Modeling the impacts of a simple meiotic gene drive on small, homeostatic populations

Kevin R. Pilkiewicz and Michael L. Mayo
U.S. Army Engineer Research and Development Center, Vicksburg, Mississippi 39180, USA

(Received 29 July 2019; accepted 16 January 2020; published 18 February 2020)

Gene drives offer unprecedented control over the fate of natural ecosystems by leveraging non-Mendelian inheritance mechanisms to proliferate synthetic genes across wild populations. However, these benefits are offset by a need to avoid the potentially disastrous consequences of unintended ecological interactions. The efficacy of many gene-editing drives has been brought into question due to predictions that they will inevitably be thwarted by the emergence of drive-resistant mutations, but these predictions derive largely from models of large or infinite populations that cannot be driven to extinction faster than mutations can fixate. To address this issue, we characterize the impact of a simple, meiotic gene drive on a small, homeostatic population whose genotypic composition may vary due to the stochasticity inherent in natural mating events (e.g., partner choice, number of offspring) or the genetic inheritance process (e.g., mutation rate, gene drive fitness). To determine whether the ultimate genotypic fate of such a population is sensitive to such stochastic fluctuations, we compare the results of two dynamical models: a deterministic model that attempts to predict how the genetics of an average population evolve over successive generations, and an agent-based model that examines how stable these predictions are to fluctuations. We find that, even on average, our stochastic model makes qualitatively distinct predictions from those of the deterministic model, and we identify the source of these discrepancies as a dynamic instability that arises at short times, when genetic diversity is maximized as a consequence of the gene drive's rapid proliferation. While we ultimately conclude that extinction can only beat out the fixation of drive-resistant mutations over a limited region of parameter space, the reason for this is more complex than previously understood, which could open new avenues for engineered gene drives to circumvent this weakness.

DOI: 10.1103/PhysRevE.101.022412

I. INTRODUCTION

It has often been said that "history is written by the victors," and nowhere is that more true than in the state of nature itself, where natural selection drives unfavorable genetic alleles and the organisms that possess them to extinction, leaving behind only the genotypic winners to inscribe the chronicle of their species' survival—their genome—in the universal language of paired nucleic acids. Gene drive technologies have armed synthetic biologists with the capability of subverting this paradigm by enabling alleles with a significant phenotypic fitness cost to selectively propagate across a population [1]. Although many mechanisms of gene drive action have been reported [2], most involve insertion of a "drive" allele on one diploid chromosome that can trigger a wild-type allele on the corresponding gene of the paired chromosome to be excised and replaced with a copy of the drive. Whenever this cut-and-replace mechanism succeeds, the drive is preferentially passed down to all offspring instead of half, as expected with Mendelian inheritance.

New gene-editing methods, such as those that appropriate bacterial DNA sequences known as clustered regularly interspaced short palindromic repeats (CRISPR) [3,4], now permit mankind to choose which organisms live and which die, a power fraught with ethical quandaries [5–7]. Synthetic gene drives have been hypothesized as a means for annihilating pests and controlling the spread of invasive species [8–11]; but the release of genetically modified organisms into the wild is, at present, an irreversible action that could have dire ecological and environmental consequences, especially if the drive were to escape beyond the targeted population or species. Some progress has been made in addressing these risks [12,13], but as long as such dangers exist, predictive modeling will remain the principal tool for assessing the impact of a gene drive release on the long-term fate of wild animal populations.

Although many mathematical and computational models have already been developed to study the spread of gene drives, they are generally insufficient for understanding how the changing distribution of alleles within a population of organisms will impact the size of the group over time. The most common modeling approach, which we also adopt, has been to study the generational dynamics of genotypic frequencies within a well-mixed population whose genome has been reduced to only those genes relevant to the action of the drive. Most such models implicitly assume that populations are either fixed [14,15] or formally infinite [11,16], but both assumptions preclude the possibility of the population being pushed to extinction by a drive with a high fitness cost.

Furthermore, a constantly large or infinite population all but guarantees the emergence and eventual fixation of mutated alleles with no inherent fitness cost that are immune to the action of the drive. This is despite the fact that the processes producing such alleles, such as the nonhomologous end-joining of cut DNA strands, can be extremely rare events that may not even emerge in a small to moderate population before the gene drive itself can either fixate or extinguish the population entirely. Other models that account for variable

populations have been reported, but they model population fluctuations in terms that are specific to a single species of interest [17,18].

We address these shortcomings with a deterministic model that quantifies the expected response of a finite, homeostatic animal population to the introduction of a simple, meiotic gene drive [19,20]. Note that although "meiotic drive" is a general term for any mechanism that manipulates meiosis to enhance the likelihood of one allele being inherited over another [21–24], we use the variation "meiotic gene drive" to restrict our considerations to synthetic, gene-editing constructs like those based on CRISPR. To model the population as homeostatic, we set its death rate proportional to its size, which leads to stable population dynamics that always tend toward a steady state in which deaths balance births. We also employ integer rounding to ensure that mutations do not inevitably fixate in the population through the gradual and unphysical accumulation of fractional animals. Using this model, we demonstrate how the presumed fitness cost of the drive allele and the frequency of drive-resistant mutations both influence the long-time fate of the population, resulting in a broad range of allelic outcomes, including population extinction.

For a small to moderate-sized population, long-time genotypic outcomes may be sensitive to stochastic fluctuations in mating and inheritance rates, so we test the robustness of our deterministic model's predictions by simulating the generational dynamics of an agent-based population model whose behavior, averaged over many replicate trajectories, is designed to agree with the deterministic results when the dynamics are stable to fluctuations. While this equivalence holds reasonably well in the absence of mutations, it is generally destroyed by even a modest nonzero mutation rate, and the fascinating diversity in long-term outcomes predicted by the deterministic model is lost as well. Instead, the agent-based model predicts the long-time fixation of drive-resistant mutations across the vast majority of the relevant parameter space, which we trace to a tipping point that emerges in our populations at short times due to the interplay between population and inheritance dynamics. Near this tipping point, rare events like mutations become more likely due to the genetically diverse composition of the population. Interestingly, the occurrence of only one or two individuals with mutated alleles can be sufficient to tip the dynamics towards a disparate genetic fate.

II. MODELS

Our deterministic and stochastic models both initialize with a population of $N(0)$ diploid, sexually reproducing organisms whose genotypes are characterized by two relevant genes. The first is a sex gene, which we assume has an X and Y allele so that an organism possessing an XY pair is male and one possessing an XX pair is female. Since mating can only occur between a male and female organism, standard Mendelian inheritance guarantees that, on average, half of all offspring will be male and the other half female. The second relevant gene is the one targeted by our hypothetical gene drive. This gene is assumed to have only a single wild-type allele, which we denote W, and an engineered drive variant D.

We assume the drive can only activate during the process of meiosis, when haploid gametes are produced. If an organism possesses one wild-type allele and one drive allele, then there is a homing probability h that the drive triggers the cutting of the wild-type DNA strand and attempts to replace the W allele with another D allele. When this occurs, there is a second probability m that the cut DNA is repaired improperly through nonhomologous end joining, resulting in the creation of a mutated allele that is immune to further action of the gene drive. For simplicity, we label all such mutated alleles as M. As a result of these considerations, the gamete of a wild-type-drive heterozygote will possess the W allele with probability $(1-h)/2$, the M allele with probability $hm/2$, and the D allele with probability $[1+h-hm]/2$. In deriving these probabilities, we have neglected the possibility of segregation distortion [25,26]; any deviations from Mendelian inheritance are due solely to the action of the gene drive.

