

## Optimality Pressures toward Lateralization of Complex Brain Functions

Luís F. Seoane<sup>✉\*</sup>

*Departamento de Biología de Sistemas, Centro Nacional de Biotecnología (CSIC),  
C/ Darwin 3, 28049 Madrid, Spain  
and Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain*

 (Received 23 March 2023; accepted 7 July 2023; published 13 September 2023)

Most animal brains present two mirror-symmetric sides, but closer inspection reveals a range of asymmetries (in shape and function) that seem more salient in more cognitively complex species. Sustaining symmetric, redundant neural circuitry has associated metabolic costs, but it might aid in implementing computations within noisy environments or with faulty pieces. It has been suggested that the complexity of a computational task might play a role in breaking bilaterally symmetric circuits into fully lateralized ones; however, a rigorous, mathematically grounded theory of how this mechanism might work is missing. Here, we provide such a mathematical framework, starting with the simplest assumptions but extending our results to a comprehensive range of biologically relevant scenarios. We show mathematically that only fully lateralized or bilateral solutions are relevant within our framework (dismissing configurations in which circuits are only partially engaged). We provide maps that show when each configuration is preferred depending on costs, fitness contributed, circuit reliability, and task complexity. We discuss evolutionary paths leading from bilateral to lateralized configurations and other possible outcomes. The implications of these results for evolution, development, and rehabilitation of damaged or aging brains is discussed. Our work constitutes a limit case that should constrain and underlie similar mappings when other aspects (besides task complexity and circuit reliability) are considered.

DOI: [10.1103/PhysRevX.13.031028](https://doi.org/10.1103/PhysRevX.13.031028)

Subject Areas: Biological Physics, Complex Systems,  
Interdisciplinary Physics

### I. INTRODUCTION

In Bilateria (an ample clade of animals that includes humans), the body displays an overall mirror-symmetric disposition. Mirror (or bilateral) symmetry mathematically captures our day-to-day experience that reflected objects look the same but are inverted sidewise (an object's left side appears as the reflection's right side). A mirror symmetry exists in bilaterian bodies with respect to our sagittal central plane (which separates our left and right sides). For example, both of our hands look like the mirror reflection of each other with respect to that plane.

This symmetry is present in most parts of bilaterian central nervous systems—including the human brain, where it also appears broken at a range of levels [1–7]. From a structural perspective, some brain areas grow bigger than their symmetric counterpart—for example, several regions of the frontal and temporal left hemispheres are

usually thicker than their contralateral opposites [5,6]. Some of these structural differences underlie an asymmetry in function as well—human language is a preferred example as it depends on the development, in the dominant side only, of a series of areas (Broca's, Wernicke's, etc.) [8–15]. Functional asymmetry may appear without such salient morphological differences—e.g., the right hemisphere usually dominates high-level visual processing, taking care of discerning fine details, while the left hemisphere processes coarser visual aspects [16,17]. Whether linked to structure and function or not, behavior can break the mirror symmetry as well—for example, hand dominance in humans, which is also present at different levels (and varying in choice of dominant side) in other vertebrates [18–21].

The discovery of Broca's area proved that brain function is localized (which was unclear at the time) and that its bilateral symmetry is broken in humans. An assumption lingered that the symmetry breaking was due to complex human cognitive abilities; thus, it was assumed that increased complexity would generally favor lateralization. Indeed, brain asymmetry was deemed a human trait that other species lacked [22–24]. As this was proved wrong (and brain asymmetry was found to be widespread throughout animals [18–21,25,26]), neuroscientists worried less about the role of sheer complexity in symmetry breaking. They focused on more tangible mechanisms, such as the

\*brigan@gmail.com

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

input of light during hatching in birds [20] or faster processing within a single hemisphere [27]. The hypothesis of complexity as a driving force of brain symmetry breaking survived with nuances [2,28], often linked to mechanistic explanations (e.g., that each hemisphere could have specific abilities and more complex tasks could recruit subunits in each hemisphere differently [26], or that asymmetric brains would allow a more optimal packing, supporting more functions [28]). However, the following questions remain: Can cognitive complexity *per se* be a driving force behind the breaking of mirror symmetry in neural systems? If so, is it possible to infer thresholds of complexity beyond which bilaterality is doomed? Are there parsimonious ways through which the lost symmetry of a neural circuit could reemerge? A rigorous mathematical formalism to answer these questions is lacking.

We tackle these issues from a computational framework, within which notions of complexity can be well defined and quantified. We assume that neural circuits carry out some computational job—whether relaying signals, taking in inputs, or transforming them in some way. All these actions have metabolic costs (incurred by engaged neurons) but also more abstract thermodynamic costs that relate to the complexity of the computation itself or the reliability of neural circuits [29–31].

Reliability is important since real neural systems work within noisy environments. How to perform computations with unreliable units is a problem that has worried researchers since the inception of computer science [32–34]. Redundancy (introducing units that perform the same operations in parallel or that can substitute a damaged component) is often a preferred strategy. However, too much redundancy incurs in unnecessary costs (e.g., as responses from parallel circuits need to be integrated) and multiplies the energetic metabolic expenses. When is redundancy preferred, depending on circuitry reliability and task complexity?

Throughout evolution, bilateral symmetry has been a source of redundancy that provided us with neural circuits in duplicated pairs. While our discussion is centered on lateralization versus bilaterality, our results are general for any pair of redundant circuits. Optimality in wiring and regimes of efficient activity can appear linked in neural systems [35]. As Darwinism proceeds, we expect that optimality constraints will guide, allow, or prevent certain evolutionary paths. These are the possibilities that we intend to illuminate in this paper. Matters of optimality, reliable computation, and selection are also relevant during development—as learning engages in a Darwinian process of its own [36]. Finally, the optimality of lateralized or bilateral neural configurations might be challenged again as we age [37] or as our nervous system becomes damaged. A landscape of optimal configurations for neural circuits helps us navigate these cases and even suggests treatments for damaged or aging brains.

In this paper, we lay out a minimal mathematical model to map optimal configurations (bilateral versus fully

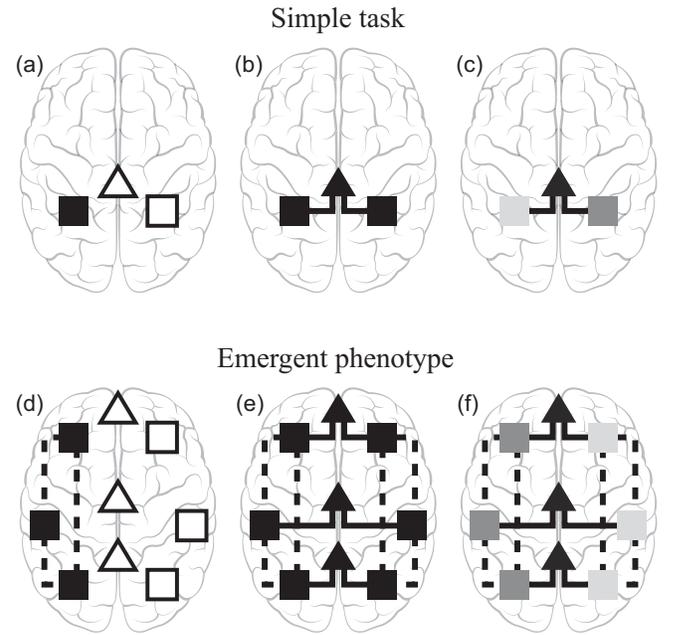


FIG. 1. Simple model for brain bilaterality versus lateralization. In panel (a), we show that a simple task can be carried out by a lateralized circuit (black square). A similar unit in the right hemisphere (white square) and circuitry to integrate both symmetric homologues (white triangle) are left unused. In panel (b), we show that if both mirror-symmetric units are engaged to solve that simple task (black squares), the integration mechanism (black triangle) is needed. This is a bilaterally symmetric configuration. In panel (c), we show that it might be possible to implement the task by a graded engagement (shades of gray) of both units and the integration mechanism. (d)–(f) Modeling an emerging phenotype that recruits  $M$  (in this case,  $M = 3$ ) different modules, each implementing a simpler subtask. Lateralized (d) and fully (e) or partial (f) bilateral solutions are possible again. We determine which configurations are optimal in each case.

lateralized—allowing any intermediate designs) of computing neural units as a function of their running costs, fitness gain, reliability, and complexity of the task at hand. All results are analytical. Mathematical details are developed in the appendixes. In Sec. II, we discuss our most important insights. In Sec. II A, we study the least-complicated case, in which a *simple task* is implemented by a module. This module consists of either a faulty, irreducible neural circuit, or a faulty circuit and its mirror-symmetric counterpart [Figs. 1(a)–1(c)]. Graded engagement of both sides is allowed [Fig. 1(c)], but we show that only fully lateralized [Fig. 1(a)] or fully bilateral [Fig. 1(b)] solutions matter. In Sec. II B, we study complex computations that need the cooperation of several modules, each taking care of a different subtask [Figs. 1(d)–1(f)]. We call these emergent or complex phenotypes. We assess what optimality constraints towards lateralized [Fig. 1(d)] or bilateral [Fig. 1(e)] configurations operate now—finding, again, only all-or-nothing engagement. In Sec. II C, we juxtapose

both sets of optimality pressures (those operating on simple tasks and those on complex phenotypes) to illuminate how biases towards laterality or mirror symmetry might have changed as complex cognition evolved. We use the simplest possible model to guide our discussion, but we prove in the appendixes a broad generality of our results. An ample range of biologically and computationally meaningful choices should share the most salient constraints—notably, that only fully lateralized or completely bilateral configurations matter. Deviations from this should inform us about important scenarios not captured by our model.

