbol{x}, t) + [V(\boldsymbol{x}, t) - \lambda] n(\boldsymbol{x}, t), \tag{2}
$$

where, for notational simplicity, we retain the symbol n to represent the average population *density* in the continuum limit of $\boldsymbol{x}$. We also assume the death rate λ and the diffusion coefficient D [obtained from the underlying mutation rates via Eq. (A7)] are constants independent of $\boldsymbol{x}$. Note that, as described earlier, we consider vaccines with either single antigen types or cocktails wherein the administered antigens are not substantially different from each other. Consequently, at any given time point, the most cross-reactive BCRs to the administered vaccine antigens—namely, BCRs located at $\boldsymbol{x}_V(t)$ in the center of the fitness—may not correspond to a bnAb sequence. This is somewhat distinct from the setting described in Ref. [23], where the administered antigens are different enough that bnAb sequences are always the most fit. The situation we consider here is akin to the experimental scenarios described in Refs. [19,40]. In these experiments, similar to our model, the sequentially administered antigens bridge the gap between epitopes that efficiently activate the appropriate germlines and those for which bnAbs are needed to target the conserved region.

Solving the mean-field dynamics [Eq. (2)] requires specifying the initial germline distribution, which typically has little overlap with the target bin in shape space. We thus assume $n_0(\boldsymbol{x})$ is centered at a distant location $\boldsymbol{\mu}_0$, and it has a width σ_0 that is an indication of the diversity of the germline BCRs (see Fig. 1). In the spirit of the central limit theorem, we model $n_0(\boldsymbol{x})$ by a normal distribution,

$$
n_0(\boldsymbol{x}) = N_0 \frac{e^{-\frac{(\boldsymbol{x} - \boldsymbol{\mu}_0)^2}{2\sigma_0^2}}}{\left(2\pi \sigma_0^2\right)^{d/2}} \equiv N_0 \mathcal{G}(\boldsymbol{x} - \boldsymbol{\mu}_0, \sigma_0^2), \tag{3}
$$

where $\mathcal{G}$ denotes the Gaussian function. Here, N_0 represents the total number of activated B cells at $t = 0$.

Our objective is to determine an optimal vaccination strategy, characterized by $\{\boldsymbol{x}_V(t), \sigma_V(t)\}$, that maximizes the final bnAb population denoted by $n(\boldsymbol{x} = 0, t_f) \equiv n_f(0)$ while adhering to a constrained total vaccine dose, i.e.,

$$\int_0^{t_f} dt \int d^d\boldsymbol{x}\, V(\boldsymbol{x}, t) = \text{const.} \tag{4}$$

Equation (2) resembles the imaginary-time Schrödinger equation with a time-dependent potential [41]. Typically, such equations lack analytical solutions, rendering a direct approach to the maximization problem infeasible.

Before delving into the analysis of the model, we summarize the assumptions made to simplify the model and acknowledge their associated limitations:

(1) We interpret Eq. (4) as imposing a constraint on the total amount of administered antigens in vaccine doses. This interpretation implies a linear relation between the fitness function V [appearing on the left-hand side of Eq. (4)] and the amount of the administered antigen. However, $V(\boldsymbol{x}, t)$ can generally be a complicated function of the antigen concentration, incorporating nonlinear terms that arise from various factors such as competition between B cells for GC entry, positive selection within GCs, and epitope masking (see, e.g., Refs. [5,42–44]). These effects are ignored in the present model. Nonetheless, the linear relation between fitness and vaccine dose remains a plausible approximation if selection based on the amount of internalized antigen is not overly stringent. This approximation is consistent with the linear structure of the birth-death-mutation model of Eq. (1).

(2) We assume that mutational jumps occur independently of replication events, exert incremental effects on BCR affinity, and they are equally probable in any direction (i.e., they are isotropic). These assumptions underlie the symmetric form of the diffusion term in the mean-field dynamics, as included in Eq. (2), along with the vanishing drift term (see Appendix A). In a more realistic model, mutations primarily occur during B cell replication events. Point mutations in a BCR sequence typically reduce its affinity to a given antigen, making them more likely to be deleterious or silent rather than beneficial [45]. Consequently, the corresponding coarse-grained dynamics in the shape space would feature a nonhomogeneous diffusion term and a nonvanishing drift term. Additionally, since deleterious mutations can significantly reduce the affinity of a BCR to the antigen [43,45,46], potentially long-ranged steps (i.e., nonlocal in the shape space) should generally be included in a more accurate coarse-grained mean-field dynamics.

III. ANALYTICAL RESULTS

As mentioned earlier, solving the mean-field equation (2) explicitly for a general vaccine fitness function, and thus direct optimization of bnAbs, is not feasible. In this section, we propose a method to simplify this maximization by applying a coordinate change to the path integral representation of the solution to the mean-field dynamics [Eqs. (7) and (10) below].

A. Path integral representation and the optimal location of the fitness peak

We begin constructing the path integral formulation by rearranging Eq. (2) for an infinitesimal timestep δt:

$$n(\boldsymbol{x}, t + \delta t) = \{1 + \delta t\, D\nabla^2 + \delta t[V(\boldsymbol{x}, t) - \lambda]\}n(\boldsymbol{x}, t)$$
$$= e^{\delta t D\nabla^2} e^{\delta t[V(\boldsymbol{x}, t) - \lambda]} n(\boldsymbol{x}, t) + O(\delta t^2). \tag{5}$$

Equation (5) provides a formal method for evolving the population profile in infinitesimal time steps. By repeatedly applying Eq. (5) to evolve the population over time, the final population $n(\boldsymbol{x}, t_f) \equiv n_f(\boldsymbol{x})$ can be expressed by the formal time-ordered product [41]

$$n_f(\boldsymbol{x}) = \lim_{M \to \infty} \prod_{j=0}^{j=M} e^{\delta t D\nabla^2}\, e^{\delta t[V(\boldsymbol{x}, j\delta t) - \lambda]}\, n_0(\boldsymbol{x}). \tag{6}$$

Here, we have defined $\delta t = \frac{t_f}{M}$ with M as the number of time increments, and $V_j(\boldsymbol{x})$ denotes the vaccine-induced fitness function at time $j\delta t$. The products in Eq. (6) act in the operator sense:[2] starting from the germline population n_0, at the jth infinitesimal timestep δt, the B cell population is locally expanded/contracted according to the vaccine operator $e^{\delta t[V(\boldsymbol{x}, j\delta t) - \lambda]}$, followed by diffusion broadening under the operator $e^{\delta t D\nabla^2}$.

The effect of the diffusion operator on a test function $\eta(\boldsymbol{x})$ is given by the convolution

$$e^{\delta t D\nabla^2} \eta(\boldsymbol{x}) = \int d^d\boldsymbol{x}'\, \mathcal{G}(\boldsymbol{x} - \boldsymbol{x}', 2D\delta t)\, \eta(\boldsymbol{x}').$$

Due to the nonlocal nature of this term, the final population at any location in shape space results from all potential ways that different parts of the initial population spread and expand until they reach that location at the final time. Upon inserting the diffusion operators in Eq. (6) and integrating over all intermediate points, the final population can be more explicitly expressed as

$$\int \prod_{j=0}^{M} d^d\boldsymbol{x}_j\, \mathcal{G}(\boldsymbol{x}_{j+1} - \boldsymbol{x}_j, 2D\delta t)\, e^{\delta t(V(\boldsymbol{x}_j, j\delta t) - \lambda)}\, n_0(\boldsymbol{x}_0),$$

with $\boldsymbol{x}_{M+1} = \boldsymbol{x}$. This expression is then rearranged into the standard path integral form [41,52]

$$n_f(\boldsymbol{x}) = e^{-\lambda t_f} \int d^d\boldsymbol{x}_0\, n_0(\boldsymbol{x}_0) \int_{\boldsymbol{r}(0) = \boldsymbol{x}_0}^{\boldsymbol{r}(t_f) = \boldsymbol{x}} \mathcal{D}\boldsymbol{r}(u)\, e^{-S[\boldsymbol{r}, \dot{\boldsymbol{r}}]}, \tag{7}$$

where $\mathcal{D}\boldsymbol{r} \propto d\boldsymbol{x}_1, \ldots, d\boldsymbol{x}_M$ stands for the measure of all paths connecting the end points, each weighted according to its "action,"

$$S[\boldsymbol{r}, \dot{\boldsymbol{r}}] = \int_0^{t_f} du \left(\frac{\dot{\boldsymbol{r}}^2(u)}{4D} - V(\boldsymbol{r}(u), u) \right). \tag{8}$$

---

[2]This product formula is also known as the first-order Suzuki-Trotter decomposition [47,48] and it forms the basis for the path integral formulation of quantum mechanics [41], quantum simulators [49,50], and the so-called operator splitting methods [51].
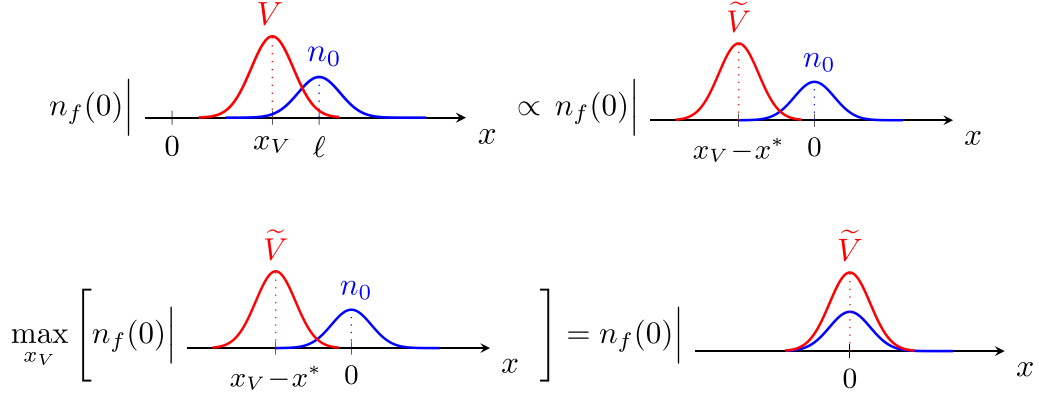
FIG. 2. A visual representation of the path integral result. Top panel: the final number of bnAbs, $n_f(0)$, for a Gaussian germline n_0 that starts away from the target bin ($x = 0$) and evolves under a general vaccine fitness V, is directly proportional to the final bnAbs count when n_0 starts at the target bin and evolves under the modified fitness $\widetilde{V}$ defined in Eq. (11). The proportionality relationship is governed by Eq. (10). Bottom panel: To optimize $n_f(0)$ in the mapped problem, the modified fitness $\widetilde{V}$ needs to be centered at the target, which is achieved by adjusting the optimal vaccine center according to Eq. (12).

Equation (7) indicates that evolutionary paths passing through regions with high vaccine fitness tend to have larger weights in the path integral. In principle, Eq. (7) provides a formal expression for the solution to the mean-field dynamics, applicable to any arbitrary BCR germline and vaccination fitness landscape.

We intend to utilize the path integral expression (7) for $x = 0$ to deduce features of the vaccine fitness V that maximizes the bnAb count. Intuitively, we expect the optimal fitness to move the germline population gradually from its initial position $x = \mu_0$ towards the target bnAb state at $x = 0$. To achieve this, the optimal fitness must remain sufficiently close to the peak of the existing B cell population at every instant to prevent B cell extinction by apoptosis. On the other hand, the fitness should also be positioned far enough from the B cell population to optimally transport it toward $x = 0$ within the finite time window t_f. Based on these considerations, we anticipate the optimal fitness to initially focus slightly away from the peak of the B cell population and, over time, move toward the target bin. However, this argument only provides a qualitative picture of the motion of the optimal fitness and does not contain enough information to uniquely determine the time dependence of the fitness function. A simple scenario for the motion of the optimal fitness center involves linear motion with constant velocity. More complex scenarios may involve nonlinear motion over time. Given the linear structure of the mean-field dynamics (2) and the corresponding quadratic form of the path integral action (8), this simple case seems a plausible candidate for being optimal and thus warrants more systematic investigation.