For each sex, $s \in \{XY, XX\}$, there are thus six relevant genotypes, $g \in \{WW, DD, MM, WD, WM, DM\}$, whose inheritance dynamics must be considered. Only the gametes of the WD genotype will have a non-Mendelian allelic distribution. We define $n_{s,g}(t)$ to be the number of organisms after t generations that have sex s and genotype g, and we define $N(t)$ as the total number of organisms after t generations, with the requirement that

$$N(t) = \sum_{s,g} n_{s,g}(t). \tag{1}$$

In both models, the population dynamics obey the following general master equation:

$$n_{s,g}(t+1) = n_{s,g}(t) + b_{s,g}(t) - d_{s,g}(t), \tag{2}$$

where $b_{s,g}(t)$ is the number of new organisms with sex s and genotype g born during the t^{th} generation and $d_{s,g}(t)$ is the number that die. The differences between our analytic and computational approach lie principally in how these two quantities are modeled.

A. Deterministic model

Our analyic model assumes that the inheritance dynamics of a homeostatically stable population consist of small, stochastic fluctuations about some steady state. When a small initial fraction of gene-drive-inoculated organisms is introduced into this wild population, the size and genotypic makeup of the population will shift towards a new steady state, and we assume that this transition occurs relatively predictably, with stochasticity accounting only for minor fluctuations about some average trajectory that our model will attempt to characterize. Because the gene drive under consideration impacts both sexes equally, and because we have assumed Mendelian sex inheritance, we will further assume that $n_{XX,g}(t) = n_{XY,g}(t) \equiv n_g(t)/2$ for all g and t. Even if an initial sex imbalance existed, the former two considerations ensure its transience in the population. The number of births per generation in our analytic model can thus be expressed independently of sex as

$$b_g(t) = \sum_{g_1,g_2} 2\left\lfloor \frac{\omega_B f_{g_1} f_{g_2} p(g|g_1,g_2) n_{g_1}(t) n_{g_2}(t)}{4N(t)} \right\rfloor. \tag{3}$$

The meaning of each quantity to the right of the equal sign will be considered in turn.

First, the parameter ω_B represents the average number of organisms that a female births each generation. This parameter can be thought of as the average number of times a female mates in a generation multiplied by the average number of offspring produced per coupling. We shall assume that every newly born organism is sexually active in the generation following its birth, subsuming the maturation time scale of an organism in our definition of generation.

The parameter $f_g \in [0, 1]$ is the fitness of an organism with genotype g. The lower the fitness coefficient, the more significant the phenotypic penalty the allele g inflicts on an organism's ability to reproduce. For simplicity, we assume that only organisms that are homozygous in the drive allele suffer such a penalty, so $f_g = 1$ unless $g = DD$. The number of organisms birthed by couples with one drive homozygous parent is thus reduced by a factor of $f_{DD} < 1$, and that produced by two such parents is reduced by a factor of f_{DD}^2.

The function $p(g|g_1, g_2)$ is the conditional probability that an offspring of genotype g is born, given it had parents of genotypes g_1 and g_2. As an example, consider the probability $p(DM|WD, WD)$. The probability of one parent passing down a drive allele while the other passes down a mutated allele is $hm[1 + h - hm]/4$. Since both parents can pass down either allele in this case, the total probability of two wild-type-drive heterozygotes giving birth to an offspring with genotype DM is double this: $p(DM|WD, WD) = hm[1 + h - hm]/2$.

The remaining factors in Eq. (3) can be accounted for as follows. First, we assume that each female chooses a single mate each generation and that her choice is equally likely to be any male, i.e., the population is generationally monogamous and panmictic. Since we have already assumed a sexually symmetric population, this makes the probability of a female with genotype g_1 choosing a male with genotype g_2 during generation t equal to $n_{g_2}(t)/N(t)$. Since there are $n_{g_1}(t)/2$ such females in that generation, the expected number of pairings between organisms with genotypes g_1 and g_2 is just the product $[n_{g_1}(t)n_{g_2}(t)]/[2N(t)]$. Note that which sex we associate with g_1 is arbitrary, since both g_1 and g_2 are summed over.

The floor function is taken so that no fractional organisms are born and the accumulation of unlikely genotypes (typically those with mutant alleles) is not made inevitable by the rounding up of fractional values less than unity. If we did not perform this rounding, it would be impossible for the model to account for situations in which the drive allele fixates faster than mutations can emerge. We divide by two inside the floor brackets and then multiply by two outside to ensure that the population always consists of an even number of organisms, as required by our assumption of sex symmetry. The total summand is thus the average number of organisms with genotype g born to mothers with genotype g_1 and fathers with genotype g_2, or vice versa.

The organism death rate $d_g(t)$ is constrained by our requirements that the population be homeostatic but not perfectly resilient. In other words, the population should resist changes to its stable population, but a sufficiently large drop in organism fitness should still be capable of driving it to extinction. This implies that the initial state of the population, in the absence of the drive, should be a stable steady state, i.e., if the population is initialized as entirely wild type, then $N(t) = N(0)\ \forall t \geqslant 0$.

The simplest way of modeling death that is consistent with this requirement would be to assume that each organism has some Poisson-distributed likelihood to die during any generation t, implying a linear death rate $d_g(t) = \zeta n_g(t)$. This is inadequate, however, because our initial steady-state assumption would force us to choose $\zeta = \omega_B/2$, precluding us from using ζ to fit empirical population data and, more importantly, causing any population with fitness less than unity to rapidly go extinct.

An alternative death model would be that of a logistic population model, $d_g(t) = \alpha(N(t)/K)n_g(t)$, where K is the carrying capacity of the local environment. Again, however, our steady-state assumption would require that $K = N(0)$ and $\alpha = \omega_B/2$, depriving us once more of a free fitting parameter and resulting in a population that can be shown to *never* go extinct, except in the extreme case of zero fitness.

The Poisson death model leads to populations that are not sufficiently homeostatic, but the logistic death model produces populations that are too resilient; we find that a linear combination of the two allows us to strike a balance between the two extremes that even leaves us a free parameter:

$$d_g(t) = \left\lceil \left[\zeta + \left(\frac{1}{2}\omega_B - \zeta \right) \frac{N(t)}{N(0)} \right] n_g(t) \right\rceil. \tag{4}$$

Note that whereas we rounded births down, the ceiling function brackets indicate that we round the number of deaths up. This seems like an arbitrary and inconsistent choice, but it is simply a quantification of the physically sensible notion that a partial animal is a dead animal, i.e., neither a partially born animal nor a partially dead animal should be counted as viable members of the population in subsequent generations. Furthermore, since we want a model that can account for the possibility of population extinction, rounding deaths up is necessary to prevent a small population of two to four organisms from surviving indefinitely due to the suppressed death rate at population levels well below the carrying capacity. Also note that although ζ is not a fixed parameter, the steady state of a wild-type population will only be stable to fluctuations if $\zeta \leqslant \omega_B/2$.