## II. RESULTS

### A. Charting bilaterality and lateralization for simple tasks

Assume that some neural circuit has to solve a relatively simple task. Assume that this circuit is faulty, such that it fails with a probability  $\varepsilon$  at each attempt. Assume also that, thanks to our bilaterian body, we possess two such circuits: one at the left,  $L$ , and one at the right,  $R$  [squares in Figs. 1(a)–1(c)]. We refer to them as left and right units or circuits, and we say that they constitute a mirror-symmetric or bilateral module. We also say that this module can remain bilaterally symmetric or that it can become lateralized to some degree. By now, these modules are mathematical contraptions. In Appendix F, we briefly discuss the relationship of the mathematics in our model and feasible neural biology.

Does it pay off to keep both mirror-symmetric circuits? Two units together might perform a task more reliably. However, keeping each circuit running has a metabolic cost—as active neurons consume energy. In addition, if both units function simultaneously, they might interfere with each other—one side providing a wrong answer might spoil the other’s correct outcome. Some additional structure or mechanism is needed to transport and cross-check parallel activity [triangles in Figs. 1(a)–(c)]. Let us attempt to capture, with the simplest equation possible, all the costs and benefits of keeping just one unit of the module running [Fig. 1(a)] versus retaining both sides [Fig. 1(b)] versus keeping some intermediate engagement [Fig. 1(c)].

Whenever computing this task is required, unit  $L$  is switched on with a probability  $l \in [0, 1]$ , and unit  $R$  is activated with a probability  $r \in [0, 1]$ . These probabilities do not need to be independent. For a computation that needs to be carried out uninterrupted throughout the day, we can think of  $l$  and  $r$  as fractions of time that each unit stays active. We introduce a cost  $c$  paid for each occasion in which either unit is switched on, such that

$$C = c(l + r) \quad (1)$$

are the total expenses of running both units independently. Throughout this paper, when naming costs and fitness gains, lowercase letters denote benefits earned or costs

incurred as a task is attempted once. Uppercase symbols denote averages over repeated attempts with circuits engaged for fractions of time,  $l$  and  $r$ .

Here,  $C$  represents the metabolic cost of use. The structural existence of each unit should have a great cost as well—but it is paid independently of use and only once as the brain develops. Adding such additional costs does not alter our results qualitatively. Our costs could also depend, e.g., on the desired accuracy, such that lower error rates are more costly. One such case was partially explored in Ref. [7]. In Appendix E, we show that our results generalize to these and other broad scenarios.

The coordinating mechanism has an additional cost  $k$ , and we assume that it is paid whenever both bilateral units function simultaneously. Hence,

$$K = klr \quad (2)$$

are the total coordination expenses. This captures the metabolic expenses of the coordinating structure but also losses due to interference between both units that result in a faultier functioning. We can think of this last possibility as an average loss due to insufficient coordination. This would happen only when both units are active—and thus can be absorbed within  $k$ .

As for the benefits, we assume that some fitness,  $g^S$ , is gained if and only if the task is successfully implemented. Our results are more general—e.g., they hold if an imperfect implementation still contributes some fitness (see Appendix E 2). Focusing on the simplest Ansatz, we obtain

$$G^S = g^S[(1 - \varepsilon)l(1 - r) + (1 - \varepsilon)(1 - l)r + (1 - \varepsilon^2)lr]. \quad (3)$$

Note the convention of lowercase and uppercase letters again, as well as the superscript in  $g^S$  and  $G^S$  (that stands for simple task), which we did not use for costs. Simple tasks and emergent phenotypes will report different fitness gains. Operating costs might also differ (as we discuss below), but this is not as relevant for our results.

The first term within square brackets in Eq. (3) is the probability that  $L$  is active,  $l$ , and works properly,  $1 - \varepsilon$ , and that  $R$  is switched off,  $1 - r$ . The second term is the probability that  $R$  is active and working properly and  $L$  is switched off. The third term is the probability that both units are active and at least one of them produces the correct answer. In this simplest Ansatz, we assume that the correct answer cannot be reached if neither unit reaches it on its own. Alternatively, two faulty but nearly correct answers could improve each other (especially thanks to the coordinating mechanism). To capture this possibility, the third term would display an alternative function of  $\varepsilon$  instead of  $1 - \varepsilon^2$  (see Appendix E 2). Returning to our simplest equation, the third term assumes that the coordinating

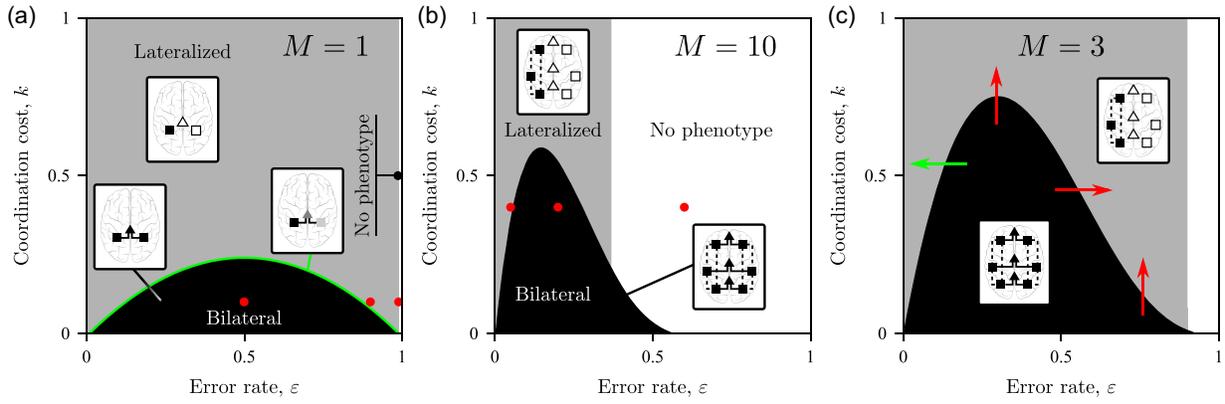


FIG. 2. Optimal lateral vs bilateral configurations. For  $(\varepsilon, k)$  within the white areas, the sought phenotype is so costly that it never pays off. Within the gray regions, it is optimal to lateralize. Within the black areas, bilaterality is preferred. (a) Model for simple cognitive tasks with  $g^S = 1$ ,  $c = 0.01$ . This is the only scenario in which we find graded solutions but only in a negligible part of phase space (green boundary between regions). Utility functions at the red dots are plotted in Fig. 5. (b) Model for emergent phenotypes with  $g^E = 1$ ,  $c = 0.01$ ,  $M = 10$ . Utility functions at the red dots are plotted in Fig. 7. (c) Trajectories on a morphospace (e.g., increased coordination costs,  $k$ , due to pathology or larger brains, vertical arrows; or increased fallibility due to aging, rightward arrows), which might prompt changes in optimal configurations.

mechanism can obtain the correct outcome if it has been produced by at least one unit. This still allows for important synergy—e.g., units failing 50% of the time still work together with about 75% fidelity [Fig. 10(e)].

Subtracting costs from fitness gains renders a utility function:

$$\begin{aligned} \rho^S(l, r) &\equiv G^S - C - K \\ &= g^S(1 - \varepsilon)[l(1 - r) + (1 - l)r + (1 + \varepsilon)lr] \\ &\quad - c(l + r) - klr. \end{aligned} \quad (4)$$

Given some values of our model parameters— $g^S$ ,  $c$ ,  $k > 0$ , and  $\varepsilon \in [0, 1)$ —the maximum utility as a function of  $(l, r)$  tells us the optimal configuration of our module—i.e., which activity each mirror-symmetric side should keep.

We draw optimal configurations in a map (or morphospace) that charts how optimality changes as parameters vary. Figure 2(a) shows such a map for Eq. (4) with fixed  $g^S = 1$  and  $c = 0.01$  and varying  $\varepsilon \in [0, 1)$  and  $k \in [0, 1]$  (see Appendix A for the mathematical derivation). For  $\varepsilon > 1 - c/g^S$  (thin white stripe at the right of the map), both units are so faulty that the fitness gain does not pay off enough to keep the module. Such tasks are not viable. In the large gray area, coordination is costly enough that it pays off to lateralize, indistinctly keeping  $L$  or  $R$  and shutting down the opposite side. In the black region, it is always convenient to keep both sides. At the boundary separating both regions (green curve), the optimal solution has one unit that is always active and the other one is active any arbitrary fraction of time. This is the only graded configuration (in which both units are not either completely off or completely on)—otherwise, Eq. (4) only has all-or-nothing solutions.