To examine the optimality of the fitness function whose center moves with a constant velocity, we use the path integral formula, Eq. (7), in a straightforward but somewhat nonstandard manner. We begin by substituting the germline distribution (3) into Eq. (7), and subsequently we rewrite the resulting expression in the new coordinate frame with a moving origin. Denote the location of the origin of this frame by $z(t) = z_0 - vt$, where z_0 and v indicate the initial location at $t = 0$ and the constant velocity of the moving origin, respectively. As we outline in Appendix B, for the special case

of a Gaussian germline distribution, it is possible to select z_0 and v such that the transformed path integral expression written in the moving frame would have the same mathematical structure as the path integral expression in the original frame. By doing so, we obtain the path integral expression in the moving frame for a modified germline that is centered at the target bin (see Fig. 2).

In particular, for the moving frame whose origin is defined by z_0^* and v^* given in Eq. (B4) of Appendix B, the shape-space coordinates are transformed by applying the linear map $x \to x + x^*(t)$, where

$$x^*(t) \equiv \mu_0 \frac{2D(t_f - t)}{2Dt_f + \sigma_0^2}. \tag{9}$$

After some algebraic manipulations of the path integral expression as detailed in Appendix B, we arrive at the following exact relationship for the bnAb counts:

$$\frac{n_f(0)\big|_{\mu_0 = \ell; V}}{n_f(0)\big|_{\mu_0 = 0; \widetilde{V}}} = \exp -\frac{\ell^2}{2(\sigma_0^2 + 2Dt_f)}. \tag{10}$$

Here, the numerator on the left-hand side denotes the final population at $x = 0$ (i.e., the number of final bnAbs) given that the normally distributed initial population (3) is centered at $\mu_0 = \ell$ and evolves under the vaccine fitness $V(x, t)$. The denominator, on the other hand, represents the final bnAb count for a germline population that is centered at $\mu_0 = 0$ and is evolved under the transformed fitness $\widetilde{V}$ defined as

$$\widetilde{V}(x, t) \equiv V(x + x^*(t), t). \tag{11}$$

Therefore, the exact relationship (10) simply relates the number of bnAbs generated by fitness V, with the germline population initially centered away from the target in shape space, to the number of bnAbs developed under the modified fitness $\widetilde{V}$, assuming the germline is already centered at the target (see the top panel of Fig. 2).

The crucial observation here is that the right-hand side of Eq. (10) is independent of the choice of the vaccine parameters. Thus, the maximum of the numerator on the left-hand side as V is varied [by changing $\{x_V(t), \sigma_V(t)\}$], must coincide

with the maximum of the denominator as $\widetilde{V}$ is varied. As a result, Eq. (10) effectively converts the maximization of $n_f(0)|_{\boldsymbol{\mu}_0=\boldsymbol{\ell};V}$ into the maximization of $n_f(0)|_{\boldsymbol{\mu}_0=0;\widetilde{V}}$. This mapping simplifies the optimization problem: since for $\boldsymbol{\mu}_0 = 0$ the starting germline population is already peaked at, and symmetric around, the target bin $\boldsymbol{x} = 0$, it makes sense that for $\widetilde{V}$ to be optimal, it should also be peaked at the target and symmetric with respect to it at all times (see Fig. 2). Using this argument, and recalling that the maxima of the numerator and denominator in Eq. (10) coincide, we then conclude that

$$\boldsymbol{x}_V(t)\big|_{\text{optimal}} = \boldsymbol{x}^*(t) = \boldsymbol{\mu}_0 \frac{2D(t_f - t)}{2Dt_f + \sigma_0^2}. \quad (12)$$

It is worth noting that for a narrow germline population ($\sigma_0 \ll \sqrt{Dt_f}$) with less BCR diversity, the optimal vaccine center given by Eq. (9) simply starts close to the naive population at $\boldsymbol{\mu}_0$ and reaches the target state ($\boldsymbol{x} = 0$) at the final time. However, for a wider initial population ($\sigma_0 \gtrsim \sqrt{Dt_f}$)—corresponding to a more diverse germline—the optimal center is shifted closer to the target bin. This reflects a change in the importance of paths contributing to the bnAb count $n_f(0)$ in Eq. (7). As described intuitively earlier, one needs to position the vaccine fitness over time such that it gradually moves the germline population from $\boldsymbol{x} = \boldsymbol{\mu}_0$ towards the target state at $\boldsymbol{x} = 0$, while also remaining sufficiently close to the peak of the existing B cell population to avoid extinction by apoptosis. If $V(\boldsymbol{x}, t)$ is positioned too close to the population $n(\boldsymbol{x}, t)$, it may not optimally move the population to reach $\boldsymbol{x} = 0$ within the finite time interval t_f; conversely, if V is centered too far from the peak of $n(\boldsymbol{x}, t)$, there is a risk of population collapse as the vaccine-induced replication is directed to locations with fewer B cells to begin with. The optimal center [Eq. (9)] aligns with this intuitive understanding and also demonstrates how the germline width σ_0 influences the optimal position of the vaccine fitness. In addition, Eq. (12) reveals that the mean-field optimal center moves with a constant velocity from its starting position in the shape space towards the target bin. These features are further illustrated in the numerical and simulation results discussed in Sec. IV and Appendix D.

B. Operator approximation and the optimal spread of the fitness function imposed by vaccine antigens

As discussed above, Eq. (10) maps the problem of maximizing bnAbs to finding the optimal modified fitness function $\widetilde{V}$ that maximizes $n_f(0)|_{\boldsymbol{\mu}_0=0;\widetilde{V}}$. Upon replacing V by $\widetilde{V}$, Eq. (2) describes the expansion of the transformed germline population centered at $\boldsymbol{x} = 0$ under the modified fitness $\widetilde{V}$, coupled with diffusive spreading that arises from BCR mutations.

To investigate the optimal vaccine spread, we employ operator approximations. Starting from Eq. (6), we rearrange the operators in the time-ordered product formula using a second-order truncation of the celebrated Baker-Campbell-Hausdorff (BCH) formula [48], namely

$$e^{\epsilon A} e^{\epsilon B} = e^{\epsilon B} e^{\epsilon A} e^{\epsilon^2[A,B]} (1 + O(\epsilon^3)), \quad (13)$$

where A and B are generally noncommuting operators, and $C \equiv [A, B] = AB - BA$. We assign $\epsilon \to \delta t$, $A \to D\nabla^2$, and $B \to \widetilde{V}_j(\boldsymbol{x})$, and then we utilize this formula to reorder the

alternating operator product in Eq. (6) such that all vaccine fitness operators $e^{\delta t \widetilde{V}_j}$ are grouped on the left in the product, all the diffusion operators $e^{\delta t D\nabla^2}$ are collected on the right, and the commutator terms that result from swapping these operators appear in between. The apoptosis operator $e^{-\lambda t_f}$ commutes with all the other operators.

To achieve such an ordering of the operators, we start by moving $e^{\delta t \widetilde{V}_0}$ through the M diffusion operators $e^{\delta t D\nabla^2}$ on its left in Eq. (6), resulting in a net commutator term $e^{(\delta t)^2 M C_0}$, where the commutator operator reads

$$C_0 = [D\nabla^2, \widetilde{V}_0(\boldsymbol{x})] = 2D\widetilde{V}_0'(\boldsymbol{x})\nabla + D\widetilde{V}_0''(\boldsymbol{x}). \quad (14)$$

Subsequently, the next fitness operator, $e^{\delta t \widetilde{V}_1}$, needs to pass through $(M - 1)$ diffusion operators on its left, thereby creating a commutator contribution $e^{(\delta t)^2(M-1)C_1}$. This process continues for all fitness operators, and in the limit of small increments $\delta t = \frac{t_f}{M} \to 0$, we obtain[3]

$$n_f\big|_{\boldsymbol{\mu}_0=0;\widetilde{V}} \approx N_0 e^{-\lambda t_f} e^{\int_0^{t_f} dt\, \widetilde{V}(\boldsymbol{x},t)}$$

$$\times\, e^{D \int_0^{t_f} dt\,(t_f - t)\left(2\widetilde{V}'(\boldsymbol{x},t)\nabla + \widetilde{V}''(\boldsymbol{x},t)\right)}$$

$$\times\, e^{t_f D\nabla^2} \mathcal{G}(\boldsymbol{x}_0, \sigma_0^2). \quad (15)$$

Unlike the exact expression in Eq. (6), where the fitness and mutation (diffusion) alternate in the operator product, Eq. (15) suggests that n_f can be approximated by first diffusing the starting profile n_0 under the (full) diffusion operator $e^{t_f D\nabla^2}$. Then, the commutator operators act on this diffused population profile, and finally, the vaccine fitness operators (and the death operator) expand the profile to its final form. The action of $e^{t_f D\nabla^2}$ on the Gaussian initial population widens its profile and yields $\mathcal{G}(\boldsymbol{x}_0, \sigma_0^2 + 2Dt_f)$, which is independent of the fitness profile. The symmetry of this updated profile with respect to $\boldsymbol{x}_0 = 0$ still implies a similar symmetry for the optimal $\widetilde{V}(\boldsymbol{x}, t)$, resulting in $\widetilde{V}'(\boldsymbol{x} = 0, t) = 0$. Consequently, Eq. (15) simplifies to

$$n_f(0)\big|_{\boldsymbol{\mu}_0=0;\widetilde{V}} \approx N_0 e^{-\lambda t_f} e^{\int_0^{t_f} dt\, \widetilde{V}(0,t)}$$

$$\times\, e^{D \int_0^{t_f} dt(t_f - t)\widetilde{V}''(0,t)} \mathcal{G}(0, \sigma_0^2 + 2Dt_f). \quad (16)$$

For any choice of the vaccine-induced fitness profile, Eq. (16) offers a starting point to investigate the optimal vaccine spread by providing an explicit expression whose optima can be derived by straightforward differentiation. Narrowing our focus, we consider the Gaussian vaccine-induced fitness

$$V_{\mathcal{G}}(\boldsymbol{x}, t) \equiv A_V\, \mathcal{G}(\boldsymbol{x} - \boldsymbol{x}_V(t), \sigma_V(t)^2). \quad (17)$$

---

[3]Each truncation of the BCH formula introduces an error of order $(\delta t)^3$. It can be shown that the number of such error terms grows at least as $\sim M^3$, rendering this truncation an uncontrolled approximation. Despite this, the result obtained through this approach matches that of a perturbative solution in powers of the nondimensionalized diffusion coefficient. The second-order truncation remains valid for small diffusion limits and as long as the total time t_f is sufficiently short.

Recall that this choice is consistent with our initial assumption that, in the shape-space representation, the fitness can be parametrized by the location of its center and its spread. Furthermore, Eq. (17) also satisfies the total-dose constraint given by Eq. (4). However, it is important to note that Eq. (17) does not represent the most general Gaussian fitness profile allowed by the constraint, as it only considers a fixed *vaccine strength* A_V. The more general case would involve a vaccine dose that varies over time, potentially leading to larger and more diverse antibody responses [53–55]. We defer the investigation of such scenarios to future studies.