B. Stochastic model

The objective of our agent-based model is to determine whether stochastic fluctuations, which we assumed were small and could be averaged out in our deterministic model, can in fact exert any influence over the genetic fate of a population targeted by a meiotic gene drive. This model is initialized as follows. At $t = 0$, $N(0)$ agents are uniformly distributed throughout a two-dimensional box of area L^2, upon which we impose periodic boundaries that approximate bulk conditions. Finally, each agent is assigned a genotype from the same set of possibilities considered in the deterministic model.

At every time step, each female randomly chooses a mating partner, with equal probability, from all males lying within some characteristic interaction radius of her position. (A female does not mate during any generation in which a sufficiently proximate male cannot be found.) If the genotype of the female is g_1, and the genotype of her selected partner is g_2,

then the coupling produces ω_B-many offspring, each of which is assigned a sex at random and a genotype g with probability $p(g|g_1, g_2)$, which is defined the same as in Eq. (3). The total number of these new agents is labeled by $b_{f=1}(t)$, which is the number of agents that would be born during generation t if all genotypes were of unit fitness. A uniformly distributed random number on the interval [0,1] is then chosen for each new agent, and the agent is added to a random location within the simulation box only if this number is less than or equal to $f_{g_1} f_{g_2}$, which is the fitness product of its parents' genotypes.

After births have been tabulated, a number of agents equal to

$$\left\lceil \left[\frac{2\zeta}{\omega_B} + \left(1 - \frac{2\zeta}{\omega_B}\right) \frac{N(t)}{N(0)} \right] b_{f=1}(t) \right\rceil \tag{5}$$

are selected at random from the existing adult population and removed from the box; newly created agents are eligible for selection only if all existing agents have already been removed. A slight modification to Eq. (4) is required for the stochastic model in order for the dynamics of a purely wild-type population to consist of small fluctuations about a stable steady state of $N(0)$ organisms. This is because at steady state, Eq. (4) fixes the total number of deaths to a value of $\omega_B N(0)/2$; but in the stochastic model, this only equals the number of births *on average*. If this equation were used for the stochastic model, the stable steady-state population that the system relaxes towards would vary dynamically as a function of the stochastically fluctuating gender ratio, since the number of females controls the number of mating events. While the overall average population level would still tend towards $N(0)$, as desired, the fluctuations about this state would be significant. The modifications made in Eq. (5) resolve this issue, and the total number of deaths in the deterministic model can be recovered by setting $b_{f=1}(t)$ equal to its average value of $\omega_B N(t)/2$.

After births and deaths have been taken into account, the remaining agents are allowed to move around the box according to some predefined set of rules. Simple possibilities include translating each agent a fixed distance in a random direction or allowing each agent to perform a fixed-length random walk with a specified step size. This makes our computational model capable of studying density-dependent effects on inheritance dynamics, although our goal here is to establish a correspondence between this model and our deterministic alternative. As such, we set the interaction radius to exceed a length of approximately $L/\sqrt{2}$ (half the simulation box diagonal), in which case the population will be panmictic regardless of the motility model used. After new positions are chosen for the agents, each female chooses a new mate within her interaction radius, and the algorithm repeats.

The dynamics of the model are considered to be stable to stochastic fluctuations if the trajectories of many replicate simulations all cluster about a single, sharply defined mean, thereby implying that stochastic variations in the genotypic subpopulations at most lead to small changes in the long-term genotypic fate of the population. An unstable system, by contrast, would be akin to a chaotic mechanical system in physics, wherein small variations to the trajectory early on are able to propagate into large variations at long times. If the simulated panmictic population dynamics of the agent-based model are indeed stable in this sense, then the replicate-averaged trajectory should, by construction, be very close to the result of the deterministic model, whose averaging is predicated on an assumption that the stochasticity of the system consists only of small, random fluctuations about some mean trajectory. When the mutation rate is sufficiently small for the population size, this close correspondence between the models seems to hold, as illustrated in Fig. 1. Both models predict the same long-time genotypic fate for the population, although the stochastic model, on average, seems to relax towards steady state a little faster than the deterministic model. This difference likely derives from the rounding scheme used in the deterministic model, which will always round a partial birth down, even in cases where the extra whole animal might be more likely to be born than not. For larger mutation rates, the dynamics can become unstable to stochastic fluctuations, resulting in the two models making disparate long-time predictions. We document and ultimately account for these differences in the proceeding section.

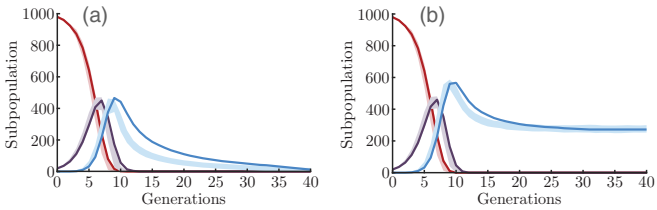

FIG. 1. Comparison of the deterministic model (thin curves) and the stochastic model (thick curves) in the absence of mutations ($m = 0$) for two different values of the drive fitness: (a) $f_{DD} = 0.70$ and (b) $f_{DD} = 0.80$. In both panels, $n_{WW}(0) = 980$, $n_{WD}(0) = 20$, $\omega_B = 2$, $\zeta = 0.5$, and $h = 0.9$. The colors red, blue, and purple correspond to the genotypes WW, DD, and WD, respectively. The stochastic curves were generated from 200 replicate simulations, with the center of each curve determined by the replicate-averaged genotypic subpopulations, and the thickness at each time point delineating one standard deviation above and below that mean value.

III. RESULTS AND DISCUSSION

In order to study the effects of introducing a gene drive into a small, homeostatic, wild population, we initialize both of our models with a population of $N(0) = 1000$ mostly homozygous wild-type organisms. Some small percentage of the initial population is assigned to be heterozygous in the gene drive (genotype WD), and we then iterate Eq. (2) (with the birth and death rates set by the appropriate model) and track the evolution of the genotypic subpopulations until a new steady state is reached.

If our model gene drive has a homing rate h that is sufficiently close to unity and is assumed to be incapable of producing genetic mutations ($m = 0$), our deterministic model predicts that it will always grow to fixation within a population, though the number of organisms that remain at long times will depend sensitively on the drive fitness f_{DD}. For values below a critical threshold, the population will die off completely. We can determine this threshold by combining Eqs. (2)–(4) with the assumption that at steady state, the drive homozygous subpopulation should equal the total population:

$n_{DD}(t \to \infty) = N(t \to \infty)$. This leads to the expression

$$2\left\lfloor \frac{1}{4}\omega_B f_{DD}^2 N(t\to\infty)\right\rfloor = \left\lceil\left[\zeta + \left(\frac{1}{2}\omega_B - \zeta\right)\frac{N(t\to\infty)}{N(0)}\right]N(t\to\infty)\right\rceil. \quad (6)$$

If we assume that $N(t \to \infty) \gg 0$, then the algebraic error inherent to dropping the floor and ceiling brackets in Eq. (6) will be minimal, and we can simplify it to the following:

$$\frac{1}{2}\omega_B f_{DD}^2 - \zeta = \left(\frac{1}{2}\omega_B - \zeta\right)\frac{N(t\to\infty)}{N(0)}.$$

We have already noted that a stable steady state requires that the term in parentheses on the right-hand side of the above equation must be greater than zero. In that case, a positive solution for $N(t \to \infty)$ is only possible if the following condition holds:

$$f_{DD} > \sqrt{\frac{2\zeta}{\omega_B}}. \quad (7)$$

If f_{DD} is much smaller than $\sqrt{2\zeta/\omega_B}$, the poor fitness of the gene drive will crash the population. This threshold is only approximate, due to our dropping the floor and ceiling brackets in Eq. (6); the fate of the population must be determined more carefully in cases where the drive fitness is just above or just below the threshold value.