The boundary between bilateral and lateralized solutions is a parabola with a maximum of  $k = g^S/4 - c$  located at  $\varepsilon = 1/2$ . Note that if  $c > g^S/4$ , the bilateral configuration disappears [Fig. 6(c)], so the lateralized solution is preferred for any  $(\varepsilon, k)$ . If  $c = 0$ , the maximum of the parabola is at  $g^S/4$ . This imposes a very stringent limit: Coordination costs can never be larger than a fourth of the fitness contributed.

## B. Charting bilaterality and lateralization for emergent phenotypes

Conceive now a complex phenotype that, in order to be implemented, needs to recruit a series of brain regions (akin to our modules), each one carrying out a different subtask. A good example is human language (see Appendix F), which requires the successful functioning of Broca’s, Wernicke’s, and other regions [8–15] [Fig. 11(a)]. Full-fledged language only emerges if all regions perform correctly. Failure at any critical subtask results in specific pathology related to the malfunctioning area. Let us call such cognitive ability an *emergent* phenotype. Note our use of the word “emergent” here: We will not study how such phenotypes came together. Rather, by “emergent” we mean that the phenotype is compounded of simpler parts. Assuming that each subtask can be implemented as before (i.e., by either unit within a mirror-symmetric module), then we ask again, when is it favorable to lateralize and keep just one side active [Fig. 1(d)], keep all bilateral circuits functioning [Fig. 1(e)], or have them running at some intermediate level [Fig. 1(f)]?

Let us assume an emergent phenotype that involves  $M$  modules—i.e., it requires the correct implementation of  $M$  subtasks. We take  $M$  as a proxy for the phenotype’s cognitive complexity. Let us assume that each module incurs similar costs as before:

$$\begin{aligned} C &= c(l+r) \cdot M, \\ K &= klr \cdot M. \end{aligned} \quad (5)$$

Specific values  $c$  and  $k$  might change with respect to simple phenotypes (see section above), but the functional dependency is the same, as a cost  $c$  is incurred by running circuits individually and another  $k$  by coordination within each module. As before, we can interpret  $l$  and  $r$  as the average time that  $L$  or  $R$  units are active. Alternatively, we can say that a fraction  $l$  of the  $M$  left units is always switched on (and similarly for the right side). We assume that all subtasks are equally costly. In Appendix E 1, we explore different costs, error rates, and engagements for each module. This gives rise to a richer phenomenology, with lateralization for each module depending on other modules' parameters and configurations, but all the main results that follow for the homogeneous case are retained.

Regarding the fitness gain, now it is only cashed in if all independent subtasks are successful; thus,

$$G^E = g^E \cdot M[(1-\varepsilon)l(1-r) + (1-\varepsilon)(1-l)r + (1-\varepsilon^2)lr]^M. \quad (6)$$

The superscript in  $g^E$  and  $G^E$  stands for an emergent phenotype. Here, we see the same probability of implementing each subtask as before, now raised to the  $M$ th power—giving us the likelihood that no subtask is lacking. We assume a fitness gain  $g^E \cdot M$  (we could absorb the  $M$  within  $g^E$ , but it is not convenient).

We can now define the following utility function:

$$\begin{aligned} \rho^E(l, r) &\equiv G^E/M - C/M - K/M \\ &= g^E(1-\varepsilon)^M[l(1-r) + (1-l)r + (1+\varepsilon)lr]^M \\ &\quad - c(l+r) - klr. \end{aligned} \quad (7)$$

Figure 2(b) charts its optimal solutions with  $g^E = 1$ ,  $c = 0.01$ ,  $M = 10$ , and varying  $\varepsilon \in [0, 1)$  and  $k \in [0, 1]$  (see Appendix B for the mathematical derivation).

The area in which the phenotype fails to emerge [now,  $\varepsilon > 1 - \sqrt[M]{c/g^E}$ ; white region in Fig. 2(b)] is much wider than before, while the lateralization region (gray) has shrunk. We take relatively low running costs ( $c = 0.01$ ). If these costs become smaller ( $c \rightarrow 0$ ), the “no-phenotype” region would become negligible as  $\varepsilon > 1 - \sqrt[M]{c/g^E} \rightarrow 1$ . This is noteworthy, even though we expect realistic scenarios to have non-negligible costs ( $c > 0$ ). Even if we do not approach the  $c \rightarrow 0$  limit, complex phenotypes contribute a greater fitness than simpler tasks (i.e.,  $g^E \gg g^S$ ). Since we set  $g^E = 1$  to generate our maps, we should rescale running costs accordingly, resulting in smaller  $c$ .

The region where bilaterality is optimal has shifted to lower  $\varepsilon$  and deformed with respect to the parabola. Unlike

before, graded engagement is not optimal along the boundary. Instead, both the lateralized and fully bilateral configurations are simultaneously optimal at the boundary of the black and gray regions.

In Appendix B, we prove that the curve describing this boundary has only one maximum, so all morphospaces have a similar shape as  $M$  changes. As the emergent phenotype's complexity increases (i.e., as more modules need to be recruited, thus  $M$  grows), the no-phenotype region grows (as per  $\varepsilon > 1 - \sqrt[M]{c/g^E}$ ). The bilaterality region shifts further left while its peak reaches higher in the  $k$  axis. For simple tasks, there is a harsh limit ( $k < g^S/4$ ) for bilaterality. For complex phenotypes, a much higher coordination cost can be tolerated.

### C. Optimality pressures as new phenotypes emerge

How do optimality constraints change as complex phenotypes evolve? The morphospaces derived above give us static pictures at two endpoints of an evolutionary process. Consider the set of neural modules in early hominins that constituted the precursors of language areas (Broca's, Wernicke's, etc.). Let us assume that, before languages appeared, these modules worked independently—each tending to its own, relatively simpler task. Each module's optimality constraints would be captured by simple task morphospaces [Fig. 2(a)]. Now, take full-fledged human language. Some evolutionary process forced those precursor modules to work together to produce the emergent phenotype. This advanced cognitive ability depends upon the correct implementation of all subtasks; thus, optimality pressures now resemble the complex phenotype morphospace [Fig. 2(b)].

In this section, we juxtapose these two static pictures to visualize how optimality constraints might change the preferred configuration as complex cognition evolves. This addresses our main question: Are there evolutionary pressures towards brain lateralization associated with increased cognitive complexity? Some issues will linger regarding the evolutionary process: In what order were the modules forced to cooperate? In the example above, precursor modules likely started working together well before language appeared. Some complex phenotypes might have evolved more gradually than others—perhaps forcing modules into a hierarchy. Then, successive morphospaces would be relevant at different stages (e.g., first  $M = 3$ , then  $M = 5$  modules, etc.). Evolutionary paths would need to adapt to changing optimality constraints. In addition, as modules are brought together, the circuitry within might become fine-tuned to their subtasks—changing their reliability  $\varepsilon$  and operating costs  $c$  and  $k$ . This would amount to trajectories across morphospaces, which might prompt additional changes in optimal configurations. Exploring these scenarios is beyond the scope of this paper, but our model is a guide to navigating such evolutionary paths.

Finally, morphospaces for emerging phenotypes revealed broad regions at which complex cognition is not viable according to the cost-benefit calculation [white areas in Fig. 2 (b)]. This is alleviated by assuming that, after the appearance of a complex phenotype, simpler tasks still contribute some fitness on their own [Fig. 2(c) shows one such resulting morphospace]. This is also more realistic. Think about how some language regions can process symbol patterns unrelated to language without activating all language modules. In this section, we review cases in which simpler phenotypes coexist alongside emergent ones. Our results are quite general, as we explore in Appendix E. There we consider the effect of heterogeneous costs, error rates, and engagements across modules (Appendix E 1); increased synergy across sides of the brain and arbitrary dependencies on error rates (Appendix E 2); nonlinear dependencies on unit engagement (Appendix E 3); and dependency across modules in implementing their simpler subtasks (Appendix E 4).

### 1. Complex phenotypes implemented alongside simple ones

Complex phenotypes evolve upon a previously existing computational substrate. In it, modules exist that already implemented their own, individual, simpler task. In the example above, we mentioned *precursors* of language

regions. An emerging phenotype is likely not engaged at all times. When not engaged, the constituting modules are freed up to perform their more ancient, simpler tasks, which can still contribute some fitness. In our example, recognition of nonlinguistic symbol patterns can be carried out by some language areas without engaging, e.g., the semantic map.

Let us assume an emergent phenotype that is engaged a fraction  $\tau \in [0, 1]$  of the time. Let us assume that ancient, simpler tasks are engaged for the remaining time,  $1 - \tau$ . Averaging the corresponding fitness gains accordingly, we compute the resulting morphospaces (see Appendix C for details). Figure 3 explores these optimality constraints taking  $g^S = 1 \cdot (1 - \tau)$  and  $g^E = 2 \cdot \tau$  (i.e., the emerging phenotype allows us to gain twice as much fitness per module). We illustrate emerging phenotypes of distinct complexity [i.e., number of modules recruited,  $M$ ; Figs. 3(a)–3(d)] and emerging phenotypes that are engaged for different fractions of time,  $\tau$  [Figs. 3(e)–3(h)]. These constraints are juxtaposed upon the map for simpler tasks, revealing how new optimality biases operate when the emerging phenotype appears.