By selecting the optimal vaccine center according to Eq. (12), the Gaussian fitness of Eq. (17) transforms according to Eq. (11), resulting in the modified fitness $\widetilde{V}_{\mathcal{G}}(\boldsymbol{x}, t) = A_V \mathcal{G}(\boldsymbol{x}, \sigma_V^2(t))$, which is still a Gaussian profile but now centered at the origin. Substituting $\widetilde{V}_{\mathcal{G}}$ into Eq. (16), we obtain

$$
n_f(0)\big|_{\mu_0 = 0; \widetilde{v}_{\mathcal{G}}} \approx \frac{N_0 e^{-\lambda t_f}}{\sqrt{2\pi\left(\sigma_0^2 + 2Dt_f\right)}}
$$
$$
\times \exp \int_0^{t_f} dt \left[\frac{A_V}{\sqrt{2\pi}\,\sigma_V(t)} \left(1 - \frac{D(t_f - t)}{\sigma_V^2(t)}\right) \right].
\tag{18}
$$

To maximize this expression, we set its functional derivative with respect to σ_V to zero. Assuming $\sigma_V(t) \geqslant \sigma_{\min}$ at all times (see the discussion below), we finally obtain

$$
\sigma_{\mathrm{BCH}}(t) = \max\{\sqrt{3D(t_f - t)}, \sigma_{\min}\}.
\tag{19}
$$

In deriving Eq. (19), we have assumed that $\sigma_V(t) \geqslant \sigma_{\min}$, which prevents $V_{\mathcal{G}}$ from reaching arbitrarily large values by imposing a maximum fitness value $A_V/(\sqrt{2\pi}\sigma_{\min})$. As discussed in Sec. II, the fitness function exhibits a minimal yet finite spread for vaccine shots containing a single antigen type, whereas cocktail shots typically exhibit a wider spread. Moreover, in Appendix C, we demonstrate that the optimal spread retains a finite minimum value when employing the BCH approximation for B cell dynamics on discrete similarity bins (before taking the continuum limit in shape space). Additionally, the minimum spread can also reflect practical constraints on how precisely the vaccine-induced fitness can be concentrated in different regions of shape space. We posit that $\sigma_{\min}$ encompasses all of these factors.

Equation (19) can be understood by noting that within a temporal window of duration Δt, diffusion affects a range of order $\sqrt{D\Delta t}$ of bins on the x axis. To maximize $n_f(0)$ while adhering to a constrained total dose, it would thus be sensible to concentrate the fitness on BCRs in a region around the target bin that has the potential to reach the target location by mutation within the available time. Replication events at larger mutational distances are unlikely to reach the target within the finite allotted time. This suggests that at any time $0 < t < t_f$, the optimal $\widetilde{V}$ should at most cover BCRs within a range of order $\sqrt{D(t_f - t)}$ while remaining peaked at $\boldsymbol{x} = 0$. Once the vaccine spread reaches the smallest possible value $\sigma_{\min}$, it should remain at that value until the final time. In a qualitative sense, Eq. (19) suggests that an optimal vaccination protocol should initially have a broader spread and contain shots with more diverse antigens to cover a larger range of BCRs, which would provide the germline population

a chance to grow and diversify. As time progresses, vaccine epitopes gradually narrow down according to Eq. (19), concentrating the fitness on expanding those BCRs within an accessible mutational distance from the target bin.

It is also worth noting that the right-hand side of the BCH approximate result (16) depends only on time integrals of $\widetilde{V}(0, t)$ and $\widetilde{V}''(0, t)$, making it a strictly local function of the transformed vaccine fitness $\widetilde{V}$ at $\boldsymbol{x} = 0$. This locality aligns with the second-order truncation of the BCH formula, as higher-order terms in BCH's nested commutator expansion would introduce contributions involving more derivatives of the fitness, leading to nonlocal expressions in $\boldsymbol{x}$. Since the mean-field Eq. (2) aims to describe the coarse-grained temporal evolution of the B cells toward the bnAb state, and considering that bnAbs are typically produced over long timescales in natural circumstances [15,17], it may be reasonable to assume the slow diffusion regime, where local approximations such as the one above remain reliable at the mean-field level. However, it is important to acknowledge that even in the slow diffusion limit, this approximation might break down if either the vaccine fitness or the B cell population profile is sharply peaked, as higher-order derivatives may become significant in the expansion.

Furthermore, as noted in Sec. II B, our approach differs from Refs. [22,23] in our assumption that antigens administered at each time point share sufficient similarity, that bnAb sequences are not inherently the fittest. Given this distinction and the discussion of the previous paragraph, while Eq. (19) bears resemblance to the findings of Ref. [23]—where the optimal fitness function appeared narrower in subsequent vaccinations—we approach the interpretation of Eq. (19) with caution. In the next section, we will compare the performance of protocols employing a width σ_{BCH} with those utilizing a narrow fitness width $\sigma_{\min}$ (see Fig. 3) through numerical results and simulations. The comparative analysis aims to shed light on the effectiveness of different vaccine spread strategies and will elucidate a distinct interpretation of the optimal spread in our study.

IV. COMPUTATIONAL RESULTS

In this section, we delve into the dynamics of the B cell population within a one-dimensional shape space ($d = 1$) and investigate both the numerical solution of the continuum mean-field model [Eq. (2)] and simulations of the discrete stochastic dynamics [Eq. (1)]. It is important to recall that while the master equation deals with discrete shape-space bins, representing BCR occupancy with integer values, the mean-field model offers a continuum description, neglecting population fluctuations. Consequently, it is not immediately clear whether discrete sample averages would align with the mean-field results. To establish a comparison between the two approaches, we employ the stochastic simulation algorithm (SSA) in the manner of Gillespie [57,58] to generate realizations of stochastic dynamics and then compare the sample averages with those obtained from the mean-field treatment [59]. This numerical examination sheds light on how effectively the optimal vaccine fitness obtained from the mean-field model translates to the realm of discrete stochastic dynamics.
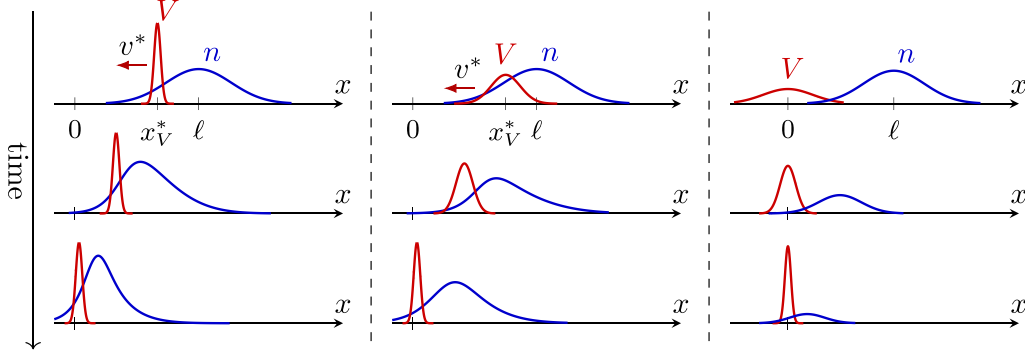
FIG. 3. Schematic of three different vaccination protocols, where the red and blue curves represent the vaccine-induced fitness and the B cell population, respectively. Left: the optimal protocol in which the fitness center follows $x_V^*(t)$ given in Eq. (9) while the fitness spread is fixed at the minimum value $\sigma_{\min}$. Center: the BCH protocol, where the fitness center follows $x_V^*(t)$ as in the optimal protocol but the fitness spread narrows over time as described in Eq. (19). Right: the protocol of Refs. [22,56], where the antigens administered at each time point are sufficiently different such that the bnAb sequence would have the highest fitness, and the fitness spread becomes narrower over time (while still reaching the left tail of the B cell population).

To set up our numerical experiments and simulations, we must define parameters for (i) the germline population, specifically N_0, μ_0, and σ_0; (ii) the dynamics, including D, t_f, λ; and (iii) the vaccine's strength, denoted by A_V. To make these parameter choices more grounded, we draw from empirical observations and rough estimates.

First, it is crucial to consider the slow and delayed production of bnAbs in HIV patients [15]. This implies that the germline population should be initially distant from the target bin ($x = 0$). For instance, it is known that numerous mutations are typically required for germline B cells to evolve into bnAbs that can target CD4 binding sites [14]. Therefore, we choose the simulation parameters for the time required to reach the bnAbs state by mutations only, denoted by t_{mut}, to be longer than the total vaccination time:

$$t_{\mathrm{mut}} \equiv \frac{\mu_0^2}{D} \gtrsim t_f. \tag{20}$$

Note that t_{mut} represents the (mean-field) time required for the BCR population to reach the bnAbs state by mutations.

Next, in the absence of vaccine-induced replication, the apoptotic B cell population is expected to go extinct [1,37], presumably on a timescale shorter than t_{mut}. To turn this observation into a relationship among the parameters, we utilize the mean-field approximation, within which the average population size decays exponentially as $N_0 e^{-\lambda t}$. We then crudely define the extinction as a population size falling below some minimum threshold, say 1.[4] The extinction condition may then be expressed as

$$\ln N_0 \lesssim \lambda t_{\mathrm{mut}} = \frac{\lambda \mu_0^2}{D}. \tag{21}$$

Drawing from the findings of Ref. [23], we further posit that an effective vaccine would provide sufficient fitness to prevent the extinction. We thus choose the vaccine strength

---

[4]The threshold is needed since, in the mean-field approximation, the average population size never truly reaches zero.

A_V so that the total vaccine-induced fitness given at each time exceeds the total death rate across all shape-space bins, i.e.,

$$A_V \gtrsim \lambda N_{\mathrm{bin}} \Delta x_{\mathrm{bin}}, \tag{22}$$

where N_{bin} represents the number of bins and Δx_{bin} is the length of a bin, assumed to be 1. [Note that as per Eq. (17), A_V has the dimension of (shape-space) length per time.]

Lastly, we note that a further condition to avoid extinction under the optimal vaccine center is for x^* to remain close to the population's center; otherwise, most of the provided fitness is expended on bins with few to no B cells. Given the difficulty of analytically solving for the population center as a function of time, we only impose this condition on the germline at $t = 0$, which corresponds to $|\mu_0 - x^*(t = 0)| \lesssim \sigma_0$. Substituting for $x^*(t = 0)$ from Eq. (9) then yields the fourth condition:

$$t_f \gtrsim \frac{\sigma_0(\mu_0 - \sigma_0)}{D}, \tag{23}$$

which, for a given diffusion constant D, sets a lower bound for the total vaccination time t_f. Typically $\sigma_0 \ll \mu_0$, ensuring that this bound is significantly smaller than the diffusive timescale μ_0^2/D required for the germline to reach the target by mutations alone [cf. Eq. (20)].