The capability of the gene drive to fixate within a population is dependent upon the efficiency of the drive, as measured by the homing rate, and we can quantify the critical efficiency needed for drive fixation as the threshold homing rate needed to cause an initial increase in the subpopulation of drive heterozygotes. If the conditions are favorable enough for the drive that such an increase occurs within the first generation of the simulation, the larger frequency of the drive allele in the next generation will only make the conditions for drive propagation that much more favorable, guaranteeing eventual fixation. Once again using Eqs. (2)–(4), we can derive the exact condition for initial drive propagation as

$$2\left\lfloor \frac{1}{8}\omega_B(1-h^2)n_{WD}(0)\frac{n_{WD}(0)}{N(0)}\right\rfloor + 4\left\lfloor \frac{1}{8}\omega_B(1+h)n_{WD}(0)\frac{N(0)-n_{WD}(0)}{N(0)}\right\rfloor > \left\lceil \frac{1}{2}\omega_B n_{WD}(0)\right\rceil. \quad (8)$$

This expression is just a statement that the number of WD births must outweigh the number of their deaths. The first term on the left side of the inequality comes from the average number of heterozygote births resulting from matings between pairs of WD organisms, and the second term is the average number of heterozygote births resulting from WD organisms mating with their wild-type homozygote brethren. If we assume that $n_{WD}(0) \ll N(0)$, then the first term on the left-hand side will, under most reasonable circumstances, be zero due to the action of the floor brackets. Our assumption of sex symmetry guarantees that $n_{WD}(0)$ will be evenly divisible by 2, so as long as ω_B is an integer, the ceiling brackets on the right-hand side can be formally removed. Noting that $\lfloor x \rfloor > y$ is equivalent to $x > y + 1$, we can reduce Eq. (8) to the following much simpler condition:

$$\frac{1}{8}\omega_B(1+h)n_{WD}(0)\frac{N(0)-n_{WD}(0)}{N(0)} > \frac{1}{8}\omega_B n_{WD}(0) + 1.$$

Solving the above inequality for the homing rate, we find that drive fixation requires

$$h > \left[\left(1 + \frac{8}{\omega_B n_{WD}(0)}\right)\frac{N}{N - n_{WD}(0)}\right] - 1. \quad (9)$$

For homing rates below this threshold, the drive allele will either reach a low-frequency steady state or be driven out of the population entirely. Experimental CRISPR-based gene drive constructs developed for *Anopheles gambiae* mosquitos have effective homing rates between 0.90 and 0.99 [9,18], but this threshold may be a practical concern for other species.

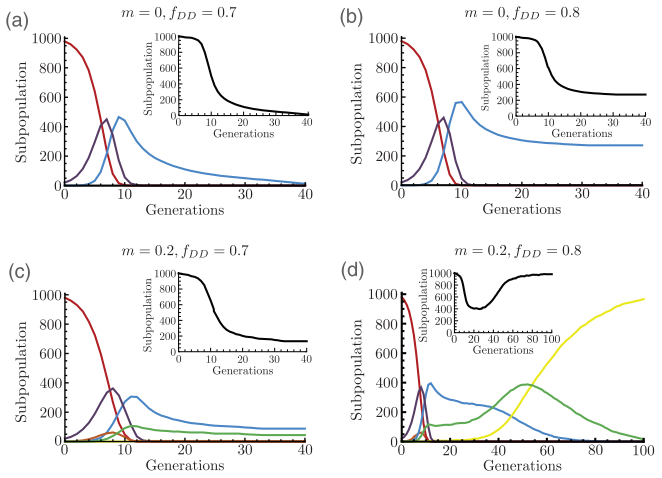

FIG. 2. The generational time evolution of genotypic subpopulations in the deterministic gene drive model for (a) $f_{DD} = 0.70, m = 0$; (b) $f_{DD} = 0.80, m = 0$; (c) $f_{DD} = 0.70, m = 0.2$; (d) $f_{DD} = 0.80, m = 0.2$. In all cases, the system was initialized with the same parameters as those used in Fig. 1. The primary colors red, blue, and yellow represent the homozygous genotypes WW, DD, and MM, respectively. The secondary colors purple, orange, and green represent the heterozygous genotypes WD, WM, and DM, respectively. The insets plot the total population vs time (in generations) for each case.

In Fig. 2 we illustrate representative trajectories of the deterministic model for both the case of population extinction [panel (a)] and drive fixation [panel (b)]. For our initial population of 1000 organisms, we assumed that 2% of them were heterozygous drive carriers ($h = 0.9$, $m = 0$). For an average litter size $\omega_B = 2$ and a mortality parameter $\zeta = 0.5$, the threshold value of the drive fitness given by Eq. (7) is $1/\sqrt{2} \approx 0.707$. Consistent with this, we observe population extinction for $f_{DD} = 0.70$ and drive fixation, at a reduced population level, for $f_{DD} = 0.80$. Note that these are the same conditions shown in Fig. 1, replotted here to facilitate a comparison with the analogous results for $m > 0$.

When the mutation rate is nonzero, the behavior of the deterministic model can become much richer. In Fig. 2(c), the same conditions as in Fig. 2(a) are reevaluated for $m = 0.2$, and we find that the emergence of drive-immune mutated