In Fig. 3, we find regions marked **B**→**B** (black) and **L**→**L** (light gray) in which, respectively, the bilateral and lateralized solutions are optimal, both for simple and complex phenotypes. In such cases, the evolution of complex cognition does not alter the optimal configuration of the recruited neural modules.

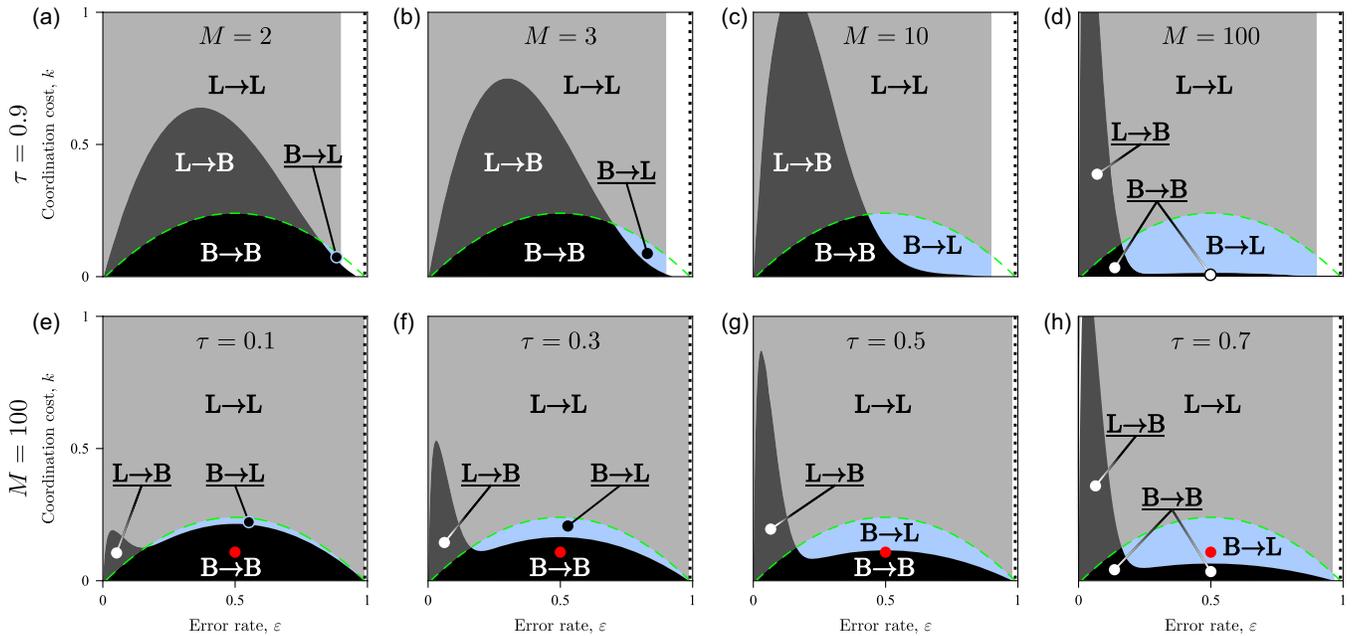


FIG. 3. Evolutionary paths to complex phenotypes that emerge upon substrates that retain their simpler functions, with  $g^S = 1$ ,  $g^E = 2$ , and  $c = 0.01$  in all cases. (a)–(d) We explore an emergent phenotype that occupies the neural substrate 90% of the time ( $\tau = 0.9$ ). As the complexity of the emergent phenotype increases [(a)  $M = 2$ , (b)  $M = 3$ , (c)  $M = 10$ , and (d)  $M = 100$ ], it becomes unavoidable that the mirror symmetry breaks apart. In panels (e)–(h), we explore a notably complex phenotype ( $M = 100$ ) that gradually increases the fraction of time during which it makes use of the neural substrate [(e)  $\tau = 0.1$ , (f)  $\tau = 0.3$ , (g)  $\tau = 0.5$ , and (h)  $\tau = 0.7$ , with (d)  $\tau = 0.9$  completing the progression]. This resembles developmental situations in which higher brain functions are assembled gradually and displace simpler computations in the same neural substrate. A circuit sitting where the red dot is would become lateralized by such a process.

We also find regions marked  $\mathbf{B} \rightarrow \mathbf{L}$  (blue). In them, bilateral symmetry is optimal for simpler tasks but it breaks apart as more complex phenotypes emerge. These  $\mathbf{B} \rightarrow \mathbf{L}$  regions grow larger for more complex phenotypes (i.e., when more modules  $M$  are recruited), as the  $\mathbf{B} \rightarrow \mathbf{B}$  region vanishes [Figs. 3(a)–3(d)]. The same thing happens for complex phenotypes that are engaged for larger fractions of time,  $\tau \rightarrow 1$  [Figs. 3(e)–3(h)]. In other words, the emergence of more complex cognition entails increased evolutionary pressure for brain lateralization, thus settling the main question of this paper.

Additionally, Fig. 3 shows salient regions labeled  $\mathbf{L} \rightarrow \mathbf{B}$  (dark gray). In them, a lateralized solution is optimal for simple tasks, but bilaterality is preferred for the emergent phenotype. This is an evolutionary pressure towards recovering lost bilaterality. Depending on an organism's history, the symmetric counterpart of a lateralized circuit might have been lost (or diverged towards other tasks—see below). Recovering bilaterality might not be possible. The organism is then stuck with the suboptimal lateralized solution—a frozen accident. Alternatively, the evolutionary pressure could foster the appearance of a duplicate that is not mirror symmetric—we focus on lateralization, but our results are valid for duplicated circuits, in general. The evolution of such duplicates has been observed in the

mammalian brain [38]. Our  $\mathbf{L} \rightarrow \mathbf{B}$  regions suggest ample pressures favoring this evolutionary route.

We assume that all modules have the same reliability  $\varepsilon$ , that they incur exactly the same costs  $c$  and  $k$ , and that all simpler tasks report the same fitness  $g^S$ . This is unlikely in real brains. Different modules will perform with different accuracy, effort, and rewards. We show (Appendix E 1) that our results are valid when modules are heterogeneous. Figures 10(a)–10(d) show, indeed, that this heterogeneity increases the pressure towards lateralization.

## 2. Segregating functions

When exposed to languages, babies show engagement in both sides of the brain (e.g., Broca's area and its contralateral homologue). As they mature, this symmetry is broken, and only one side (usually the left) is engaged during language input or production [39]. Early loss of this hemisphere can result in the development of language regions on the opposite side [40], indicating that contralateral homologues retain (at least temporarily) a potential for language. But language recovery is usually not possible after hemispherectomy or stroke when they happen later in life. One possibility is that contralateral homologues become specialized in other tasks—e.g., processing nonsyntactic musical patterns [41].

