

Appendix to Sex as an Algorithm: The Theory of Evolution under the Lens of Computation.

A. Additional Background and Literature

Because of CACM space limitations, in the main text of our paper we did not cite as much relevant and related previous material as we could, and here we give a few more references.

A.1. Introduction

Jacque Monod's comment was made in [43].

Dobzhansky made his point in a 1964 presidential address to the American Society of Zoologists and used it in the title of an essay [13].

A good introduction to the scientific crisis that arose from the rediscovery of Mendel's laws—"the biometry-Mendelism debate"—is [52].

The reader interested in Fisher, Wright and Haldane's works may start, for example, by reading [21, 61, 62, 23].

A.2. Evolution and Computer Science

See references in main text.

A.3. The Role of Sex

On sex and recombination:

From an evolutionary theoretical perspective, sex and recombination have been studied in two principal ways. The first consisted of models attempting to answer the question of what benefit sex may bring (to individuals, populations or species). Here, some main 20th-century hypotheses have been the

Fisher-Muller theory [21, 45], Muller's ratchet [46], the deterministic mutation hypothesis [30] and the parasite hypothesis [38, 28, 24]. For other hypotheses, see, e.g., [14, 7], and for reviews, see [20, 1, 31, 19, 5]. The other main approach has been modifier theory, focusing not on what benefits sexual recombination may bring, but on whether and how recombination rates evolve within a traditional population genetic framework [48, 16, 17, 18, 2, 6, 3, 10, 34, 22]. Finally, awareness of the costs of sex [56] has become part of the investigation since Maynard-Smith's formalization of the two-fold cost, according to which an all-female lineage can reproduce twice as fast than a sexual lineage because it spends no effort on male production [40, 41].

Despite nearly a century of research, problems abound. First, each of the hypotheses on the benefit of sex required specific conditions to be met, in contrast to the prevalence of sex in nature [19, 58, 56]. Thus, in 1999, West et al. proposed the pluralistic hypothesis [58], according to which different benefits explain sex in different species—arguably an unparsimonious answer to a very weighty question. Second, in the realm of modifier theory, the reduction principle has demonstrated the general difficulty of maintaining non-zero sexual recombination rates [48, 36, 16, 17, 18, 2, 6, 19], and specific parameters are needed in order to reverse it [3, 10,

22], which is again incongruent with the prevalence of sex in nature.

The issue of the two-fold cost also suffers from many difficulties. First, it holds only in the case where males make no parental investment and where the production of males adds an investment equal in magnitude to that of the production of females. Since this condition excludes many organisms, such as the majority of the flowering plants, it is not clear how useful a concept it is for understanding the problem of sex in its proper generality. Second, for the same reason, if the two-fold cost were important, one would have expected to see more sex in isogamous than anisogamous organisms (in the former, both mating types invest in offspring), whereas in nature, in a sense, the opposite is true [26]. Third, there are still many other potentially important costs of sex [56], making it so that their sum total is neither “two-fold” nor practically quantifiable. Indeed, the fact that it has become a fashion to use “two-fold” as a benchmark in theoretical studies of sex (such that if and only if a benefit of sex is found that overcomes it, sex is explained), may have had more to do with the fame of Maynard-Smith and the catchiness of his formalization than with well-thought-out scientific reasons. Fourth, if sex had no costs, we would still need to understand its role, making it so that the investigation of sex is not inseparable from the

investigation of its costs.

Not all experts on the theory of the evolution of sex and recombination agree with us that the problem is wide open [5]. However, a word of caution is that even Nobel Laureate and legendary geneticist Hermann J. Muller wrote in 1932 that genetics has finally settled the *raison d’être* and function of sex [45], even though other main hypotheses on these questions were brought forth later besides his own. Our paper demonstrates that there is much that is of interest outside of the traditional boundaries within which sex has been studied, and which is informed by the computational lens—from the robustness of mixability, to the view of sexual evolution as a coordination game between genes played according to MWUA, to the view of sex as a randomized algorithm, to the connection between sex and a computational view of genetic change.