Guided by these conditions, we choose $N_{\mathrm{bin}} = 43$ corresponding to the shape-space bins $x \in \{-21, -20, \ldots, 0, \ldots, 20, 21\}$ for the stochastic simulations. This range is associated with the interval $(-21, 21)$ for the continuum mean-field numerics. We set the death rate to $\lambda = 0.001$ and the mutation rate to $\mu = 0.1$ in SSA, corresponding to a mean-field diffusion coefficient of $D = 0.05$ (see Appendix A). We assume the germline population is normally distributed according to Eq. (3), with an initial population size of $N_0 = 10$. In addition, the imposed vaccine fitness follows the Gaussian shape of Eq. (17), where we fix the vaccine strength at $A_V = 0.08$. For the stochastic dynamics, we utilize SSA to take averages over 1000 realizations of Eq. (1) for a total time of $t_f = 400$. To facilitate the simulation, we implement SSA along with an operator splitting scheme [51] where we alternate between
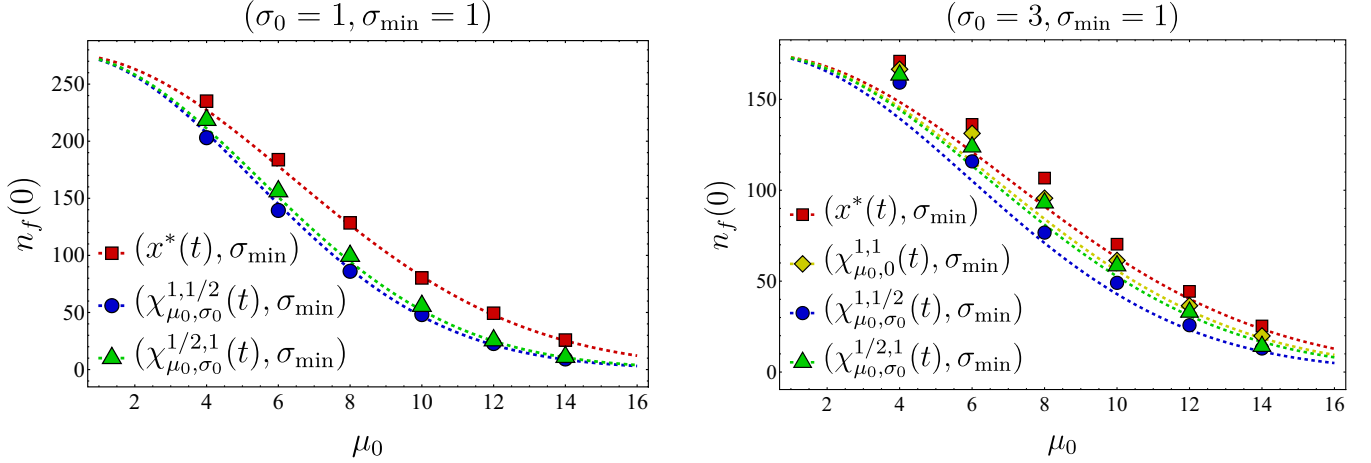
FIG. 4. Final bnAb counts for different vaccination protocols in a one-dimensional shape space as obtained from the continuum mean-field numerics (dashed lines) and SSA averages (solid points), with the choice of parameters $A_V = 0.08$, $D = 0.05$, $\lambda = 0.001$, and $t_f = 400$. Left: final bnAb count as a function of the germline center μ_0 assuming a fixed germline width $\sigma_0 = 1$. The vaccine-induced fitness is assumed to follow the Gaussian form of Eq. (17) with the vaccine center located at $\chi_{\mu_0,\sigma_0}^{q,p}(t)$ [see Eq. (24)] and the vaccine spread set to the minimum allowed $\sigma_{\min} = 1$. Different colors correspond to different choices for q and p, with $q = p = 1$ corresponding to the optimal center $x^*(t)$. The $x^*(t)$ protocol indeed outperforms the other choices for the vaccine center. Right: similar to the left panel but for a wider germline width ($\sigma_0 = 3$). We have also included the results obtained from setting the vaccine center according to $\chi_{\mu_0,0}^{1,1}(t)$. Here, SSA averages deviate from the continuum mean-field results for $\mu_0 \sim \sigma_0$, indicating the breakdown of the continuum approximation.

periods of stochastic replication events and stochastic mutation or death events [cf. Eq. (13) for the corresponding mean-field expression]. In the following, we present the SSA results for $M = 200$ steps, although consistent results are obtained for $M = 400$. Note that the mean-field numerics presented below are obtained by direct numerical solution of the continuous-time dynamics described given by Eq. (2).

Fixing the vaccine spread at $\sigma_{\min} = 1$ and the germline width at $\sigma_0 = 1$, we first examine the effect of different vaccine centers on the final bnAb count by selecting the vaccine center from the family

$$\chi_{\ell,\kappa}^{q,p}(t) \equiv \frac{\ell}{1 + \frac{\kappa^2}{2Dt_f}}\left[1 - \left(\frac{t}{t_f}\right)^q\right]^p. \qquad (24)$$

In particular, we compare the final bnAb count for $\chi_{\mu_0,\sigma_0}^{1,1}$—corresponding to x^* as defined in Eq. (9)—with that of $\chi_{\mu_0,\sigma_0}^{1,1/2}$ (vaccine center accelerating towards the target), $\chi_{\mu_0,\sigma_0}^{1/2,1}$ (vaccine center decelerating towards the target), and $\chi_{\mu_0,0}^{1,1}$ (uniform motion similar to x^* but not accounting for the germline's width). The left panel in Fig. 4 shows the final bnAb counts as a function of the germline distance from the target for three different vaccine centers $x^*(t)$ (red), $\chi_{\mu_0,\sigma_0}^{1,1/2}(t)$ (blue), and $\chi_{\mu_0,\sigma_0}^{1/2,1}(t)$ (green). We observe that in each case, the SSA sample averages (solid points) closely match the continuum mean-field numerics (dashed lines). To understand this, note that with $\sigma_0 = 1$, the germline population is narrowly distributed, so the entire population would roughly move together in shape space. When this population is relatively far from $x = 0$, moving between neighboring (discrete) bins would only change the distance to the target bin incrementally, thus making the continuum limit a reliable approximation of the discrete dynamics. Additionally, with the chosen vaccine strength [Eq. (22)], the population size increases and makes

the population number fluctuations less relevant, thus rendering the mean-field description even more accurate. This is also reflected in the fact that the optimal center $x^*(t)$—obtained from the continuum mean-field approximation—outperforms the other protocols even in the discrete stochastic dynamics.

The right panel in Fig. 4 presents a similar comparison of the final bnAb counts, but for a broader germline with $\sigma_0 = 3$. Here, we include a fourth choice for the vaccine center, namely $\chi_{\mu_0,0}^{1,1}(t)$ (in yellow), which differs from $x^*(t)$ in that its starting position and velocity are not adjusted by σ_0. We still observe that $x^*(t)$ outperforms the other protocols, although the final bnAb counts (vertical axis) are generally smaller compared to the $\sigma_0 = 1$ case. This is because the population does not move toward the target as synchronously when the initial population is spread more widely. It is also seen that when μ_0 becomes comparable with σ_0, SSA averages deviate from the continuum mean-field results, which might indicate the breakdown of the continuum approximation for that range of parameters. Nevertheless, the qualitative trend of the SSA sample averages is still captured by the mean-field solution. In Appendix D, we present temporal snapshots of the BCR population evolved under various vaccine protocols, and we explore additional aspects of the population dynamics, such as population size and the location of the population peak over time. Notably, these observations indicate that, as discussed below Eq. (12), the optimal center gradually shifts the germline population from its initial position in shape space toward the target while maintaining an optimal distance from it.

Assuming the optimal choice for the vaccine center, we proceed to compare the effectiveness of two different vaccine spread strategies: maintaining the fitness as narrow as possible, or adjusting the fitness spread based on the time-dependent BCH result from Sec. III B. Figure 5 depicts the final bnAb counts as a function of $\sigma_{\min}$. It shows that
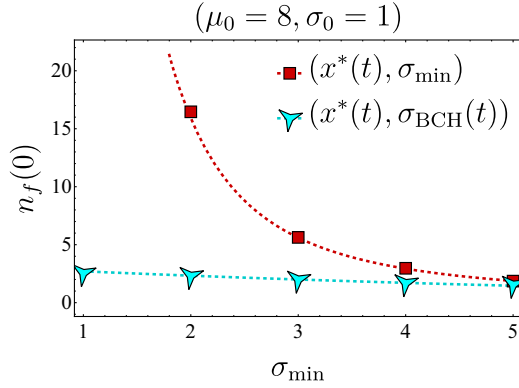
FIG. 5. Final bnAb numbers for a germline centered at $\mu_0 = 8$ and with width $\sigma_0 = 1$ in the one-dimensional shape space. The population is evolved under two different vaccine fitnesses, both optimally centered at $x^*(t)$ but with different spreads. The dashed lines represent results from continuum mean-field numerics, while the solid points depict SSA averages [with the choice of parameters $A_V = 0.08$, $D = 0.05$, $\lambda = 0.001$, and $t_f = 400$]. This plot illustrates that maintaining the vaccine spread at the minimum allowed ($\sigma_{\min}$) yields a larger bnAb count compared to using the time-dependent spread prescribed by the BCH result, (19), and it may indicate that a sequential vaccination with a single antigen type administered at a time would outperform a cocktail regimen.

maintaining the spread at $\sigma_{\min}$ (in red) results in higher final bnAb numbers compared to setting the vaccine spread according to the time-dependent BCH result given in Eq. (19) (in cyan). In other words, when aligning the fitness center with the optimal position defined in Eq. (9), maximizing the focus by keeping the fitness spread narrow results in the largest bnAb numbers ultimately. This computational observation clearly shows that Eq. (19) does not provide the optimal fitness width, likely due to the breakdown of the BCH approximation for narrow profiles, as noted earlier in Sec. III B. Furthermore, drawing from the discussion on the minimum fitness spread in Sec. III B (also see Sec. II A), a plausible interpretation of this finding suggests that sequential immunization with a single antigen type administered at any time—where the fitness remains narrow and focused—outperforms immunization with cocktails of antigens, where the fitness spread is larger. This is consistent with the findings reported in Ref. [25]. Therefore, the implication here is that even a cocktail vaccine that is optimally designed to promote bnAbs, as found in Ref. [26], would be outperformed by a sequential immunization regimen containing appropriately selected single antigen types at each time point, similar to the sequential protocol with gradually changing antigens as reported in Ref. [19].

It is also noteworthy to contrast Fig. 5 with the conclusions drawn in Refs. [22,23], where the administered vaccine antigens were considered sufficiently distinct in their variable residues such that the fittest BCR sequence at every time point corresponded to bnAbs, while the germline started far from the bnAb sequence. In such settings, Ref. [23] found the optimal spread of the vaccine-induced fitness function to be one that is sufficiently wide to reach the least cross-reactive BCRs among the existing B cells; a narrower fitness spread would thus have led to B cell extinction. As time

progresses and mutations widen the BCR distribution along the "breadth" axis, the optimal fitness landscape becomes narrower in subsequent immunizations, as smaller fitness spreads can prevent the extinction. This narrowing of the optimal fitness appears to coincide with the narrowing nature of the BCH spread [Eq. (19)]. However, it is worth emphasizing that this is merely a coincidence. As discussed in the preceding paragraph, the results presented here relax the assumption of distinct antigens used in Refs. [22,23]. Instead, these results pertain to vaccine regimens containing similar antigen types at each time point, corroborating the effectiveness of sequential immunization schemes with gradually changing antigens considered in Refs. [19,31].

V. DISCUSSION

We have explored the optimal vaccination protocol for producing bnAbs through a minimal birth-death-mutation model described by the master equation [Eq. (1)] and its continuum mean-field approximation [Eq. (2)]. Assuming a B cell germline that is initially distant from the desired bnAb state in the shape space, this optimization problem comprises two aspects: first, determining the optimal time-dependent center of the vaccine-induced fitness function, and second, its optimal time-dependent spread. The center of the vaccine fitness function—i.e., where the fitness peaks—is determined by the epitopes shared among the antigen(s) presented at each time point; the spread of the fitness function, on the other hand, is controlled by diversity of the presented epitopes in each shot. We assumed the administered antigens are sufficiently similar to allow for approximating the vaccine-induced fitness landscape with a coarse-grained smooth function whose peak can be distal from the BCR sequence corresponding to bnAbs (Fig. 1). An example of such a scenario is the sequential immunization scheme considered in Refs. [19,31], where the structure of the administered antigens change gradually with subsequent boosts.