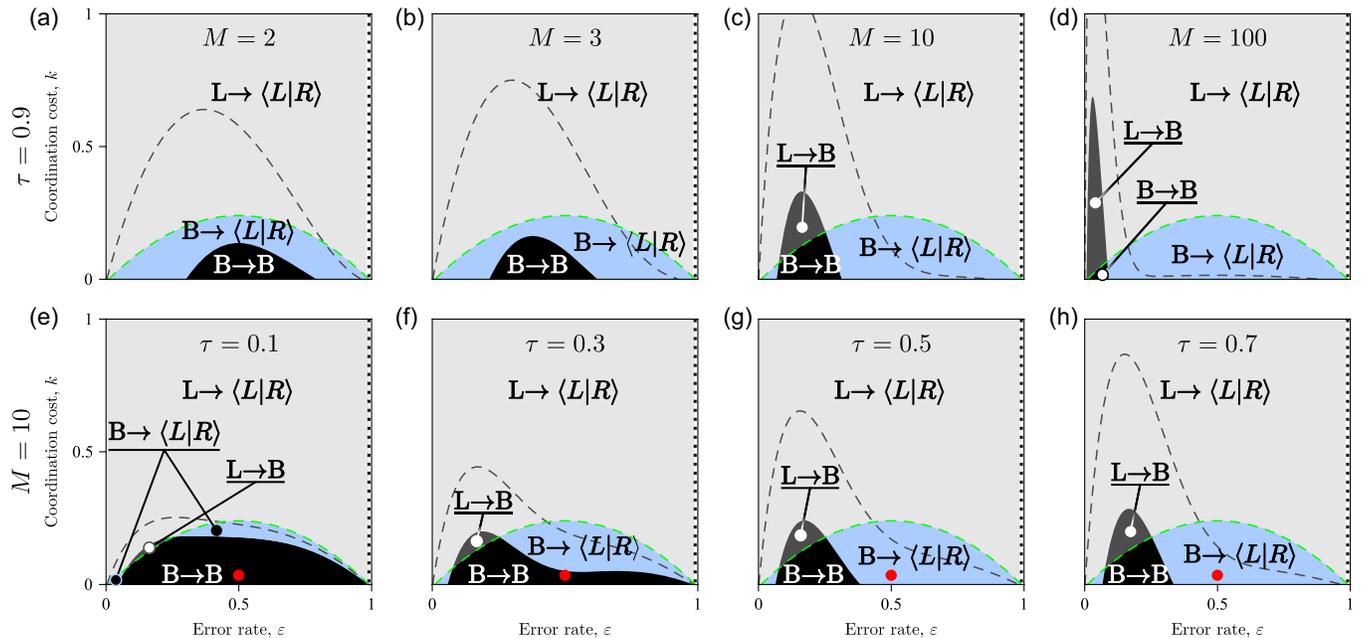


FIG. 4. Optimalty pressure towards function segregation. In all cases,  $g^S = 1$ ,  $g^E = 2$ , and  $c = 0.01$ . Gray dashed lines represent the boundary between bilateral and lateralized solutions from Fig. 3. In panels (a)–(d), as before, we explore emerging phenotypes that occupy the neural substrate 90% of the time ( $\tau = 0.9$ ). Again, as the complexity of the emerging phenotype increases [(a)  $M = 2$ , (b)  $M = 3$ , (c)  $M = 10$ , and (d)  $M = 100$ ], it becomes unavoidable that bilateral circuits lose their mirror symmetry. This time, when bilaterality is lost, each mirror symmetric counterpart becomes specialized in either the simple or complex tasks ( $\mathbf{B} \rightarrow \langle \mathbf{L}|\mathbf{R} \rangle$ ). We explore a moderately complex phenotype ( $M = 10$ ) that gradually increases the fraction of time during which it makes use of the neural substrate [(e)  $\tau = 0.1$ , (f)  $\tau = 0.3$ , (g)  $\tau = 0.5$ , (h)  $\tau = 0.7$ , with (c)  $\tau = 0.9$  completing the progression]. Again, this allows us to model developmental situations in which higher brain functions are assembled gradually and displace simpler computations in the same neural substrate. A circuit sitting where the red dot is would become lateralized by such a process.

This exemplifies (over development, rather than evolution) a mirror symmetric module that becomes lateralized by devoting each side to a different function. When does such a configuration pay off?

Figure 4 shows optimality pressure towards this lateralized-segregated solution (denoted  $\langle L|R \rangle$ ; see Appendix D for mathematical details). This is amply preferred both as the complexity  $M$  of the emergent phenotype increases [Figs. 4(a)–4(d)] and as the fraction  $\tau$  of time devoted to the complex task grows [Figs. 4(e)–(h)].

The  $L \rightarrow \langle L|R \rangle$  pathway (light gray) starts with already-lateralized modules for simple tasks. If the unused counterparts persist through evolution, they might be available for  $\langle L|R \rangle$  recruitment. However, those unused circuits might have been lost over the course of evolution—hence, this pathway would not be straightforwardly available. Again, since our results are valid for sets of duplicated circuits (not only mirror-symmetric modules), this segregated solution is yet another evolutionary pressure towards the duplication of existing structures within one hemisphere [7,38].

The  $B \rightarrow \langle L|R \rangle$  pathway (blue) is more amply preferred the more complex the emergent phenotype is. This shows, again, how increased cognitive complexity is a strong evolutionary driver of lateralization—and (since our results extend beyond bilaterality) of symmetry breaking in the brain.

The literature discusses how function segregation can be convenient—e.g., to have specialized hemispheres that complex cognition can recruit from [26] or to allow more efficient packing [28]. Even if allowing segregation, bilaterality remains optimal in ample regions of parameter space. Note that segregation would not happen in the bilateral solution since it engages both mirror-symmetric circuits for simple and complex phenotypes. The persistence of bilateral designs enables the  $B \rightarrow B$  (black in Fig. 4) and  $L \rightarrow B$  [dark gray, Figs. 4(c)–(h)] pathways. In them, the fitness gained from combining faulty circuits with robust modules overcomes the advantages of segregating tasks into equally faulty circuits. The  $L \rightarrow B$  pathway is again a pressure towards duplication of existing circuits or (if still available) reenacting lost mirror symmetry. Variations of our model (Appendix E 2) show that increased within-module synergy results in broader  $B \rightarrow B$  and  $L \rightarrow B$  regions [Figs. 10(b)–10(d)].

### III. DISCUSSION

Since very early in the history of computer science, redundancy was acknowledged as an efficient strategy to perform computations with faulty parts [32–34]. The bilaterian body plan is a source of redundancy for many organs, including the central nervous system. Brains are equipped with mirror-symmetric duplicates of most cortical regions, ganglia, etc. This can result in more robust neural computations but can have an excessive metabolic cost.

We have developed a concise yet comprehensive mathematical framework to study the optimality of bilateral versus lateralized solutions. We have built a series of morphospaces that summarize optimal configurations as a function of (i) costs of running lateralized neural circuits independently, (ii) costs of coordinating efforts across sides of the brain, (iii) how error prone these units are, (iv) the fitness gained by successful neural computations, and (v) the complexity of the tasks at hand.

A first, strong result is that only all-or-nothing configurations are optimal in our framework: It is either better to engage both mirror-symmetric sides, just one lateralized circuit, or none at all. A graded engagement is never exclusively optimal. We might think that this happens because of our fairly stringent assumptions—e.g., the idea that modules are strictly independent, irreducible units; or the choice of implementation of synergies across the hemispheres [as given by the  $(1 - \varepsilon^2)lr$  term in Eq. (3)]. However, this result is very general. In Appendix E, we show that it holds (i) if we allow heterogeneous modules (i.e., modules that work together yet have different costs, levels of engagement, and error rates; Appendix E 1); (ii) if we allow different levels of synergies across hemispheres, or if we consider more arbitrary dependencies of the costs on the error rates (Appendix E 2); (iii) if we allow non-linear, monotonically increasing dependencies between the time that a module is engaged and its performance (e.g., because it learns, becoming more efficient; Appendix E 3); and (iv) if we allow modules to aid each other in implementing simpler tasks, even before emergence of a new, complex phenotype (Appendix E 4). Each of these variations is an effort to bring the model closer to the diverse biological scenarios that exist in nature. However, our assumptions remain stringent in many senses. The robustness of our results upon variations suggests that our approach (from the simplest towards more complex scenarios) is promising, and we propose that we can also use our optimality maps to investigate very detailed models (e.g., with realistic neurons, accounting for metabolism, etc.). Returning to the all-or-nothing engagement, while no assumption has changed this result in any model variations, it seems reasonable that neural circuits might be gradually engaged. A great follow-up question is: What change in our equations would result in graded engagement, and what would that mean for a real brain? A possible way forward is that, while our maps reveal *pressures* towards optimality, biology might ignore optimality, and historical contingencies might prevent us from reaching it.

Early findings of localized language functions suggested that lateralized brain activity was exclusive of complex human cognition [22–24]. This idea was debunked after finding lateralized activity in other animals, but the hypothesis that cognitive complexity might prompt brain-symmetry breaking has survived, with nuances. A solid mathematical understanding of such a mechanism has been missing.

Our model provides the lacking framework. We show mathematically how different optimality pressures towards lateralization are fostered by increasing cognitive complexity. This happens in scenarios in which complex phenotypes emerge alongside simpler tasks in a hierarchical (Fig. 3) or segregated (Fig. 4) fashion. With these well-grounded mathematical results, we conclude that the evolution of more complex cognition can be a paramount driver of brain lateralization.

However, we also provide strong evidence in the opposite direction: For large combinations of model parameters, there exists a pressure upon formerly lateralized circuits to evolve a duplicate again. These scenarios are different from those for which it is optimal to lose bilaterality. Hence, emerging complex cognition can act as a source of new symmetries or break older ones, depending on properties of the neural substrate (e.g., its fallibility or metabolic needs) and other conditions (e.g., complexity of the emerging phenotype). While we focus our discussion on mirror symmetry, our results apply to any sets of duplicated neural structures. When pressure to develop redundancy is present, it might be more parsimonious to create duplicity within the same hemisphere. Evidence of such duplicates in the mammalian brain has recently been described [7,38]. The alternative (re-recruiting the actual mirror symmetric circuitry) might be impossible if this circuitry has diverged over evolutionary time.

We visualize optimal configurations as morphospaces over model parameters. Morphospaces were introduced to describe the shape of shells as a function of factors that affected their formation [42,43]. They have been expanded to map complex systems [44–50], including how optimality guides the evolution and development of neural substrates [37,51,52]. These morphospaces remind us of phase diagrams that map solid, liquid, etc. phases of matter subjected to different physical conditions. Similarly, changes in optimal configurations as we move around our morphospaces or as we juxtapose different conditions remind us of phase transitions. It seems natural to extend these tools and concepts from statistical physics as we do here.