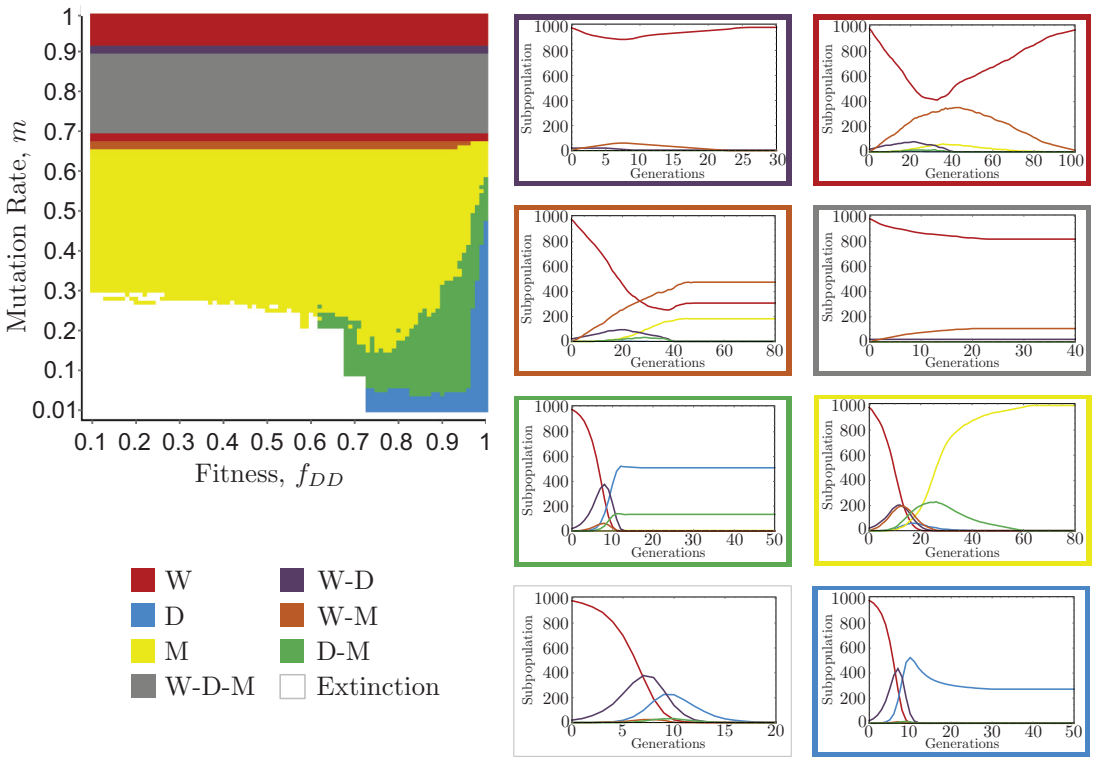

FIG. 3. The deterministic model phase diagram. To the right of the phase diagram, representative portraits of the genotypic subpopulation dynamics are shown for each of the eight phases. From left to right and bottom to top, the portraits represent the phases E, D, D-M, M, W-M, W-D-M, W-D, and W.

alleles can protect the population from extinction, resulting in a long-time, steady-state population consisting principally of a mixture of drive homozygotes and drive-mutation heterozygotes. A single pair of mutation homozygotes also survives in each generation at steady state. This seems like an anomalous artifact of our rounding schemes, since one would normally expect mutation homozygotes to outcompete their less fit drive counterparts given enough time, but any stochastic uptick in mutant homozygote births during one generation could just as easily be canceled out by a stochastic uptick in their deaths in a later generation. Our assumption in this model is that, over long periods of time, the overall trajectory of the population dynamics is stable to these occasional fluctuations, in which case a small, persistent population of mutant homozygotes surviving alongside a much large population of drive homozygotes is not impossible. As we shall demonstrate with our stochastic model, however, this assumption of stability is precisely what breaks down in the presence of mutations.

Shifting the fitness of the drive up to 0.8 enables the subpopulation of DM organisms to grow much larger, introducing enough M alleles into the total population to ensure their eventual fixation, as shown in Fig. 2(d). The assumed lack of a fitness cost for the MM genotype also allows the population to recover to its capacity.

The four trajectories depicted in Fig. 2 are only half the picture, as four more allelic outcomes become observable in the deterministic model as the mutation rate is increased further. The eight qualitatively distinct outcomes of the deterministic inheritance dynamics can be differentiated uniquely by which of the three alleles survive at steady state, and Fig. 3 is a phase diagram depicting the ranges of f_{DD} and m values over which each of these outcomes, or phases, may be observed. Representative genotypic evolution profiles for each of these phases are shown to the right of the figure, each boxed by a colored frame matching the color used in the phase diagram for that allelic outcome. The phase diagram itself was generated numerically by computing the long-time, steady-state population of the system at every point (f_{DD}, m) on a square mesh of side length 0.01. All other model parameters [$N(0)$, ω_B, etc.] were the same as those used in Fig. 1.

The genotypic subpopulation plots shown in Fig. 3 for the extinction (E), pure drive (D), drive-mutation (D-M), and pure mutation (M) phases represent the same four cases depicted in Fig. 2, albeit for different values of f_{DD} and m. In all of these cases, the wild-type allele is swiftly driven out of the population, supplanted by the drive. When the mutation rate gets sufficiently high, however, the drive can be supplanted by mutations before it can substantially reduce the frequency of wild-type alleles, resulting in steady-state populations that remain predominantly wild type. The remaining four phases all share this characteristic, achieving steady states at long times that reflect different balances between the WW, WM, MM, and WD genotypes, which all have unit fitness in our model. Thus, while an increasing mutation rate initially facilitates the dominance of the mutated alleles, as one might expect, there is a turnover point (at $m \approx 2/3$ for a one-thousand-organism population) where more frequent mutations ultimately favor production of wild-type homozygotes.

We would like to compare and contrast the genotypic phases observed at long times in our deterministic model