On genetic interactions:

So far, there have been two fundamental ideas on what drives adaptive evolution. One is that effective selection acts at its core on additive variance—on the part of the contribution of an allele to fitness that is independent of the genetic context [21]. This idea is tied to the problem of sex: Williams argued that effective selection can only act on sufficiently small stretches of the genome—on single loci as independent actors—because only a small stretch

lasts through the generations undisturbed by sexual recombination as to be able to experience selection effectively [59]. However, already from Fisher's days it has been recognized that a view based on non-interacting loci is unrealistic, hence Wright put forth the second major idea—that chance alone may drive the fixation of alleles under some conditions, forming the basis for new beneficial genetic interactions by chance [61, 62]. This chance fixation of alleles is called “random genetic drift,” and Wright's view of adaptive evolution based on this drift, selection on additive variance and migration is called the “shifting balance theory” [61, 62]. Note that Wright believed that genetic interactions were key, but did not offer a selection-based mechanism to explain their formation. Thus, following Fisher and Wright, researchers were left with the main principles of the chance arising of a mutation, the chance fixation of it and selection acting on additive variance, among other issues such as migration [21, 61, 62]. Besides these principles, many models in population genetics have examined the selective effect of genetic interactions, or “epistasis” [60]. For example, Fisher was interested in genetic interactions for his theory of the evolution of dominance [21]. As another example, Kondrashov's hypothesis on the role of sex required (perhaps unrealistically) the assumption that a particular type of genetic interaction is

nearly ubiquitous [30, 32]. Note, however, that in this nearly century-long research, genetic interactions have been studied as a subsidiary element in the theory—as a correction to the core theory of selection acting on additive variance.

Thus, while the obvious effect of sexual recombination is that it creates transient genetic combinations—indeed it creates individual genotypes—it has not been considered that selection acting directly on these transient combinations—on individual genotypes as complex wholes—can be at the core of the adaptive evolutionary process. This has not been considered precisely because of the transience of these combinations. In this paper we have discussed two ways by which selection on transient, complex genetic combinations can be important: one is that such selection favors mixability—the ability of alleles to perform well across many different genetic combinations—and the other is that mutation as a computational event allows information to be transmitted across loci and thus changes the nature of heredity in the presence of recombination.

A.4. A Game Between Genes

On the search for an optimization principle:

Further results in the line of thinking originated by Wright [63] are given in [27].

Fisher's fundamental theorem is often invoked in discussions about the meaning of the evolutionary process. It demonstrates, unfortunately in a partial sense that has been criticized [51, 15, 47], that the rate of increase of the population mean fitness is equal to the additive variance in fitness [21].

Following the works of Fisher and Wright, and despite the severe limitations on their results, the population mean fitness has often been tacitly treated in theoretical studies as the "end goal of evolution." Specifically in theoretical research on the benefit of sex, investigators have often assumed that the question to be answered is "how does sex facilitate the increase in the population mean fitness measure" [19]. One important criticism of this assumption is based on the levels-of-selection debate and has already been explained elsewhere [19]. Another important point is that the term "fitness" has two different meanings. In one, it is the harmony between the organism's complex adaptive structure and behavior, and its environment. In the other, it is the organism's expected number of surviving offspring. For Darwin, the former fitness was the end and the latter was a means of adaptive evolution, yet the mathematization of evolutionary theory has subtly shifted the discourse to the latter. Principles such as reviewed here may be part of an expanded understanding of the evolution-

ary process that is a work in progress, and thus of the increase in fitness in the original sense.

On evolutionary game theory:

Game theory has been used in evolutionary theory before to study strategic individual behavior [42, 57]. The concepts of it are also invoked at the genetic level by the selfish gene view [11]. The use of game theory described here is very different. We show that the basic equations of evolution in the presence of natural selection and sexual reproduction can be described as a repeated coordination game between genes, and furthermore the game is played according to the powerful MWUA. Thus, game theoretic concepts are applied here at a fundamental level, prior to the consideration of any conflict—the main ingredient of past evolutionary game theory.

A.5. Are Mutations Random?

It is agreed today by biologists that there exist complex biological influences on mutation. The question is whether these biological influences are important for the evolutionary process or not. The brief argument made in this paper on the topic is fundamentally different than those made before.