The ultimate goal of these morphospaces is to portray real-world systems and explain actual phenomenology. Let us try this with our mathematical framework:

- (i) Neural damage or pathology could alter the operating parameters of our modules. A damaged circuit can become faultier (increased  $\varepsilon$ ) or more costly to engage or coordinate (growing  $c$  and  $k$ ). These changes could push bilateral circuits outwards from a bilaterality region [Fig. 2(c), red arrows]. Aging could also lead to increased fallibility and costs. We should expect more asymmetry with age or pathology, as is the case [5,53,54]. Even if lateralization becomes optimal due to fallibility or cost changes, mature brains might retain mirror-symmetric configurations as if stuck on a frozen accident. In such
- cases, blocking one side of the brain might help achieve the optimal configuration. Noninvasive techniques such as transcranial magnetic stimulation (TMS) might help us test this possibility.
- (ii) Larger brains pay higher coordination costs due to limits on callosal information transfer. In our model, increased  $k$  moves mirror-symmetric circuits towards lateralized configurations [Fig. 2(c), vertical arrows]. We expect increased asymmetry in larger brains, as is the case [5,55]. This agrees with more asymmetric brains presenting less or thinner trans-callosal fibers [56]. These brains might have renounced some coordination efforts and embraced lateralization. Tasks that demand short reaction times should show similar effects—they would penalize interhemispheric communication delays (i.e., larger  $k$ ). Our model subsumes this route to lateralization, which has been explored before [27].
- (iii) Hemisphere dominance is the process through which, while both sides are engaged in some function, one side takes a leading role—often acting as a controller or coordinator of both sides. Handedness is a paramount example. There is a strong bias towards mirror symmetry due to our bilateral body [Fig. 11(b), item 1]. However, a trend towards lateralization exists in mammals, with handedness increasing with behavioral complexity [18–21]. Studies of unilateral hemiplegia further show that the dominant hemisphere is needed for complex movement of the unaffected hand [22,57] (while the dominated side is not indispensable). Dominance is observed in other neural systems such as visual processing [16,17] or the Theory of Mind network [58]. This suggests graded engagement of the dominant side, which should be rare after our optimality constraints. However, our results apply to circuits that carry out *exactly* the same computations. They do not preclude a circuit commanding the other or even delegating specific, unshared tasks on it. Indeed, some routes to lateralization (e.g.,  $B \rightarrow \langle L|R \rangle$ ) might promote such controller-controlled configurations.
- (iv) Human language is the most paradigmatic example of higher-brain-function lateralization. Language usually involves left hemispheric regions [8–15]. Recent fMRI evidence shows that language engagement starts out as more symmetric in babies and becomes fully lateralized as children grow [39]. Similar trajectories can be seen in our model: Take neural circuits sitting at the red dots in Figs. 3(e)–3(h) and 4(e)–4(h), for example. Within this framework, as language would gradually recruit its neural substrate for longer times (i.e., increasing  $\tau$ ), the initial bilateral configuration would become suboptimal. This example illustrates how our model naturally accommodates

such developmental pathways—we cannot determine if this is the actual process during language maturation.

- (v) Our morphospaces predict that nearing perfect performance [ $\varepsilon \rightarrow 0$ , Fig. 2(c), green arrow] always results in lateralization. Duplicated circuits are superfluous, yet costly, if  $\varepsilon = 0$ . Musicians with perfect pitch exemplify this scenario: They have increased asymmetry in the planum temporale, notably owed to the reduction of the nondominant side for this task [2,59–61].
- (vi) Our model can be extended to account for other lateralization mechanisms, such as hemisphere specialization [26] or optimal packing [28]. Both these proposals are qualitative. We have built a very general quantitative framework that allows us to mathematically understand how these mechanisms might operate.

Empirical measures of brain asymmetry are costly and scarce [5,6]. In the examples just provided, qualitative observations match scenarios from our model. Efforts should follow to bring quantitative empirical measurements to our theoretical framework. Measuring model parameters in real neural circuits seems difficult, but very realistic compartment models [62] might help us simulate exact real-world conditions.

Alternatively, we can induce transitions from bilaterality to lateralized solutions in experimental setups, quantifying the transition thresholds, and thus constraining model parameters. This might be feasible with neuronal preparations *in vitro*, or even in behaving animals, including humans. We could manipulate task complexity while monitoring neural activity, or we could interfere with cortical regions to raise error rates—e.g., using TMS. In that sense, fMRI studies have recently been used to track brain lateralization during language development [39] or to map how complex cognitive tasks span a space of abstract neural activity depending on task novelty or on learning [52]. It seems a straightforward extension to combine both approaches and measure mirror symmetry and lateralization in the brain as a function of task complexity, thus seeking an empirical test of our model. Such an approach would rely on the brain behaving optimally, in the sense that it would renounce mirror symmetry if it became suboptimal. While this can be argued over evolutionary time (because Darwinism likely penalizes suboptimal designs), nothing guarantees that brain activity will behave optimally in a short timescale. However, such experiments might also allow us to measure performance and metabolic stress across participants and check how they correlate with optimal configurations (e.g., participants retaining suboptimal mirror symmetry for more complex tasks might perform the worst in a task).

We tried to make our mathematical framework as general as possible, but, unavoidably, some costs and effects have been left out. Future models should explore how our

morphospaces change as new aspects come into play. Some of the omissions might actually make our results more robust. For example, as complex phenotypes emerge, we have not demanded that all subtasks be implemented on the same side. However, we observe clear constraints towards lateralization. Including a penalty for interhemispheric communication should strengthen this trend. We have not discussed structural and lasting costs either. For us, building a neural circuit is free; we only pay to keep it running. Such additional costs should exacerbate some results—e.g., by making lateralization more definitive, which is relevant for the  $\mathbf{L} \rightarrow \mathbf{B}$  route. This could introduce path dependency, which reminds us of hysteresis in physical phase transitions. Other structural constraints might favor mirror symmetry. We mentioned our bilateral body. It should be feasible to modify our model accordingly and see how the morphospaces are updated.

Finally, our model is not only agnostic regarding bilaterality versus other sources of redundancy. It is also independent of the computational substrate. Our results should matter for designing efficient computing devices. In such cases (such as a chip that could choose to engage several microprocessors depending on task complexity), it should be helpful to extend our framework to arbitrary redundancies (i.e., not just two symmetric circuits). A similar modeling might be useful to study gene duplication [63,64], especially as more computational and cognitive approaches to the functioning of cells are explored. To achieve this, it might be convenient to expand our mathematical framework. The field of redundant coding offers interesting mathematical tools grounded in information theory [65,66] that have been used on empirical questions [67]. It should be possible to ground our model within this framework and perhaps use its concepts to expand our results and to understand, in terms of information flows, how our optimality constraints operate.

## ACKNOWLEDGMENTS

The author thanks Susanna Manrubia for support. We acknowledge enlightening discussions with Ricard Solé at the Santa Fe Institute (and elsewhere) and with Susanna Manrubia and her extended group at the Spanish National Center for Biotechnology (CNB), the Carlos III University, and other institutions in Madrid, Spain. We received funding from the Spanish National Research Council (CSIC) and the Spanish Department for Science and Innovation (MICINN) through a Juan de la Cierva Fellowship (IJC2018-036694-I), and from the Jesús Serra Foundation (Grant No. FJSCNB-2022-12-B).

## APPENDIX A: MAXIMA OF THE UTILITY FUNCTION FOR SIMPLE PHENOTYPES

In the simplest scenario, the average fitness gain reads

$$G^S = g^S(1 - \varepsilon)[l(1 - r) + (1 - l)r + (1 + \varepsilon)lr], \quad (\text{A1})$$

and the costs are

$$\begin{aligned} C &= c(l+r), \\ K &= klr, \end{aligned} \quad (\text{A2})$$

which yields the utility function

$$\begin{aligned} \rho^S(l, r) &\equiv g^S(1-\varepsilon)[l(1-r) + (1-l)r + (1+\varepsilon)lr] \\ &\quad - c(l+r) - klr. \end{aligned} \quad (\text{A3})$$

Given fixed values for our model parameters [ $g^S > 0$ ,  $c > 0$ ,  $k > 0$ , and  $\varepsilon \in [0, 1]$ ], we try to find maxima of Eq. (A3) as a function of our model variables  $l \in [0, 1]$  and  $r \in [0, 1]$ . Note how we have restricted our parameters: Without loss of generality, we could take  $g^S = 1$  and normalize all costs accordingly. It makes sense that  $c, k < g^S$ ; otherwise, it is never favorable to implement such a task (but stronger constraints will arise). Finally, the error rate takes values within  $\varepsilon \in [0, 1]$ , but we kept the interval open,  $\varepsilon \in [0, 1)$ , to avoid some algebraic problems. However, this limit is not so interesting.

To start searching for maxima, we look at the derivatives with respect to  $l$  and  $r$ :

$$\begin{aligned} \rho_l^S &\equiv \frac{\partial \rho^S}{\partial l} = (1-\varepsilon)g^S - c - r[b(1-\varepsilon)^2 + k], \\ \rho_r^S &\equiv \frac{\partial \rho^S}{\partial r} = (1-\varepsilon)g^S - c - l[b(1-\varepsilon)^2 + k]. \end{aligned} \quad (\text{A4})$$

These equations are straight lines as a function of  $r$  or  $l$ , respectively. Thus, given fixed model parameters, there is only one point at which both  $\rho_l^S$  and  $\rho_r^S$  can become 0:

$$(l, r) = \left( \frac{(1-\varepsilon)g^S - c}{g^S(1-\varepsilon)^2 + k}, \frac{(1-\varepsilon)g^S - c}{g^S(1-\varepsilon)^2 + k} \right). \quad (\text{A5})$$

Taking second derivatives

$$\rho_{ll}^S \equiv \frac{\partial^2 \rho^S}{\partial l^2} = 0 = \frac{\partial^2 \rho^S}{\partial r^2} \equiv \rho_{rr}^S \quad (\text{A6})$$

and

$$\rho_{lr}^S \equiv \frac{\partial^2 \rho^S}{\partial l \partial r} = -[g^S(1-\varepsilon)^2 + k], \quad (\text{A7})$$

yields a discriminant

$$\Delta = \rho_{ll}^S \rho_{rr}^S - (\rho_{lr}^S)^2 = -[g^S(1-\varepsilon)^2 + k]^2. \quad (\text{A8})$$

This number is always negative, meaning that the singular point is a saddle. Since Eq. (A3) is a quadratic form, the saddle spans the whole  $l$ - $r$  plane—meaning that there are no global or local extrema as a function of our variables.