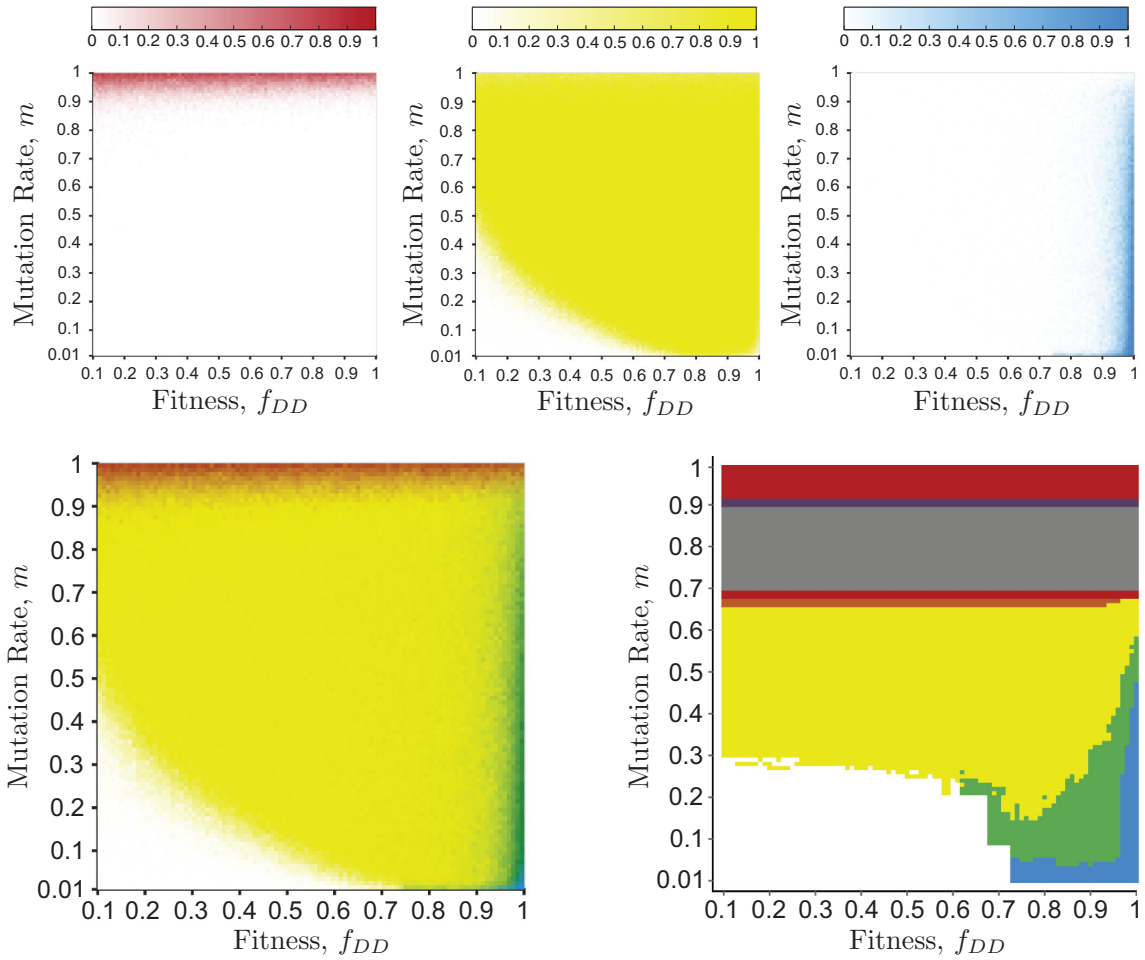

FIG. 4. The stochastic model phase diagram. The top-left plot uses a white-to-red color gradient to indicate the fraction of replicates that had surviving wild-type alleles at steady state for each pair of f_{DD} and m values. The remaining two plots in the top row illustrate the same for the mutated (yellow) and drive (blue) alleles. The bottom-left plot uses subtractive color mixing to combine the three top plots into a single phase diagram that can be compared with the deterministic result replotted in the bottom right of the figure for convenience.

with those observed in our stochastic model. To this end, we performed one hundred replicate simulations of our stochastic model for each pair of f_{DD} and m values in our mesh, once again matching all the other parameter values to those used for Fig. 1. We ran each replicate simulation for one thousand generations in order to ensure that steady state was reached in all cases, and the surviving alleles associated with at least one organism were recorded. We then plotted a separate phase portrait for each allele (see the top row of Fig. 4), using a gradient in color saturation to represent the fraction of replicates in which the corresponding allele survived. As in Fig. 2, the three primary colors red, yellow, and blue are used to represent the wild-type, mutation, and drive alleles, respectively. The three images were then combined using a subtractive color mixing model [27] to generate the bottom-left phase diagram, whose colors correspond to the same phases defined for the deterministic model in Fig. 3. (The deterministic phase diagram is replotted in the bottom right of Fig. 4 for ease of comparison.)

The contrast between the genotypic fates predicted by the two models is stark; whereas the deterministic model predicts large swaths of parameter space where the drive and wild type dominate, the stochastic model reduces those regions to the fringes of the diagram. The region of extinction is likewise reduced to occupy only the lower-left corner, and the *W-D* and *W-D-M* phases have disappeared entirely. The mutated allele now dominates nearly the entire phase diagram, even when the mutation rate is relatively low. The average behavior of the computational model is, by construction, intended to be equivalent to the analytic model, so long as the population dynamics are stable to small, stochastic fluctuations. This discrepancy thus implies a violation of the assumption of stability.

If the dynamics are indeed becoming unstable to small stochastic variations in the genotypic subpopulations, then injecting a small amount of randomness into the dynamics of the deterministic model should enable it to behave more like the stochastic model. To test this hypothesis, we amend our deterministic model by allowing the floor function in Eq. (3) to act like a ceiling function with some probability α during each iteration. This allows for a fractional organism to sometimes be rounded up to a whole one, thereby providing a simple mechanism for rare events—those births expected to occur less than once a generation—to at least occasionally transpire. (So long as α is not too large, this modification will still permit the possibility of drive fixation, the inclusion of

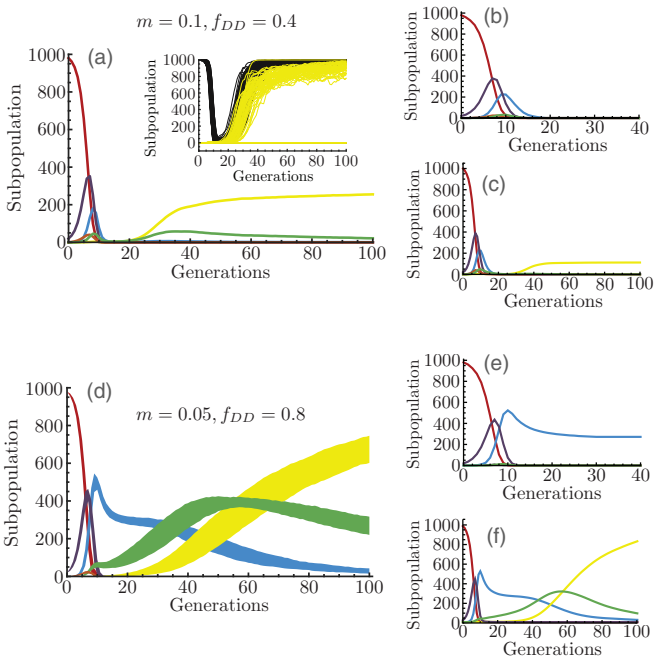

FIG. 5. Comparing the mean genotypic dynamics across three models. (a) The replicate-averaged dynamics of the stochastic model for $f_{DD} = 0.4$ and $m = 0.1$. The inset plots the trajectories of the MM subpopulation in yellow and the total population in black for all 100 replicates, illustrating the bifurcated character of the dynamics. (b) The results of the deterministic model for the same condition. (c) The result of modifying the deterministic model to account for the possibility of rare events. (d) The results of the stochastic model for $f_{DD} = 0.8$ and $m = 0.05$. Though there is no bifurcation in the long-time fate of the population, the instability of the dynamics is highlighted by plotting the replicate-averaged trajectory with a line thickness delineating one standard deviation above and below the mean. (e) The corresponding results for the deterministic model. (f) The results of the modified deterministic model for the same condition.

which was our original motivation for rounding births down in the first place.)

The effect of making this modification is demonstrated in Fig. 5. In Fig. 5(a), the replicate-averaged genotypic dynamics of the computational model are plotted for $f_{DD} = 0.4$ and $m = 0.1$. The fact that mutations do not appear to grow to fixation is a consequence of a bimodal distribution of long-time outcomes. As shown in the inset, mutations do fixate in about 28% of the trajectories, but the population goes extinct in the remaining 78% of them. This represents an extreme case of instability in which the population dynamics actually bifurcate towards two disparate outcomes. Figure 5(b) replots the deterministic model's result from Fig. 3 as a reference, and Fig. 5(c) plots the replicate-averaged result for the modified deterministic model with $\alpha = 0.1$. Though the agreement is not quantitative, our *ad hoc* correction is sufficient to reproduce the sort of dynamical bifurcation present in the fully stochastic model.

The final three panels of Fig. 5 plot the corresponding results for the condition $f_{DD} = 0.8$ and $m = 0.05$, where the deterministic model predicts fixation of the drive allele in a reduced population and the stochastic model predicts eventual fixation of the mutated allele. Although there is no dynamical bifurcation in this case, the plotted line thicknesses in Fig. 5(d), which delineate one standard deviation above and below the mean value at each time point, demonstrate that the dynamics are still relatively unstable, at least compared to the case of $m = 0$ (see Fig. 1). Allowing for rare events in the analytic model is once again sufficient to replicate the phenomenology of the fully stochastic results, including the dominance of DM organisms at intermediate times. The quantitative discrepancies presumably stem from our arbitrary assignment of a flat probability of α to all rare births, even though the likelihood of any specific type of birth will in fact depend sensitively upon the specific distribution of genotypes within the population of each generation.