We addressed the first part by employing the path integral formulation of the mean-field dynamics [Eq. (7)], and we used a linear coordinate transformation to derive the nontrivial exact relationship given in Eq. (10). This relationship is valid for a general vaccine-induced fitness if the germline population is normally distributed [cf. Eq. (3)]. Equation (10) maps the maximization of the final bnAb count $n_f(0)$ for a germline centered at $\boldsymbol{\mu}_0 \neq 0$ away from the target, into the simpler problem of finding the optimal transformed fitness, defined in Eq. (11), that maximizes the bnAb number for a germline that is instead initially centered at the target ($\boldsymbol{\mu}_0 = 0$) (top panel of Fig. 2). Based on this mapping, we then argued that the optimal vaccine center follows Eq. (12) (see the bottom panel in Fig. 2), with $\boldsymbol{x}^*(t)$ defined in Eq. (9). This result for the optimal fitness center suggests that to obtain the largest numbers of final bnAbs, the epitopes administered in the initial doses should be designed for germline B cells to bind and replicate, while the subsequent doses should contain epitopes that gradually become more similar to the target sequence (i.e., the conserved residues). As noted earlier, this approach mirrors findings in Refs. [19,31], which highlighted the effectiveness of sequential vaccination with gradually changing administered antigens.

The same mapping also allows us to investigate the second part of the bnAb optimization, i.e., the optimal spread of the vaccine-induced fitness landscape. Using a second-order truncation of the celebrated BCH formula, we obtained an approximate form for the mean-field solution [Eq. (16)] that, given a specific form of fitness, can be used to examine how the vaccine spread affects the final bnAb count. In the special case of a Gaussian vaccine-induced fitness [Eq. (17)], maximizing this expression with respect to the vaccine spread led us to Eq. (19). In Eq. (19), we incorporate $\sigma_{\min}$ as a proxy to encompass all factors constraining the narrowness of the fitness function. These factors may include the (weaker) binding of neighboring BCRs to the presented antigen, the effect of discrete similarity bins, and experimental limitations.

Finally, to test the performance of the optimal fitness as obtained from the mean-field model in the discrete stochastic dynamics, we simulated Eq. (1) using the Gillespie SSA and compared the resulting sample averages with the numerical solution of the associated mean-field dynamics. We observed that the SSA averages and mean-field numerics match closely, and the optimal vaccine center, Eq. (12), outperforms other protocols even at the level of the discrete stochastic dynamics (Fig. 4). In addition, we found that keeping the vaccine spread at the minimum allowed seems to be optimal (Fig. 5), suggesting that antigen cocktails reduce the effectiveness of the immunization procedure.

Our findings suggest that optimal vaccination protocols could potentially be designed to maximize the evolution of rare bnAbs without risking B cell extinction, by ensuring that sequentially administered vaccine antigens are not too different at each time point. This is aligned with the experimental studies of such sequential schemes studied in Refs. [19,31]. Additionally, our results provide two insights that may inform future antigen design for sequential vaccination protocols. First, each vaccine dose should contain single antigens rather than antigen cocktails, to create narrower fitness landscapes. Second, antigens in successive doses should not differ significantly; in fact, high-affinity bnAb-targeting antigens should be avoided until later stages, as they could shift the fitness peak far away from the prevailing GC B cell population and potentially lead to B cell apoptosis and extinction. This important conclusion of our work warrants further investigation by addressing several inherent limitations of the model discussed in Sec. II, as well as accounting for continuous entry of naive B cells into germinal centers [1,60]. As briefly discussed in Sec. II, during affinity maturation, B cells interact with each other in various ways, such as competing for limited survival signals from T cells and memory effects including the epitope masking [24,43]. Incorporating these effects in a birth-death-mutation model introduces nonlinear terms, and may well lead to modified perspectives on bnAb evolution, compared to the linear model employed in this study. In future, we intend to explore optimal vaccination strategies using a more realistic version of the stochastic dynamics that accounts for these crucial nonlinearities and feedback loops. Additionally, we plan to extend our analyses to encompass time-dependent vaccine dosages. Ultimately, we hope that our study will motivate further theoretical and experimental investigations by other researchers.

APPENDIX A: DERIVATION OF THE MEAN-FIELD IN THE CONTINUUM LIMIT

In this appendix, we elaborate on the derivation of the mean-field equation (2) from the master equation (1). The average BCR occupation number at bin $\boldsymbol{x}$ and time t is given by $\langle n_{\boldsymbol{x}}(t) \rangle = \sum_{\{n\}} n_{\boldsymbol{x}} P(\{n\}, t)$, where the summation is performed over all possible population configurations $\{n\}$. Note that due to the averaging process, $\langle n_{\boldsymbol{x}}(t) \rangle$ is not necessarily integer-valued anymore, while $\boldsymbol{x}$ remains discrete at this stage. The dynamics of the mean population is obtained by taking a time derivative of $\langle n_{\boldsymbol{x}}(t) \rangle$. Using Eq. (1), it reads

$$\partial_t \langle n_{\boldsymbol{x}}(t) \rangle = \sum_{\{n\}} n_{\boldsymbol{x}} \sum_{\boldsymbol{x}'} V_{\boldsymbol{x}'}(t) \, [\mathbb{E}_{\boldsymbol{x}'}^- - 1](n_{\boldsymbol{x}'} P(\{n\}, t))$$
$$+ \sum_{\{n\}} n_{\boldsymbol{x}} \sum_{\boldsymbol{x}', \boldsymbol{y}'} \gamma_{\boldsymbol{x}'\boldsymbol{y}'} [\mathbb{E}_{\boldsymbol{x}'}^+ \mathbb{E}_{\boldsymbol{y}'}^- - 1](n_{\boldsymbol{x}'} P(\{n\}, t))$$
$$+ \sum_{\{n\}} n_{\boldsymbol{x}'} \sum_{\boldsymbol{x}'} \lambda_{\boldsymbol{x}'} [\mathbb{E}_{\boldsymbol{x}'}^+ - 1](n_{\boldsymbol{x}'} P(\{n\}, t)), \quad \text{(A1)}$$

where we have introduced the operators $\mathbb{E}_{\boldsymbol{x}}^{\pm}$, defined as

$$\mathbb{E}_{\boldsymbol{x}}^{\pm}(n_{\boldsymbol{y}} P(\{n\}, t)) \equiv (n_{\boldsymbol{y}} \pm \delta_{\boldsymbol{x}, \boldsymbol{y}}) P(\{n\} \pm \mathbb{I}_{\boldsymbol{x}}).$$

Equation (A1) can be simplified in each line by an appropriate shift of the variables in the corresponding summations over $\{n\}$. For instance, it is straightforward to show that

$$\sum_{\{n\}} n_{\boldsymbol{x}} \sum_{\boldsymbol{x}'} V_{\boldsymbol{x}'}(t) \, \mathbb{E}_{\boldsymbol{x}'}^-(n_{\boldsymbol{x}'} P(\{n\}, t)) = \quad \text{(A2)}$$

$$\sum_{\{n\}} (n_{\boldsymbol{x}} + \delta_{\boldsymbol{x}, \boldsymbol{x}'}) \sum_{\boldsymbol{x}'} V_{\boldsymbol{x}'}(t) \, n_{\boldsymbol{x}'} P(\{n\}, t), \quad \text{(A3)}$$

where $\delta_{\boldsymbol{x}, \boldsymbol{x}'}$ denotes the Kronecker delta. On implementing these simplifications, Eq. (A2) turns into the following mean-field dynamics:

$$\partial_t \langle n_{\boldsymbol{x}}(t) \rangle = [V_{\boldsymbol{x}}(t) - \lambda_{\boldsymbol{x}}] \langle n_{\boldsymbol{x}}(t) \rangle$$
$$+ \sum_{\boldsymbol{y}} [\gamma_{\boldsymbol{y}\boldsymbol{x}} \langle n_{\boldsymbol{y}}(t) \rangle - \gamma_{\boldsymbol{x}\boldsymbol{y}} \langle n_{\boldsymbol{x}}(t) \rangle], \quad \text{(A4)}$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ denote different affinity bins.

To take the continuum limit of $\boldsymbol{x}$ in Eq. (A4), we make the change

$$\sum_y [\gamma_{yx}\langle n_y\rangle - \gamma_{xy}\langle n_x\rangle] \to$$

$$\int d^d\Delta\boldsymbol{x}\,[\gamma(\boldsymbol{x}-\Delta\boldsymbol{x},\Delta\boldsymbol{x})n(\boldsymbol{x}-\Delta\boldsymbol{x},t) - \gamma(\boldsymbol{x},\Delta\boldsymbol{x})n(\boldsymbol{x},t)],$$

(A5)

where $\gamma(\boldsymbol{x},\Delta\boldsymbol{x})$ denotes the mutation rate for a jump of length $\Delta\boldsymbol{x}$ that starts from $\boldsymbol{x}$. The assumption of localized jumps corresponds to negligible jump rates for large $\Delta\boldsymbol{x}$ (see the discussion in Sec. II), allowing for a Taylor expansion of the integral term in powers of $\Delta\boldsymbol{x}$. Truncating the Taylor expansion at second order then gives

$$\partial_t n(\boldsymbol{x},t) \approx [V(\boldsymbol{x},t) - \lambda(\boldsymbol{x})]\,n(\boldsymbol{x},t)$$
$$+ \nabla^2[D(\boldsymbol{x})\,n(\boldsymbol{x},t)] - \nabla\cdot[\boldsymbol{w}(\boldsymbol{x})\,n(\boldsymbol{x},t)],$$

(A6)

where we have defined the diffusion coefficient as

$$D(\boldsymbol{x}) = \frac{1}{2}\int d^d\Delta\boldsymbol{x}\,(\Delta\boldsymbol{x})^2\,\gamma(\boldsymbol{x},\Delta\boldsymbol{x})$$

(A7)

and the drift velocity as $\boldsymbol{w}(\boldsymbol{x}) = \int d^d\Delta\boldsymbol{x}\,\Delta\boldsymbol{x}\,\gamma(\boldsymbol{x},\Delta\boldsymbol{x})$. For symmetric mutation rates, we have $\boldsymbol{w}(\boldsymbol{x},t) = 0$. Assuming jump rates independent of $\boldsymbol{x}$ then simplifies Eq. (A6) into Eq. (2) of the main text.

APPENDIX B: DETAILS OF THE PATH INTEGRAL CALCULATION

In this appendix, we provide details on deriving Eq. (10) from the path integral formula (7). As discussed in the main text following Eq. (8), our goal here is to examine the optimality of the constant-velocity fitness center. To achieve this, we begin by using appropriate coordinate transformations to express the path integral formula (7) in a new shape-space reference frame. This new frame has an origin that moves linearly in time, with initial position and velocity yet to be determined. Subsequently, we leverage the Gaussian form of the germline distribution, initially centered away from $\boldsymbol{x} = 0$, to appropriately select the initial position and velocity of the new origin. This choice allows us to express the transformed path integral formula similar to its original form, but now for a modified germline centered at the new origin. Through these steps, we will derive the central result in Eq. (10).