This result further implies that, if we look at Eq. (A3) within a bounded area of the  $l$ - $r$  plane, its maxima will always lie at the boundary (Fig. 5). Thus, we just need to evaluate  $\rho^S(l, r)$  along the outer circuit of the square  $[0, 1] \times [0, 1]$ . Because of the symmetry of the problem,  $\rho^S(l=0, r) = \rho^S(l, r=0)$  and  $\rho^S(l, r=1) = \rho^S(l=1, r)$ ; thus, we observe  $\rho^S(l, r=0)$  and  $\rho^S(l=1, r)$ .

For the first segment, we obtain

$$\rho^S(l, r=0) = l[g^S(1-\varepsilon) - c]. \quad (\text{A9})$$

As a function of  $l$ , this is a straight line that passes through the origin. If the slope is positive, the largest value of  $\rho^S(l, r=0)$  will be found at  $l=1$ . If the slope is negative, the largest value will be found at  $l=0$ . In such cases, both neural units are permanently shut off, meaning that it pays

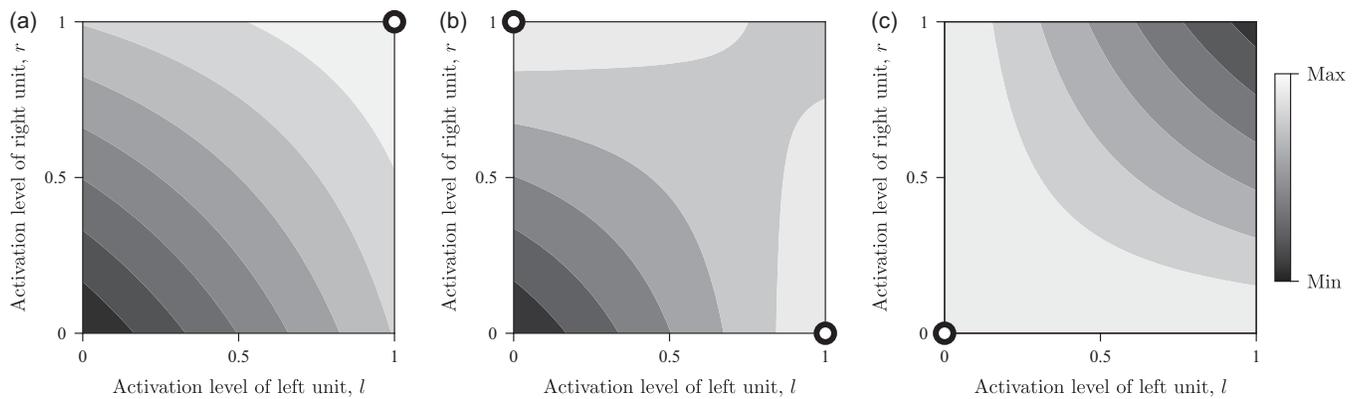


FIG. 5. Utility function for the simplest case. This utility function is a quadratic form with a negative discriminant—i.e., a saddle with no global extrema. Hence, maxima of the utility function within the region of interest are always at the enclosing circuit. With fixed model parameters  $g^S = 1$ ,  $c = 0.05$ , and  $k = 0.1$ , varying the error rate, we obtain different optimal solutions (marked by black circles): (a) bilaterality ( $\varepsilon = 0.5$ ), (b) lateralization ( $\varepsilon = 0.9$ ), and (c) no function ( $\varepsilon = 0.99$ ).

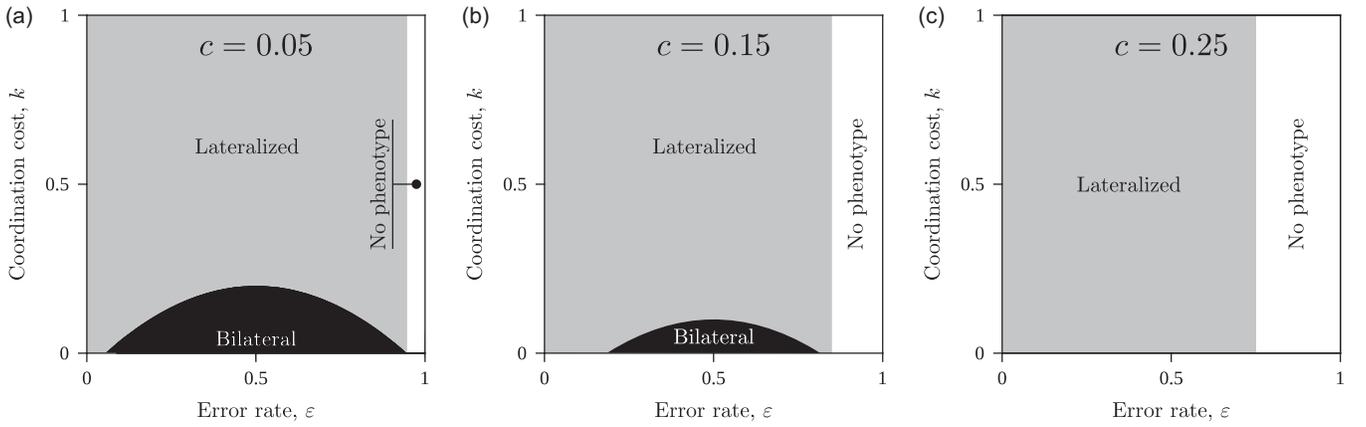


FIG. 6. Maps of optimal configurations for simple tasks. As a function of  $\varepsilon$  and  $k$ , a simple phenotype fails to emerge in the white areas because it reports too little fitness given the implementation costs. The task might be worth implementing with a lateralized (light gray) or bilateral (black) configuration. (a) Small costs of running independent units,  $c = 0.05$ . (b) Intermediate values of running independent units,  $c = 0.15$ . (c) If running independent units is larger than  $g^S/4$  (here,  $c = 0.25$ ), the bilateral configuration is not viable.

off not to implement this phenotype: It is more costly than the reward it produces. We have a negative slope if

$$c > g^S(1 - \varepsilon) \Rightarrow \varepsilon > 1 - \frac{c}{g^S}. \quad (\text{A10})$$

This condition is marked as a white rectangle in the  $\varepsilon - k$  maps in Figs. 2(a) and 6. Precisely at the point when the slope becomes zero, any level of activity  $l \in [0, 1]$  is equally optimal.

Evaluating the utility function along the second segment of interest we obtain

$$\rho^S(l = 1, r) = (1 - \varepsilon)g^S - c + r[g^S(1 - \varepsilon)\varepsilon - c - k], \quad (\text{A11})$$

which again is a straight line, now as a function of  $r$ . Take its slope,

$$m = g^S(1 - \varepsilon)\varepsilon - c - k. \quad (\text{A12})$$

If it is positive,  $\rho^S(l = 1, r)$  grows as a function of  $r$ , and its maxima within  $r \in [0, 1]$  is found at  $r = 1$ , meaning that it is optimal to keep both neural units active. If the slope is negative, the maxima is found at  $r = 0$ , meaning that it is convenient to keep only one unit (the left one in this case) switched on. Because of the symmetry of the problem, the antisymmetric solution also exists (right unit on and left unit off). If the slope is exactly zero, it is convenient to keep one neural unit always on, and it is indistinct whether the other one is on, off, or active at some intermediate level. For example,

$$m = 0 \Leftrightarrow k = -g^S\varepsilon^2 + g^S\varepsilon - c, \quad (\text{A13})$$

which is a parabola when represented in the  $\varepsilon - k$  plane. This parabola delimits the regions of parameter space in

which it is convenient to switch one or both neural units on [gray and black regions in Figs. 2(a) and 6].

The crossings of the parabola with the horizontal axis are given by imposing  $k = 0$  on Eq. (A13):

$$\varepsilon = \frac{1 \pm \sqrt{1 - 4c/g^S}}{2}. \quad (\text{A14})$$

These crossings are real numbers only if  $c < g^S/4$ . Thus, if the individual cost  $c$  is larger, a boundary separating bilateral from lateralized optima does not exist—only the lateralized solution survives [Fig. 6(c)]. The maximum of the parabola is  $g^S/4 - c$ , and it is always located at  $\varepsilon = 1/2$ . This imposes a rather harsh limit on the bilateral configuration: In the best-case scenario ( $c = 0$ ), for bilaterality to exist, coordinating efforts cannot be larger than a fourth of the total fitness gain. The lateralized configuration remains viable with much higher costs.