While the rare event probability is difficult to characterize in general, we can at least quantify how the rarity of a mutant birth depends upon the population's genetic composition in the case of a population where mutations have not yet emerged. Such a population will consist of a mixture of the three genotypes WW, WD, and DD and, in accordance with the inheritance rules of our meiotic gene drive, a mutant birth can only occur if one parent has genotype WD. We want to compute the probability p_{mut} that a WD organism will give birth to an offsrpring with a mutant allele. Assuming a panmictic, sex-symmetric population, this probability is formally equal to

$$p_{\text{mut}} = \sum_{g}^{M} \sum_{g_2}^{\neg M} p(g|WD, g_2) \frac{n_{g_2}}{N},$$

where the sum over g is restricted to genotypes containing at least one M allele, and the sum over g_2 is restricted to genotypes containing no M alleles. Note that we do not care at this point whether the mutant offspring will survive to adulthood or not, so the fitness cost of having a drive homozygote parent is not taken into account in the above. Substituting the appropriate conditional probabilities into this equation, the double restricted sum can be evaluated, yielding the following expression:

$$p_{\text{mut}} = \frac{1}{2} hm \left[1 + \left(1 - \frac{1}{2} hm \right) \frac{n_{WD}}{N} \right]. \tag{10}$$

This probability is plotted in Fig. 6(a) as a function of the fraction n_{WD}/N for several mutation rates ranging from $m = 0.01$ to $m = 0.25$. Though mutations naturally cease to qualify as rare events for larger m, the mutation probability remains $<1\%$ for the lowest value considered, even for an entire population of wild-type/drive heterozygotes. Note that the probability is nonzero even for $n_{WD}/N = 0$, since our derivation assumes at least a single heterozygote by construction. A fractional population of zero corresponds to the limiting case in which a finite number of WD organisms exist as part of an infinite population.

Of course, the total likelihood of mutations arising within a population will be much higher than p_{mut}, since every organism has some chance of producing mutations upon mating. (Wild-type and drive homozygotes can produce mutations if they mate with a WD heterozygote.) If we make the simplifying assumption that females must initiate a mating interaction and

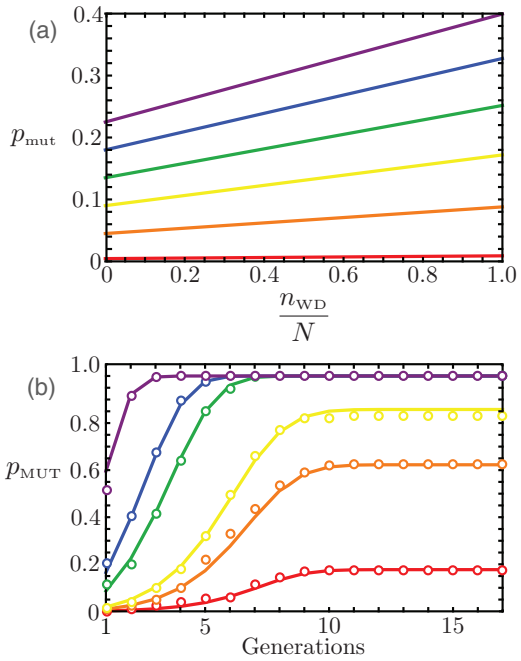

FIG. 6. (a) The probability of a WD organism in a mutation-free population giving birth to a mutated organism in the next generation, $p_{\rm mut}$, is plotted versus the fraction of WD organisms for $m = 0.01$, 0.05, 0.10, 0.15, 0.20, and 0.25 (from bottom to top). (b) The probability that at least one mutation has occurred by generation t, $p_{\rm MUT}(t)$, is plotted for $m = 0.0001$, 0.0005, 0.001, 0.005, 0.01, and 0.05 (again from bottom to top). The solid curves were generated using the deterministic model to calculate the genotypic subpopulation levels in each mutation-free generation, and the data points (circles) were computed using two hundred replicates of the stochastic model for each of the plotted mutation rates.

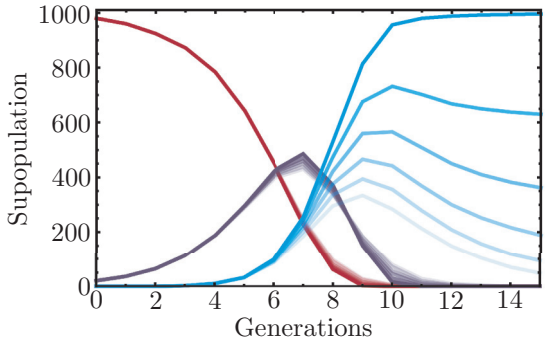

FIG. 7. The genotypic subpopulation dynamics of the deterministic model are superimposed (with varying opacities) for $m = 0$ and $f_{DD} = 0.5$, 0.6, 0.7, 0.8, 0.9, and 1.0 (from lightest to darkest). The same parameters and genotypic color scheme used throughout the paper are used here as well (see Fig. 1).

that multiple females can choose the same male mate, then the probability of each female producing mutated offspring in the next generation will be independent from that of every other female [the mean number of pairings between two genotypes will still be $(n_{g_1}/2)(n_{g_2}/N)$], and the cumulative probability that at least one mutation has occurred by generation t, $p_{\rm MUT}(t)$, can be expressed thusly:

$$p_{\rm MUT}(t) = 1 - \prod_{\tau=0}^{t} \left[\prod_{g_1} \left(\sum_{g,g_2}^{\neg M} p(g|g_1,g_2) \frac{n_{g_2}(\tau)}{N(\tau)} \right)^{\omega_B n_{g_1}(\tau)/2} \right].$$

Note that the above expression is written as the complement of the probability that no mutation alleles arise within the first t generations. As before, the explicit inheritance probabilities can be used to evaluate the bracketed expression:

$$p_{\rm MUT}(t) = 1 - \prod_{\tau=0}^{t} \left\{ \left(1 - \frac{1}{2}hm\right)^{\omega_B n_{WD}(\tau)/2} \times \left[1 - \frac{1}{2}hm \frac{n_{WD}(\tau)}{N(\tau)}\right]^{\omega_B N(\tau)/2} \right\}. \quad (11)$$

In Fig. 6(b), we plot this probability for several different values of the mutation rate. The solid curves were generated using Eq. (11) and the data points were found by counting the frequency of mutations across two hundred replicate runs of the stochastic model. Since this probability is the complement of the probability of having no mutations by generation t, the values of $n_{WD}(\tau)$ and $N(\tau)$ for each τ needed to evaluate Eq. (11) were computed from the deterministic model with $m = 0$. All the other parameters were assumed to be the same as those used in Fig. 2(b), though the result turns out to be effectively independent of the drive fitness. The lowest mutation rate considered in Fig. 6(a) is now the second highest considered in Fig. 6(b), and we see that within about five generations, the occurrence of at least one "rare" event is practically guaranteed. Even for a mutation rate of $m = 0.001$, a whole order of magnitude lower, the likelihood of observing at least one mutation plateaus at around 80% within ten generations.