Let us begin with expressing Eq. (7) in a new frame, where the origin moves according to $\boldsymbol{z}(t) = \boldsymbol{z}_0 - \boldsymbol{v}t$ (both $\boldsymbol{z}_0$ and $\boldsymbol{v}$ will be determined in the following). This transformation corresponds to adjusting the shape-space vectors as

$$\boldsymbol{x}(t) \to \tilde{\boldsymbol{x}}(t) = \boldsymbol{x}(t) - \boldsymbol{z}(t)$$

and paths as

$$\boldsymbol{r}(t) \to \tilde{\boldsymbol{r}}(t) = \boldsymbol{r}(t) - \boldsymbol{z}(t),$$

and it keeps the path integral measure unchanged [$\mathcal{D}\tilde{\boldsymbol{r}}(t) = \mathcal{D}\boldsymbol{r}(t)$]. In addition, this transformation changes the germline distribution as

$$n_0(\boldsymbol{x}_0) = N_0\mathcal{G}(\boldsymbol{x}_0 - \boldsymbol{\mu}_0, \sigma_0^2)$$
$$\to \tilde{n}_0(\tilde{\boldsymbol{x}}_0) = N_0\mathcal{G}(\tilde{\boldsymbol{x}}_0 - \tilde{\boldsymbol{\mu}}_0, \sigma_0^2),$$

where in the moving frame the germline is initially centered at

$$\tilde{\boldsymbol{\mu}}_0 \equiv \boldsymbol{\mu}_0 - \boldsymbol{z}(t=0) = \boldsymbol{\mu}_0 - \boldsymbol{z}_0.$$

It is easy to see that $\dot{\tilde{\boldsymbol{r}}}(t) = \dot{\boldsymbol{r}}(t) + \boldsymbol{v}$. Upon plugging this in Eq. (8), the path integral action is expressed as

$$S[\boldsymbol{r},\dot{\boldsymbol{r}}] = \frac{\boldsymbol{v}^2 t_f}{4D} - \frac{\boldsymbol{v}\cdot[\tilde{\boldsymbol{r}}(t_f) - \tilde{\boldsymbol{r}}(0)]}{2D}$$
$$+ \int_0^{t_f} du\left(\frac{\dot{\tilde{\boldsymbol{r}}}^2(u)}{4D} - \widetilde{V}(\tilde{\boldsymbol{r}}(u),u)\right)$$

(B1a)

$$\equiv \frac{\boldsymbol{v}^2 t_f}{4D} - \frac{\boldsymbol{v}\cdot[\tilde{\boldsymbol{r}}(t_f) - \tilde{\boldsymbol{r}}(0)]}{2D} + \widetilde{S}[\tilde{\boldsymbol{r}},\dot{\tilde{\boldsymbol{r}}}],$$

(B1b)

where we used the fact that

$$\int_0^{t_f} du\,\dot{\tilde{\boldsymbol{r}}}(u) = \tilde{\boldsymbol{r}}(t_f) - \tilde{\boldsymbol{r}}_0.$$

In Eqs. (B1), $\widetilde{V}(\tilde{\boldsymbol{r}}(u),u)$ and $\widetilde{S}[\tilde{\boldsymbol{r}},\dot{\tilde{\boldsymbol{r}}}]$ denote the transformed vaccine fitness and path integral action in the moving frame; these are simply obtained from $V(\boldsymbol{r}(u),u)$ and $S[\boldsymbol{r},\dot{\boldsymbol{r}}]$ by making the change

$$\boldsymbol{r}(u) \to \tilde{\boldsymbol{r}}(u) + \boldsymbol{z}(u).$$

To obtain the bnAb count, we need to set $\boldsymbol{r}(t_f) = \boldsymbol{x} = 0$ in Eq. (7), which corresponds to $\tilde{\boldsymbol{r}}(t_f) = -\boldsymbol{z}(t_f)$ in the transformed version. Furthermore, using $\tilde{\boldsymbol{x}}_0 = \boldsymbol{x}_0 - \boldsymbol{z}_0$ and $\tilde{\boldsymbol{r}}(0) = \boldsymbol{r}(0) - \boldsymbol{z}_0$, the lower bound of the path integral in Eq. (7) simply changes to $\tilde{\boldsymbol{r}}(0) = \tilde{\boldsymbol{x}}_0$. By substituting these expressions along with Eq. (B1b) into Eq. (7), we finally obtain

$$n_f(0) = \exp\frac{\boldsymbol{v}^2 t_f}{4D} - \frac{\boldsymbol{v}\cdot\boldsymbol{z}(t_f)}{2D}$$
$$\times e^{-\lambda t_f}\int d^d\tilde{\boldsymbol{x}}_0\,\tilde{n}_0(\tilde{\boldsymbol{x}}_0)\,e^{\frac{\boldsymbol{v}\cdot\tilde{\boldsymbol{x}}_0}{2D}}\int_{\tilde{\boldsymbol{r}}(0)=\tilde{\boldsymbol{x}}_0}^{\tilde{\boldsymbol{r}}(t_f)=\boldsymbol{z}(t_f)}\mathcal{D}\tilde{\boldsymbol{r}}(u)\,e^{-\widetilde{S}[\tilde{\boldsymbol{r}},\dot{\tilde{\boldsymbol{r}}}]}.$$

(B2)

Barring the multiplicative constant in the first line, the transformed path integral expression in Eq. (B2) retains the same mathematical structure as the original formula in Eq. (7) for a Gaussian germline. This is because for the germline given by Eq. (3), we can simplify the outer integral by combining the modified germline $\tilde{n}_0(\tilde{\boldsymbol{x}}_0) = N_0\mathcal{G}(\tilde{\boldsymbol{x}}_0 - \tilde{\boldsymbol{\mu}}_0, \sigma_0^2)$ with the exponential factor, $e^{\frac{\boldsymbol{v}\cdot\tilde{\boldsymbol{x}}_0}{2D}}$, through completing the square, which yields

$$\tilde{n}_0(\tilde{\boldsymbol{x}}_0)\,e^{\frac{\boldsymbol{v}\cdot\tilde{\boldsymbol{x}}_0}{2D}} = N_0\mathcal{G}\left(\tilde{\boldsymbol{x}}_0 - \tilde{\boldsymbol{\mu}}_0 + \frac{\boldsymbol{v}\sigma_0^2}{2D}, \sigma_0^2\right)e^{-\frac{\boldsymbol{v}\cdot\tilde{\boldsymbol{\mu}}_0}{2D} + \frac{\boldsymbol{v}^2\sigma_0^2}{8D^2}}.$$

(B3)

We note that the upper bound in the path integral in Eq. (B2), namely, $\tilde{\boldsymbol{r}}(t_f) = \boldsymbol{z}(t_f)$, is different from that in Eq. (7), i.e., $\boldsymbol{r}(t_f) = 0$. We can use the (two degrees of) freedom in choosing the initial position and velocity of the moving frame's origin to eliminate the difference in the upper bounds by imposing $\boldsymbol{z}_0 = \boldsymbol{v}t_f$. Furthermore, by imposing $\tilde{\boldsymbol{\mu}}_0 - \frac{\boldsymbol{v}\sigma_0^2}{2D} = 0$, the transformed germline (B3) will be centered at the origin of the new frame. It is straightforward to show that both conditions are satisfied for

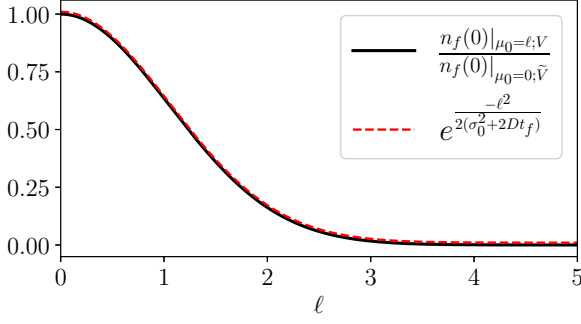$$\boldsymbol{z}_0^* = \boldsymbol{v}^* t_f = \boldsymbol{\mu}_0\frac{2Dt_f}{2Dt_f + \sigma_0^2}.$$

(B4)

FIG. 6. The ratio of $n_f(0)|_{\mu_0=\ell;V}$, obtained from numerically solving Eq. (2) in $d=1$ dimensions for three fitnesses $V(x,t)$ (see below) starting from the germline (3) $n_f(0)|_{\mu_0=0;\tilde{V}}$, to that evaluated from Eq. (2) with the modified fitness $\tilde{V}$ [see Eq. (11)] and starting from the modified initial population $N_0 \mathcal{G}(x_0, \sigma_0^2)$ instead centered on the target. The solid curve is obtained from the numerics for a set of arbitrary fitness functions [$V_{\mathrm{I}} = \max(1-|x|, 0)$, $V_{\mathrm{II}} = \min((x-2+t)^2, 1)$, and $V_{\mathrm{III}} = \mathcal{G}(x-x^*, \sqrt{3D(t_f-t)})$] and their respective $\tilde{V}$.

Substituting these back into Eq. (B2) and using Eq. (B3), we finally arrive at

$$
n_f(0) = \exp -\frac{\boldsymbol{\mu}_0^2}{2(\sigma_0^2 + 2Dt_f)}
$$
$$
\times N_0 e^{-\lambda t_f} \int d^d \boldsymbol{x}_0 \, \mathcal{G}(\boldsymbol{x}_0, \sigma_0^2) \int_{\boldsymbol{r}(0)=\boldsymbol{x}_0}^{\boldsymbol{r}(t_f)=0} \mathcal{D}\boldsymbol{r}(u) \, e^{-\tilde{S}[\boldsymbol{r}, \dot{\boldsymbol{r}}]},
$$
$$
\tag{B5}
$$

where we have reverted the dummy integration variables $\tilde{\boldsymbol{x}}_0$ and $\tilde{\boldsymbol{r}}$ back to $\boldsymbol{x}_0$ and $\boldsymbol{r}$. In Eq. (B5), the second line on the right-hand side can be recognized as the right-hand side of the original expression in Eq. (7) upon changing $V \to \tilde{V}$ and setting $\boldsymbol{\mu}_0 = 0$. Rearranging Eq. (B5) thus leads to Eq. (10) in the main text. Figure 6 shows that the left-hand side of Eq. (10), evaluated by numerically solving the one-dimensional mean-field equation for a set of arbitrary vaccine fitness functions, matches the exponential factor on the right-hand side of Eq. (10).

APPENDIX C: BCH APPROXIMATION FOR DYNAMICS ON DISCRETE BINS

Since the SSA are performed on a system of discrete bins, for completeness, this appendix demonstrates how the calculations in Sec. III B can be extended to the discrete-bin version of the mean-field dynamics.