In the past literature on non-accidental mutation in evolutionary biology one can recognize two main ideas. One idea was that of an "adaptive" or "directed" mutation (but see [55] for a

distinction between these two terms) in bacteria which responds to the immediate environment (e.g., [8]). This idea was met with resistance due to its presumed Lamarckian implications (e.g., [55, 37]). Another idea was that evolution in the long term could lead to mutational mechanisms that respond in a generic way to pressures encountered repeatedly before (e.g., [44, 53, 9, 33]). For example, it was hypothesized that an increased general mutation rate in bacteria under conditions of stress could be a result of natural selection (see [54] for a review and criticism). However, this line of thinking provides what may be a limited conceptualization and interpretation of the evidence. First, models show that it is hard to explain the evolution of such mechanisms from a traditional perspective [36, 18, 2, 4]. Second, it relies on accidental mutation as the ultimate source of novelty, confining mutational mechanisms to a later, minor role, and ignoring the possibilities that mutational mechanisms are involved in evolution in general and in a continually evolving manner specific to the currently evolving adaptation.

Importantly, neither of these preexisting lines of thinking has offered the insight that, being a result of genetic interactions, mutation can be seen as an event of information flow and computation, accepting “inputs” from multiple loci, and producing an “output” that is

the mutation. For starters, this changes the nature of heredity in a way that allows the complex, transient genetic combinations generated by sex to leave heritable effects on future generations [39]. Thus, the computational-lens [50, 29] stands to make a contribution also to our understanding of mutation.

A.6. On the preservation of variation

For modern work reaffirming the paradox of variance exposed by [49] see [35].

For overdominance, or heterozygote advantage, see [12, 25].

1. REFERENCES

- [1] In R. Michod and B. Levin, editors, *The Evolution of Sex: An Examination of Current Ideas*. Sinauer, Sunderland, 1988.
- [2] L. Altenberg and M. W. Feldman. Selection, generalized transmission and the evolution of modifier genes. I. The reduction principle. *Genetics*, 117:559–572, 1987.
- [3] N. Barton. A general model for the evolution of recombination. *Genet Res*, 65:123–144, 1995.
- [4] N. Barton, L. Partridge, et al. Limits to natural selection. *BioEssays*, 22(12):1075–1084, 2000.
- [5] N. H. Barton and B. Charlesworth. Why sex and recombination? *Science*, 281:1986–1990, 1998.
- [6] A. Bergman and M. W. Feldman. More on selection for and against

- recombination. *Theor Popul Biol*, 38:68–92, 1990.
- [7] H. Bernstein, H. C. Byerly, F. A. Hopf, and R. E. Michod. Genetic damage, mutation, and the evolution of sex. *Science*, 229(4719):1277–1281, 1985.
- [8] J. Cairns, J. Overbaugh, S. Miller, et al. The origin of mutants. *Nature*, 335(6186):142–145, 1988.
- [9] L. H. Caporale. *Darwin in the Genome: Molecular Strategies in Biological Evolution*. McGraw-Hill, New York, 2003.
- [10] B. Charlesworth. Directional selection and the evolution of sex and recombination. *Genet Res*, 61:205–224, 1993.
- [11] R. Dawkins. *The Selfish Gene*. Oxford University Press, 1976.
- [12] T. Dobzhansky. A review of some fundamental concepts and problems of population genetics. In *Cold Spring Harbor Symp Quant Biol*, volume 20, pages 1–15. Cold Spring Harbor Laboratory Press, 1955.
- [13] T. Dobzhansky. Nothing in biology makes sense except in the light of evolution. *American Biology Teacher*, 35:125–129, 1973.
- [14] I. Eshel and M. W. Feldman. On the evolutionary effect of recombination. *Theoretical Population Biology*, 1(1):88–100, 1970.
- [15] W. J. Ewens. An interpretation and proof of the fundamental theorem of natural selection. *Theoretical Population Biology*, 36(2):167–180, 1989.
- [16] M. W. Feldman. Selection for linkage modification. I. Random mating populations. *Theor Popul Biol*, 3:324–346, 1972.
- [17] M. W. Feldman, F. B. Christiansen, and L. D. Brooks. Evolution of recombination in a constant environment. *P Natl Acad Sci USA*, 77:4838–4841, 1980.
- [18] M. W. Feldman and U. Liberman. An evolutionary reduction principle for genetic modifiers. *P Natl Acad Sci USA*, 83:4824–4827, 1986.
- [19] M. W. Feldman, S. P. Otto, and F. B. Christiansen. Population genetic perspectives on the evolution of recombination. *Annu Rev Genet*, 30:261–295, 1997.
- [20] J. Felsenstein. The evolutionary advantage of recombination. *Genetics*, 78(2):737–756, 1974.
- [21] R. A. Fisher. *The Genetical Theory of Natural Selection*. The Clarendon Press, Oxford, 1930.
- [22] L. Hadany and T. Beker. On the evolutionary advantage of fitness-associated recombination. *Genetics*, 165(4):2167–2179, 2003.
- [23] J. B. S. Haldane. *The Causes of*