Though we have now quantified just how much the size of a population can amplify the likelihood of rare inheritance events, it is still not obvious how a small fluctuation in the number of rare births can so consistently control the fate of the population as a whole. To see this, we once again consider the simpler case of $m = 0$, where it is possible (albeit unlikely, due to the relative stability of the dynamics in this regime) for stochastic fluctuations in the number of organisms born with the D allele to avert the fate of extinction at long times, even when the drive fitness is below the threshold defined by Eq. (7).

In Fig. 7 we use the deterministic model to plot the genotypic subpopulation trajectories for a range of values of f_{DD} lying both below and above the extinction threshold, and we find that the divergent fate of the drive homozygotes is propagated out of comparatively small differences in the distribution of genotypes that begin to appear after only five generations. This is roughly the time where the drive heterozygote population peaks, which happens to also be the point in the trajectory where the population is at its most diverse. This diversity translates to a tipping point wherein small fluctuations in the various subpopulation sizes can have a dramatic effect on the dynamics at long times. Because this is the first point in the trajectory where the number of organisms possessing the gene drive exceed the number without it, the

genotypes most likely to be amplified by fluctuations are those with the drive, potentially resulting in long-term drive survival instead of a population crash. Looking back to Fig. 5(d), we see similar phenomenology; the genotypic trajectories show almost no variance for the first ∼5 generations, but then a small degree of variation emerges that rapidly expands as time progresses.

IV. CONCLUSIONS

In this paper we have compared a deterministic population model of meiotic gene drive transmission to a stochastic, agent-based generalization in order to gauge the stability of long-time genotypic outcomes to random fluctuations in small, homeostatic populations. What we discovered, quite surprisingly, is that while the inheritance dynamics are stable to fluctuations overall, there is an unstable tipping point at short time scales ($t \sim 5$ generations) wherein the likelihood of rare inheritance events is amplified, and the resultant uptick in these rare births is capable of drastically affecting the genotypic distributions at long times.

In a formally infinite or very large population, even a negligible frequency of drive-resistant alleles will correspond to a finite number of organisms, but this is clearly not the case for a small to moderate-sized population. To prevent low-frequency alleles from inevitably propagating into our population through the accumulation of fractional animals, our deterministic model utilized rounding functions that ensured the possibility of drive survival was not precluded *a priori*. The dominance of mutations in our stochastic model suggests that our deterministic model would have been more accurate without these rounding functions, but that would have rendered it incapable of characterizing the population extinction observed in the stochastic model for small but nonzero mutation rates and drive fitnesses. Instead of removing the rounding entirely, we found that relaxing it just enough to allow occasional rare birthing events to occur was sufficient to achieve at least qualitative agreement with the stochastic model's predictions across parameter space. What this demonstrates is that our faulty assumption was not our rounding scheme but rather our expectation that the population dynamics should always consist of small, stochastic fluctuations about a robust mean trajectory. On the contrary, the rapid proliferation of the gene drive produces an unstable tipping point where even a miniscule population of more fit alleles can dramatically alter the system's dynamical behavior. In short, the robustness of drive-resistant mutations in a small, homeostatic population is due not to their gradual accumulation over many generations but to their high likelihood of first emerging at the most unstable point in the dynamical trajectory.

This mechanistic distinction is important, as current attempts to make gene-editing constructs more viable have largely focused on making them more difficult to resist [18,28]. If multiple mutations are needed, for example, to achieve full immunity from the action of the construct, then there is a higher chance of the population becoming decimated before mutations can render the gene drive ineffective. Our work suggests that an alternative approach viable for small, isolated populations could involve engineering a construct that somehow avoids the allelic diversity of the dynamic tipping point we have identified.

ACKNOWLEDGMENTS

The theories described and the resulting data presented herein, unless otherwise noted, were obtained from research conducted under the Environmental Quality Installations Program of the US Army Corps of Engineers by the US Army Engineer Research and Development Center (ERDC). The authors express gratitude to Dr. E. Ferguson, Technical Director of the US Army ERDC Environmental Quality Installations Program, for support of this research. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the US Army. The authors would also like to thank Dr. Michael A. Rowland for insightful discussions and assistance with computational resources.

[1] M. R. Goddard and A. Burt, Proc. Natl. Acad. Sci. USA **96**, 13880 (1999).
[2] S. P. Sinkins, Nat. Rev. Genet. **7**, 427 (2006).
[3] J. van der Oost, M. M. Jore, E. R. Westra, M. Lundgren, and S. J. J. Brouns, Trends Biochem. Sci. **34**, 401 (2009).
[4] F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang, Nat. Protoc. **8**, 2281 (2013).
[5] A. L. Caplan, B. Parent, M. Shen, and C. Plunkett, EMBO Rep. **16**, 1421 (2015).
[6] B. L. Webber, S. Raghu, and O. R. Edwards, Proc. Natl. Acad. Sci. USA **112**, 10565 (2015).
[7] K. M. Esvelt and N. J. Gemmell, PLoS Biol. **15**, e2003850 (2017).
[8] N. J. Bax and R. E. Thresher, Ecol. Appl. **19**, 873 (2009).
[9] A. Hammond *et al.*, Nat. Biotechnol. **34**, 78 (2016).
[10] C. M. Leitschuh, D. Kanavy, G. A. Backus, R. X. Valdez, M. Serr, E. A. Pitts, D. Threadgill, and J. Godwin, J. Responsible Innovation **5**, S121 (2017).
[11] A. Deredec, A. Burt, and H. C. J. Godfray, Genetics **179**, 2013 (2008).
[12] O. S. Akbari *et al.*, Science **349**, 927 (2015).
[13] J. E. DiCarlo, A. Chavez, S. L. Dietz, K. M. Esvelt, and G. M. Church, Nat. Biotechnol. **33**, 1250 (2015).
[14] R. L. Unckless, A. G. Clark, and P. W. Messer, Genetics **205**, 827 (2017).
[15] C. Noble, B. Adlam, G. M. Church, K. M. Esvelt, and M. A. Nowak, eLIFE **7**, e33423 (2018).
[16] C. Noble, J. Olejarz, K. M. Esvelt, G. M. Church, and M. A. Nowak, Sci. Adv. **3**, e1601964 (2017).
[17] P. A. Eckhoff, E. A. Wenger, H. C. J. Godfray, and A. Burt, Proc. Natl. Acad. Sci. USA **114**, E255 (2016).
[18] J. M. Marshall, A. Buchman, H. M. Sánchez, and O. S. Akbari, Sci. Rep. **7**, 3776 (2017).
[19] A. K. Lindholm *et al.*, Trends Ecol. Evol. **31**, 315 (2016).
[20] J. J. Bull and H. S. Malik, PLoS Genet. **13**, e1006850 (2017).

[21] L. Sandler and E. Novitski, Am. Nat. **91**, 105 (1957).
[22] T. Prout, J. Bundgaard, and S. Bryant, Theor. Popul. Biol. **4**, 446 (1973).
[23] G. J. Thomson and M. W. Feldman, Theor. Popul. Biol. **5**, 155 (1974).
[24] D. Haig and A. Grafen, J. Theor. Biol. **153**, 531 (1991).
[25] M. W. Feldman and S. P. Otto, Am. Nat. **137**, 443 (1991).
[26] S. A. Frank, Evolution **45**, 1714 (1991).
[27] S. A. Burns, arXiv:1710.05732.
[28] Y. Yan and G. C. Finnigan, Sci. Rep. **8**, 17277 (2018).