Recall that the general mean-field dynamics for the discretized bins before taking the continuum limit is given by Eq. (A4). We set $\lambda_{\boldsymbol{x}} = \lambda$ and $\gamma_{\boldsymbol{x}\boldsymbol{y}} = \gamma \sum_{k=1}^d \delta_{\boldsymbol{x}, \boldsymbol{y} \pm \hat{e}_k}$, where $\hat{e}_k$ is the kth unit vector and d denotes the dimension of shape space. Focusing on the one-dimensional case for simplicity, we use the BCH formula (13) with A and B matrices (instead of continuous operators) chosen as $A \to \tilde{V}_{\ell\ell'}(t) = \tilde{V}_\ell(t)\delta_{\ell,\ell'}$ for $\ell \in \{0, 1, 2, \ldots, N_{\mathrm{bin}} - 1\}$, representing the discrete fitness function, and $B \to D_{\ell\ell'} = \gamma(\delta_{\ell,\ell'+1} - 2\delta_{\ell,\ell'} + \delta_{\ell,\ell'+1})$ as the discrete mutation matrix (note that the diagonal terms at $\ell = 1$ and $\ell = N_{\mathrm{bin}}$ in the mutation matrix should be set to $-\gamma$). It is then straightforward to obtain the corre-
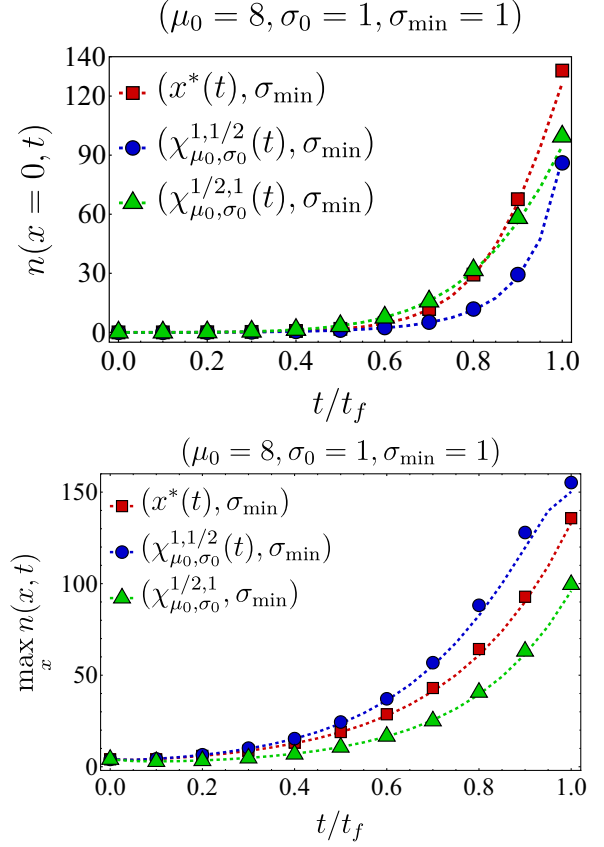

FIG. 7. Plots of bnAb count (top) and the peak size of the B cell population (bottom) as functions of time, utilizing the same parameters as in Fig. 4. The dashed lines represent the numerical solutions to the continuum mean-field equation (2), whereas the solid points depict averages over stochastic realizations of the discrete master equation (1) sampled using SSA.
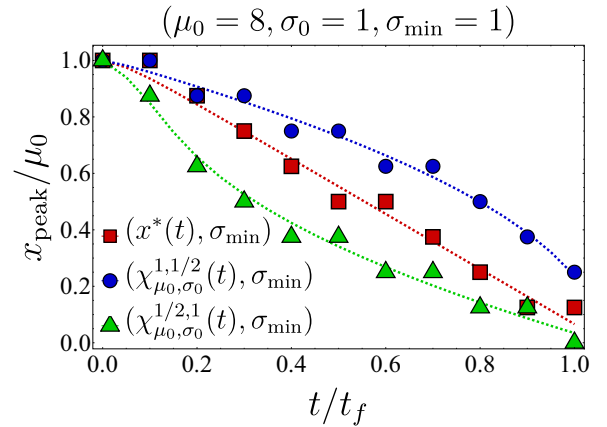

FIG. 8. The location of the population peak in the one-dimensional shape space as a function of time and for three different choices of the vaccine center. Note that in the stochastic simulations, the population peak can only move between the discrete bins, resulting in steplike jumps and periods of no motion in the corresponding data (solid points).
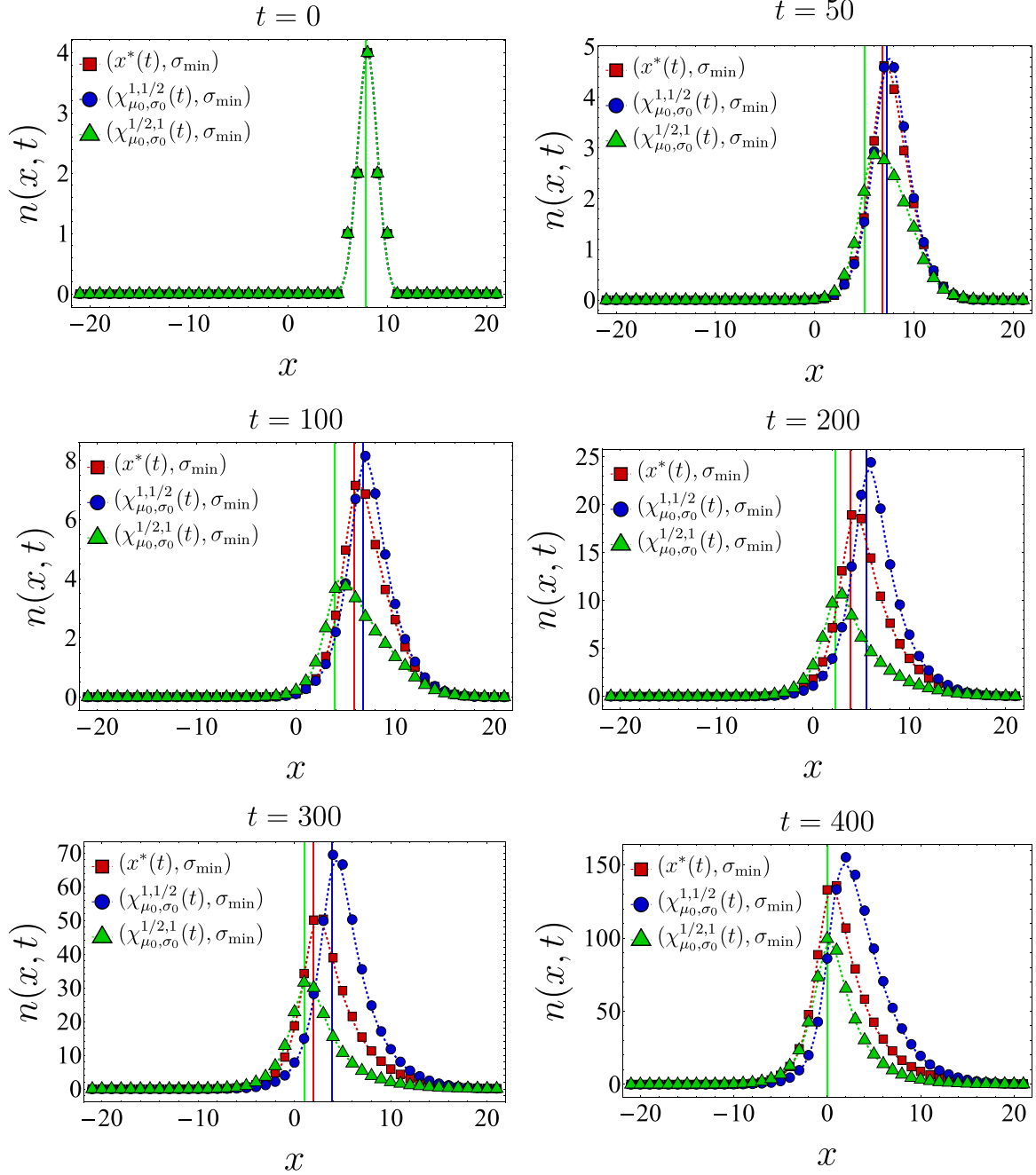
FIG. 9. Snapshots of the B cell population at different times, corresponding to the vaccination protocols of the left panel in Fig. 4. In each case, the dashed curves represent the mean-field numerical solution, while the solid data points depict the SSA sample averages taken over 1000 realizations. The vertical lines in different colors (red, blue, green) indicate the location of the (moving) vaccine center for the corresponding protocol.

sponding commutator matrix as $C_{\ell\ell'} = \gamma(\widetilde{V}_{\ell'} - \widetilde{V}_{\ell})(\delta_{\ell,\ell'+1} + \delta_{\ell,\ell'-1})$. Following through the same steps leading to Eq. (15), one can obtain the BCH approximation for the bnAb count in the discrete-bin model as

$$
\begin{aligned}
n_{\ell=0}(t_f) \approx{} & N_0 e^{-\lambda t_f} e^{\int_0^{t_f} dt \, \widetilde{V}_{\ell=0}(t)} \\
& \times e^{\gamma \int_0^{t_f} dt (t_f - t)(\widetilde{V}_{\ell=1}(t) - 2\widetilde{V}_{\ell=0}(t) + \widetilde{V}_{\ell=-1}(t))} \\
& \times [e^{t_f D_{\ell'\ell''}} n_{\ell''}(t=0)].
\end{aligned} \tag{C1}
$$

For the Gaussian fitness profile, optimizing this expression leads to the following transcendental equation:

$$
\frac{\exp \frac{1}{2\sigma_V^2(t)}}{\frac{1}{\sigma_V^2(t)} - 1} = \frac{2\gamma(t_f - t)}{1 - 2\gamma(t_f - t)}. \tag{C2}
$$

This equation has a unique solution for $2\gamma(t_f - t) \geqslant \frac{1}{1 + 2e^{-3/2}}$ and no solution otherwise, defining a temporal window preceding the final time. It can be readily verified that the optimal width that solves Eq. (C2) decreases over time until it reaches this final temporal window, beyond which the width remains

fixed. Therefore, within the discrete-bin dynamics, the optimal fitness spread automatically reaches a finite minimum value, preventing the fitness amplitude from becoming singular, as discussed in Sec. III B.

APPENDIX D: TIME EVOLUTION OF B CELL POPULATION UNDER DIFFERENT VACCINE PROTOCOLS

Here, we present further computational results on the time evolution of the B cell population in the one-dimensional shape space. Figure 7 shows the bnAb count (top) and population peak's size (bottom) as functions of time. These results correspond to the same vaccination protocols and parameters as the left panel of Fig. 4. The graphs suggest that although positioning the vaccine-induced fitness peak optimally leads to the highest final bnAb count, the resulting population peak remains smaller compared to that achieved with the $\chi_{\mu_0,\sigma_0}^{1,1/2}(t)$ protocol. This observation is reinforced by the population

snapshots in Fig. 8, indicating that while the $x^*(t)$ protocol effectively maximizes $n_f(0)$, it may not fully optimize other aspects of the final B cell population, such as its total size or peak.

In Fig. 9, we present the location of the population's peak as a function of time for the same vaccination protocols. Notably, only for the $(\chi_{\mu_0,\sigma_0}^{1/2,1}, \sigma_{\min})$ protocol does the population peak appear to reach the target bin, whereas for the other two protocols the peak remains away from $x = 0$ at the final time. For the optimal vaccine center, and after an initial transient period, the population peak consistently trails $x^*(t)$ with a constant lag. Conversely, with other choices, the population's distance from the vaccine center fluctuates over time. As discussed following Eq. (12), such fluctuations in distance may either result in a slow movement of the population towards the target when the vaccine center remains too close to the population peak (blue), or lead to limited population growth when the vaccine center moves too far from the population (green).

[1] G. D. Victora and M. C. Nussenzweig, Germinal centers, Annu. Rev. Immunol. **40**, 413 (2022).

[2] G. Altan-Bonnet, T. Mora, and A. M. Walczak, Quantitative immunology for physicists, Phys. Rep. **849**, 1 (2020).

[3] R. Sender, Y. Weiss, Y. Navon, I. Milo, N. Azulay, L. Keren, S. Fuchs, D. Ben-Zvi, E. Noor, and R. Milo, The total mass, number, and distribution of immune cells in the human body, Proc. Natl. Acad. Sci. USA **120**, e2308511120 (2023).

[4] G. Yaari and S. H. Kleinstein, Practical guidelines for B-cell receptor repertoire sequencing analysis, Genome Med. **7**, 121 (2015).

[5] L. Mesin, J. Ersching, and G. D. Victora, Germinal center B cell dynamics, Immunity **45**, 471 (2016).

[6] G. D. Victora and M. C. Nussenzweig, Germinal centers, Annu. Rev. Immunol. **30**, 429 (2012).

[7] A. K. Abbas, A. H. Lichtman, and S. Pillai, *Basic Immunology: Functions and Disorders of the Immune System* (Elsevier Health Sciences, Philadelphia, 2019).

[8] A. S. Clem, Fundamentals of vaccine immunology, J. Glob. Infect. **3**, 73 (2011).