- Evolution*. Longmans Green and Co., 1932.
- [24] W. D. Hamilton. Sex versus non-sex versus parasite. *Oikos*, 35:282–290, 1980.
- [25] D. L. Hartl and A. G. Clark. *Principles of Population Genetics*. Sinauer, 1997.
- [26] L. D. Hurst and J. R. Peck. Recent advances in understanding of the evolution and maintenance of sex. *Trends Ecol Evol*, 11(2):46–52, 1996.
- [27] Y. Iwasa. Free fitness that always increases in evolution. *Journal of Theoretical Biology*, 135(3):265–281, 1988.
- [28] J. Jaenike. A hypothesis to account for the maintenance of sex within populations. *Evol Theory*, 3:191–194, 1978.
- [29] R. M. Karp. Understanding science through the computational lens. *Journal of Computer Science and Technology*, 26(4):569–577, 2011.
- [30] A. Kondrashov. Selection against harmful mutations in large sexual and asexual populations. *Genetical Research*, 40:325–332, 1982.
- [31] A. Kondrashov. Classification of hypotheses on the advantage of amphimixis. *Journal of Heredity*, 84(5):372–387, 1993.
- [32] A. S. Kondrashov. Deleterious mutations and the evolution of sexual reproduction. *Nature*, 336(6198):435–440, 1988.
- [33] E. V. Koonin. *The Logic of Chance: The Nature and Origin of Biological Evolution*. FT Press, 2011.
- [34] A. Korol, I. Preygel, and S. Preygel. *Recombination Variability and Evolution*. Chapman Hall, London, 1994.
- [35] E. M. Leffler, K. Bullaughey, D. R. Matute, W. K. Meyer, L. Segurel, A. Venkat, P. Andolfatto, and M. Przeworski. Revisiting an old riddle: what determines genetic diversity levels within species. *PLoS Biol*, 10(9):e1001388, 2012.
- [36] E. G. Leigh Jr. Natural selection and mutability. *Am Nat*, 104:301–305, 1970.
- [37] R. E. Lenski and J. E. Mittler. The directed mutation controversy and neo-darwinism. *Science*, 259:188–194, 1993.
- [38] D. A. Levin. Pest pressure and recombination systems in plants. *Am Nat*, pages 437–451, 1975.
- [39] A. Livnat. Interaction-based evolution: how natural selection and nonrandom mutation work together. *Biology Direct*, 8(1):24, 2013.
- [40] J. Maynard-Smith. The origin and maintenance of sex. In G. Williams, editor, *Group Selection*, pages 163–175. Aldine Atherton, 1971.

- [41] J. Maynard-Smith. *The Evolution of Sex*. Cambridge University Press, 1978.
- [42] J. Maynard-Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.
- [43] J. L. Monod. On the molecular theory of evolution. In R. Harré, editor, *Problems of scientific revolution: Progress and obstacles to progress in the sciences*, pages 11–24. Oxford University Press, Oxford, 1975.
- [44] E. R. Moxon, P. B. Rainey, M. A. Nowak, and R. E. Lenski. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current Biology*, 4(1):24–33, 1994.
- [45] H. J. Muller. Some genetic aspects of sex. *Am Nat*, 66:118–138, 1932.
- [46] H. J. Muller. The relation of recombination to mutational advance. *Mutation Res*, 1:2–9, 1964.
- [47] T. Nagylaki. Error bounds for the fundamental and secondary theorems of natural selection. *Proceedings of the National Academy of Sciences*, 88(6):2402–2406, 1991.
- [48] M. Nei. Modification of linkage intensity by natural selection. *Genetics*, 57:625–641, 1967.
- [49] E. Nevo, A. Beiles, and R. Ben-Shlomo. The evolutionary significance of genetic diversity: ecological, demographic and life history correlates. *Lecture Notes in Biomathematics*, 53:13–213, 1984.
- [50] C. Papadimitriou. The algorithmic lens: How the computational perspective is transforming the sciences. In *2007 Federated Computing Research Conference*. Speech, San Diego, USA, June 8–16, 2007.
- [51] G. R. Price. Fisher’s “fundamental theorem” made clear. *Annals of human genetics*, 36(2):129–140, 1972.
- [52] W. B. Provine. *The Origins of Theoretical Population Genetics*. University of Chicago Press, 1971.
- [53] S. M. Rosenberg. Evolving responsively: adaptive mutation. *Nature Reviews Genetics*, 2:504–515, 2001.
- [54] J. R. Roth, E. Kofoid, F. P. Roth, O. G. Berg, J. Seger, and D. I. Andersson. Regulating general mutation rates: examination of the hypermutable state model for Cairnsian adaptive mutation. *Genetics*, 163(4):1483–1496, 2003.
- [55] P. D. Sniegowski. The origin of adaptive mutants: random or nonrandom? *J Mol Evol*, 40:94–101, 1995.
- [56] S. C. Stearns and R. F. Hoekstra. *Evolution: An Introduction*. Oxford University Press, New