[9] R. M. Zinkernagel, On natural and artificial vaccinations, Annu. Rev. Immunol. **21**, 515 (2003).

[10] M. Thali, J. Moore, C. Furman, M. Charles, D. Ho, J. Robinson, and J. Sodroski, Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding, J. Virol. **67**, 3978 (1993).

[11] M. Sangesland and D. Lingwood, Antibody focusing to conserved sites of vulnerability: the immunological pathways for "universal" influenza vaccines, Vaccines **9**, 125 (2021).

[12] R. Wyatt and J. Sodroski, The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens, Science **280**, 1884 (1998).

[13] S. Zolla-Pazner, Identifying epitopes of HIV-1 that induce protective antibodies, Nat. Rev. Immunol. **4**, 199 (2004).

[14] D. R. Burton and L. Hangartner, Broadly neutralizing antibodies to HIV and their role in vaccine design, Annu. Rev. Immunol. **34**, 635 (2016).

[15] P. Hraber, M. S. Seaman, R. T. Bailer, J. R. Mascola, D. C. Montefiori, and B. T. Korber, Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection, AIDS (London, England) **28**, 163 (2014).

[16] N. S. Laursen and I. A. Wilson, Broadly neutralizing antibodies against influenza viruses, Antivir. Res. **98**, 476 (2013).

[17] C. Kreer, C. Lupo, M. S. Ercanoglu, L. Gieselmann, N. Spisak, J. Grossbach, M. Schlotz, P. Schommers, H. Gruell, L. Dold *et al.*, Probabilities of developing HIV-1 bNAb sequence features in uninfected and chronically infected individuals, Nat. Commun. **14**, 7137 (2023).

[18] D. Sok and D. R. Burton, Recent progress in broadly neutralizing antibodies to HIV, Nat. Immunol. **19**, 1179 (2018).

[19] A. Escolano, J. M. Steichen, P. Dosenovic, D. W. Kulp, J. Golijanin, D. Sok, N. T. Freund, A. D. Gitlin, T. Oliveira, T. Araki *et al.*, Sequential immunization elicits broadly neutralizing anti-HIV-1 antibodies in Ig knockin mice, Cell **166**, 1445 (2016).

[20] K. E. Stephenson, K. Wagh, B. Korber, and D. H. Barouch, Vaccines and broadly neutralizing antibodies for HIV-1 prevention, Annu. Rev. Immunol. **38**, 673 (2020).

[21] A. Amitai, L. Mesin, G. D. Victora, M. Kardar, and A. K. Chakraborty, A population dynamics model for clonal diversity in a germinal center, Front. Microbiol. **8**, 1693 (2017).

[22] K. G. Sprenger, J. E. Louveau, P. M. Murugan, and A. K. Chakraborty, Optimizing immunization protocols to elicit broadly neutralizing antibodies, Proc. Natl. Acad. Sci. USA **117**, 20077 (2020).

[23] R. S. Ganti and A. K. Chakraborty, Mechanisms underlying vaccination protocols that may optimally elicit broadly neutralizing antibodies against highly mutable pathogens, Phys. Rev. E **103**, 052408 (2021).

[24] A. K. Chakraborty, A perspective on the role of computational models in immunology, Annu. Rev. Immunol. **35**, 403 (2017).

[25] S. Wang, J. Mata-Fink, B. Kriegsman, M. Hanson, D. J. Irvine, H. N. Eisen, D. R. Burton, K. D. Wittrup, M. Kardar, and A. K. Chakraborty, Manipulating the selection forces during affinity

maturation to generate cross-reactive HIV antibodies, Cell **160**, 785 (2015).

[26] J. S. Shaffer, P. L. Moore, M. Kardar, and A. K. Chakraborty, Optimal immunization cocktails can promote induction of broadly neutralizing Abs against highly mutable pathogens, Proc. Natl. Acad. Sci. USA **113**, E7039 (2016).

[27] J. G. Jardine, D. Sok, J.-P. Julien, B. Briney, A. Sarkar, C.-H. Liang, E. A. Scherer, C. J. Henry Dunand, Y. Adachi, D. Diwanji *et al.*, Minimally mutated HIV-1 broadly neutralizing antibodies to guide reductionist vaccine design, PLoS Pathog. **12**, e1005815 (2016).

[28] J. G. Jardine, T. Ota, D. Sok, M. Pauthner, D. W. Kulp, O. Kalyuzhniy, P. D. Skog, T. C. Thinnes, D. Bhullar, B. Briney *et al.*, Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen, Science **349**, 156 (2015).

[29] A. Amitai, M. Sangesland, R. M. Barnes, D. Rohrer, N. Lonberg, D. Lingwood, and A. K. Chakraborty, Defining and manipulating B cell immunodominance hierarchies to elicit broadly neutralizing antibody responses against influenza virus, Cell Syst. **11**, 573 (2020).

[30] W. Dong, Y. Bhide, F. Sicca, T. Meijerhof, K. Guilfoyle, O. G. Engelhardt, L. Boon, C. A. De Haan, G. Carnell, N. Temperton *et al.*, Cross-protective immune responses induced by sequential influenza virus infection and by sequential vaccination with inactivated influenza vaccines, Front. Immunol. **9**, 2312 (2018).

[31] J. M. Steichen, D. W. Kulp, T. Tokatlian, A. Escolano, P. Dosenovic, R. L. Stanfield, L. E. McCoy, G. Ozorowski, X. Hu, O. Kalyuzhniy *et al.*, Hiv vaccine design to target germline precursors of glycan-dependent broadly neutralizing antibodies, Immunity **45**, 483 (2016).

[32] A. S. Perelson and G. F. Oster, Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination, J. Theor. Biol. **81**, 645 (1979).

[33] D. J. Smith, S. Forrest, R. R. Hightower, and A. S. Perelson, Deriving shape space parameters from immunological data, J. Theor. Biol. **189**, 141 (1997).

[34] S. Luo and A. S. Perelson, Competitive exclusion by autologous antibodies can prevent broad HIV-1 antibodies from arising, Proc. Natl. Acad. Sci. USA **112**, 11654 (2015).

[35] A. S. Perelson and G. Weisbuch, Immunology for physicists, Rev. Mod. Phys. **69**, 1219 (1997).

[36] P. L. Yang and P. G. Schultz, Mutational analysis of the affinity maturation of antibody 48g7, J. Mol. Biol. **294**, 1191 (1999).

[37] Y. Liu, D. Joshua, G. Williams, C. Smith, J. Gordon, and I. MacLennan, Mechanism of antigen-driven selection in germinal centres, Nature (London) **342**, 929 (1989).

[38] P. Grassberger and M. Scheunert, Fock-space methods for identical classical objects, Fortschr. Phys. **28**, 547 (1980).

[39] U. C. Täuber, *Critical Dynamics: A Field Theory Approach to Equilibrium and Non-equilibrium Scaling Behavior* (Cambridge University Press, UK, 2014).

[40] A. Escolano, H. B. Gristick, R. Gautam, A. T. DeLaitsch, M. E. Abernathy, Z. Yang, H. Wang, M. A. Hoffmann, Y. Nishimura, Z. Wang *et al.*, Sequential immunization of macaques elicits heterologous neutralizing antibodies targeting the V3-glycan patch of HIV-1 Env, Sci. Transl. Med. **13**, eabk1533 (2021).

[41] R. P. Feynman, A. R. Hibbs, and D. F. Styer, *Quantum Mechanics and Path Integrals* (Dover Publications, New York, 2010).

[42] J. M. Tas, J.-H. Koo, Y.-C. Lin, Z. Xie, J. M. Steichen, A. M. Jackson, B. M. Hauser, X. Wang, C. A. Cottrell, J. L. Torres *et al.*, Antibodies from primary humoral responses modulate the recruitment of naive B cells during secondary responses, Immunity **55**, 1856 (2022).

[43] L. Yang, M. Van Beek, Z. Wang, F. Muecksch, M. Canis, T. Hatziioannou, P. D. Bieniasz, M. C. Nussenzweig, and A. K. Chakraborty, Antigen presentation dynamics shape the antibody response to variants like SARS-CoV-2 Omicron after multiple vaccinations with the original strain, Cell Rep. **42**, 112256 (2023).

[44] D. Schaefer-Babajew, Z. Wang, F. Muecksch, A. Cho, M. Loewe, M. Cipolla, R. Raspe, B. Johnson, M. Canis, J. DaSilva *et al.*, Antibody feedback regulates immune memory after SARS-CoV-2 mRNA vaccination, Nature (London) **613**, 735 (2023).

[45] J. Zhang and E. I. Shakhnovich, Optimality of mutation and selection in germinal centers, PLoS Comput. Biol. **6**, e1000800 (2010).

[46] M. Van Beek, M. C. Nussenzweig, and A. K. Chakraborty, Two complementary features of humoral immune memory confer protection against the same or variant antigens, Proc. Natl. Acad. Sci. USA **119**, e2205598119 (2022).

[47] S. Hassani, *Mathematical Physics: A Modern Introduction to Its Foundations* (Springer International Publishing, Cham, 2013).

[48] M. Suzuki, On the convergence of exponential operators—the Zassenhaus formula, BCH formula and systematic approximants, Commun. Math. Phys. **57**, 193 (1977).

[49] S. Lloyd, Universal quantum simulators, Science **273**, 1073 (1996).

[50] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter error with commutator scaling, Phys. Rev. X **11**, 011020 (2021).

[51] R. Glowinski, S. J. Osher, and W. Yin, *Splitting Methods in Communication, Imaging, Science, and Engineering* (Springer International Publishing, Cham, 2017).

[52] M. Kardar, *Statistical Physics of Fields* (Cambridge University Press, UK, 2007).

[53] H. H. Tam, M. B. Melo, M. Kang, J. M. Pelet, V. M. Ruda, M. H. Foley, J. K. Hu, S. Kumari, J. Crampton, and A. D. Baldeon *et al.*, Sustained antigen availability during germinal center initiation enhances antibody responses to vaccination, Proc. Natl. Acad. Sci. USA **113**, E6639 (2016).

[54] K. M. Cirelli, D. G. Carnathan, B. Nogal, J. T. Martin, O. L. Rodriguez, A. A. Upadhyay, C. A. Enemuo, E. H. Gebru, Y. Choe, and F. Viviano *et al.*, Slow delivery immunization enhances HIV neutralizing antibody and germinal center responses via modulation of immunodominance, Cell **177**, 1153 (2019).

[55] S. H. Bhagchandani, L. Yang, J. H. Lam, L. Maiorino, E. Ben-Akiva, K. A. Rodrigues, A. Romanov, H. Suh, A. Aung, S. Wu, A. Wadhera, A. K. Chakraborty, and D. J. Irvine, Two-dose "extended priming" immunization amplifies humoral immune responses by synchronizing vaccine delivery with the germinal center response, Sci. Immunol. **9**, eadl3755 (2024).

[56] O. H. Schnaack and A. Nourmohammad, Optimal evolutionary decision-making to store immune memory, Elife **10**, e61346 (2021).

[57] D. T. Gillespie, Stochastic simulation of chemical kinetics, Annu. Rev. Phys. Chem. **58**, 35 (2007).

[58] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. **81**, 2340 (1977).

[59] The code used for the simulation can be found online at https://github.com/saeedmsol/Gillespie-stochastic-simulation-algorithm.

[60] T. Hägglöf, M. Cipolla, M. Loewe, S. T. Chen, L. Mesin, H. Hartweger, M. A. ElTanbouly, A. Cho, A. Gazumyan, V. Ramos *et al.*, Continuous germinal center invasion contributes to the diversity of the immune response, Cell **186**, 147 (2023).