York, 2005.

- [57] J. W. Weibull. *Evolutionary Game Theory*. MIT Press, 1995.
- [58] S. A. West, C. M. Lively, and A. F. Read. A pluralist approach to sex and recombination. *J Evol Biol*, 12:1003–1012, 1999.
- [59] G. C. Williams. *Adaptation and Natural Selection*. Princeton University Press, 1966.
- [60] J. B. Wolf, E. D. Brodie, and M. J. Wade. *Epistasis and the Evolutionary Process*. Oxford University Press, 2000.
- [61] S. Wright. Evolution in mendelian populations. *Genetics*, 16:97–159, 1931.
- [62] S. Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc 6th Int Cong Genet*, 1:356–366, 1932.
- [63] S. Wright. The distribution of gene frequencies in populations. *Proceedings of the National Academy of Sciences of the United States of America*, 23(6):307, 1937.

The Experts Problem

Imagine that every day over five years you receive financial advice from ten experts. If you follow the advice of any one of them, that day you will realize a gain somewhere between -1 and $+1$ (you fill in the units). You have no idea how good each of them is, and yet each day you must choose one expert. What is a good algorithm for choosing an expert, day after day, so you will end up doing well?

This is not a well defined problem, because in our objective “to do well” we have not specified “compared to what?” Let us set a very ambitious goal. You want to choose experts in such a way that, five years later, you can look back and say: *“In retrospect, Expert 4 was the best, because if I had followed her advice throughout I would be better off than if I had followed the advice of any other single expert. But I did not know this in the beginning. Still, I managed to do almost as well as if I had followed exclusively the advice of Expert 4.”* In other words, we want to find an algorithm which picks the best expert — give or take something very small — *from the start and in hindsight!* At first sight, this might seem impossible. Note that we are not assuming anything about any probabilistic distribution of the experts’ performance. In fact, it is instructive to think that the outcomes of the experts’ advice each day are chosen by an adversary who wants you to fail in your goal.

Very surprisingly, this feat is made possible by a very simple algorithm called *multiplicative weight updates (MWU)*. It has a long history: It was discovered in the 1950s by economists, then in the 1980s in the mathematics of finance, then in the 1990s by researchers in AI (where it has been called first “Winnow,” “Hedge,” and finally “Boosting”), and finally by theoreticians as MWU. The algorithm is the following:

Fix a very small number $\epsilon > 0$ (see below for appropriate value)

Give each expert i (out of n) the same probability $p_i = \frac{1}{n}$

At each day $t = 1, \dots, T$ do the following:

 Pick an expert at random, where you choose i with probability p_i

 Let g_j be the gain you would have obtained if you had chosen expert j , for $j = 1, \dots, n$

 For $j = 1, \dots, n$, update the probabilities as follows: $p_j \leftarrow p_j(1 + \epsilon g_j)$,

 and divide all p_j ’s by $\sum_{i=1}^n p_j$ (to keep them adding to one).

That is, each day you boost (or decrease) the probability of each expert by a small amount proportional to its gain (or loss). The theorem says that, if you choose $\epsilon = \sqrt{\frac{\ln n}{T}}$, your total gain in the end of T days will be, in expectation, within $\sqrt{T \ln n}$ of the optimum. For large T this will be insignificant, compared to the range $[-1, +1]$. In our example with $n = 10$ and $T = 1825$, on an average day you are guaranteed to be only about .05 units of gain below the performance of the best expert.