## APPENDIX B: MAXIMA OF THE UTILITY FUNCTION FOR STRICTLY EMERGENT PHENOTYPES

Next, we model complex phenotypes that emerge out of the interplay of  $M$  individual tasks, each of them solved in an atomic manner by a lateralized or bilateral module. Since now we need  $M$  such modules to implement the emerging phenotype, assuming that all modules incur similar costs, we have

$$\begin{aligned} C &= c(l + r) \cdot M, \\ K &= klr \cdot M. \end{aligned} \quad (\text{B1})$$

The fractions  $l$  and  $r$  can now be interpreted either as the time that modules of a side are active or as the fraction of units of the corresponding side that are always active.

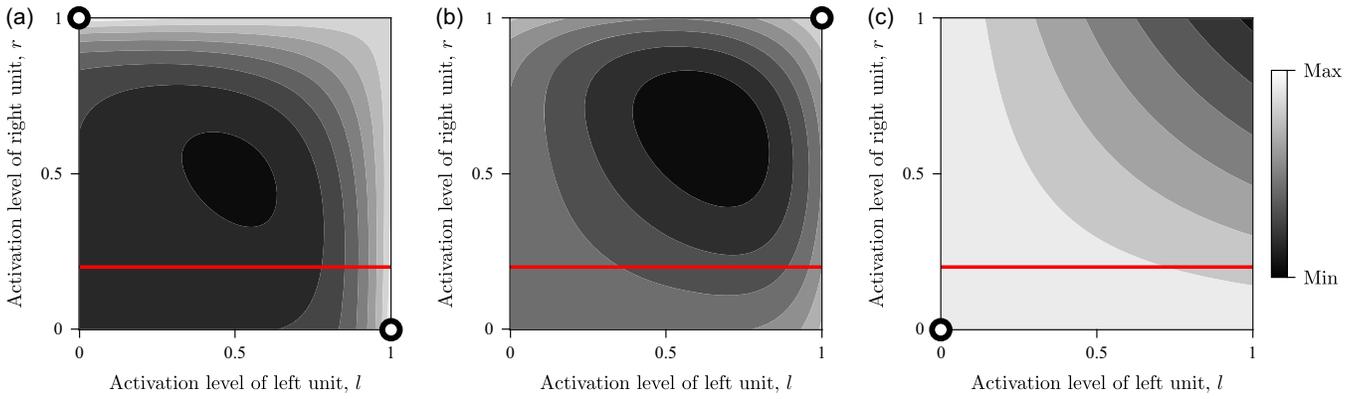


FIG. 7. Utility function for the strictly emergent phenotype. We prove in the text that this function has no maxima within the squared circuit (note, however, that global minima are present). Hence, maxima of the utility function within the region of interest are always at the boundary. With fixed model parameters  $g^E = 1$ ,  $c = 0.05$ , and  $k = 0.5$ , varying the error rate, we obtain different optimal solutions (marked by black circles): (a) lateralization ( $\epsilon = 0.05$ ), (b) bilaterality ( $\epsilon = 0.2$ ), and (c) no function emerges ( $\epsilon = 0.6$ ). The thick, red line indicates slices of the utility function with constant  $r = 0.2$  plotted in Fig. 8(b).

Regarding the gained benefit, this is only cashed in if the emergent task is implemented in full, for which we need all modules to work properly. The likelihood that this will happen is the product of the likelihood that each of the  $M$  units functions correctly:

$$G^E = g^E M \cdot (1 - \epsilon)^M [l(1 - r) + (1 - l)r + (1 + \epsilon)lr]^M. \quad (\text{B2})$$

Note that here we are modeling a fitness gain brought about strictly by the emergent phenotype—i.e., at this point, we are not considering the benefits from each of the individual tasks (we discuss this in more detail in the next section). We stipulate that the net fitness gain is  $g^E M$ , without loss of generality. We could absorb  $M$  within  $g^E$ —but we keep them separated for convenience. Hence,  $g^E$  captures the fitness gain per module recruited for the emergent phenotype.

It is also useful to introduce the polynomial

$$\begin{aligned} P(l, r) &\equiv l(1 - r) + (1 - l)r + (1 + \epsilon)lr \\ &= r + [1 - r(1 - \epsilon)]l \\ &= l + [1 - l(1 - \epsilon)]r. \end{aligned} \quad (\text{B3})$$

We have rewritten it three times, the last two just to show that, if  $r$  (respectively  $l$ ) are considered constant, the polynomial is a straight line as a function of the other variable.

We can now write the following utility function:

$$\begin{aligned} \rho^E(l, r) &\equiv G^E/M - C/M - K/M \\ &= g^E(1 - \epsilon)^M P(l, r)^M - c(l + r) - klr. \end{aligned} \quad (\text{B4})$$

Figure 7 shows this function within  $(l, r) \in [0, 1] \times [0, 1]$  in three distinct cases. Again, we are interested in finding

maxima of  $\rho^E(l, r)$  within that region of the  $l$ - $r$  plane. Therefore, let us compute its derivatives with respect to each of the variables:

$$\begin{aligned} \rho_l^E &= g^E(1 - \epsilon)^M M P(l, r)^{M-1} [1 - r(1 - \epsilon)] - c - kr, \\ \rho_r^E &= g^E(1 - \epsilon)^M M P(l, r)^{M-1} [1 - l(1 - \epsilon)] - c - kr. \end{aligned} \quad (\text{B5})$$

Because of the nonlinearity introduced by  $M$ , this case is more complicated than the one in Appendix A. To simplify it, let us look at a constant, fixed value of  $r$ , which defines a straight line parametrized by  $l$  in the  $l$ - $r$  plane (thick red lines in Fig. 7). Next, let us find singular points along this line:

$$\rho_l^E = 0 \Leftrightarrow P(l, r) = \sqrt[M-1]{\frac{c + kr}{g^E(1 - \epsilon)^M M [1 - r(1 - \epsilon)]}}. \quad (\text{B6})$$

The right-hand side of this equation is a constant for fixed  $r$ . Meanwhile,  $P(l, r)$  is a straight line as a function of  $l$  [as shown by Eq. (B3)]. This straight line can only cross the constant on the right-hand side once [Fig. 8(a)].

Let us call the term inside the root  $\gamma$ . Within the range of parameters and variables that we are interested in,  $\gamma$  is always positive, so its  $(M - 1)$ th root always exists. If  $M$  is even,  $M - 1$  is odd, and there are  $M - 1$  identical, positive roots on the right-hand side of Eq. (B6) [for  $M = 10$ , black horizontal line in Fig. 8(a)]. Thus, the equation has  $M - 1$  identical solutions. If  $M$  is odd,  $M - 1$  is even, and there are  $(M - 1)/2$  identical and negative roots and  $(M - 1)/2$  identical and positive roots [for  $M = 11$ , gray dashed horizontal lines in Fig. 8(a)]. Hence, the equation has two sets of  $(M - 1)/2$  identical solutions—one set based on the negative root and another one based on the positive root.

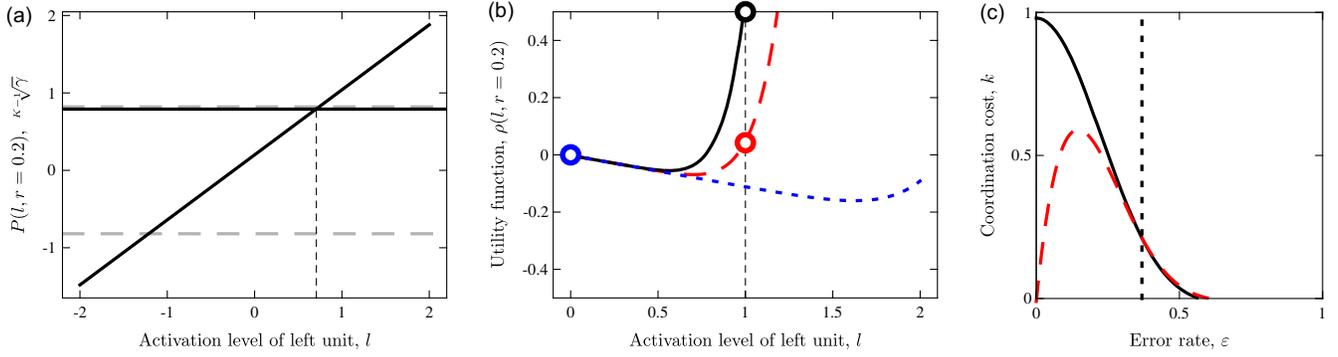


FIG. 8. Proof that strictly emergent phenotypes only have all-or-nothing solutions. (a) Polynomial  $P(l, r)$  and constant  $\varepsilon\sqrt[M]{\gamma}$ , which can only cross once, limiting the number of extrema of the utility function. (b) Utility function along lines of constant  $r$  from Fig. 7. Their maxima within our range of interest,  $l \in [0, 1]$ , can only be at  $l = 0$  or  $l = 1$ . (c) Resulting conditions for optimality of either configuration. The dashed vertical line marks when lateralized solutions pay off. The black solid curve indicates when bilateral configurations pay off. The red dashed curve delimits when bilateral configurations are preferred over lateralized ones.

Introducing the actual dependency of  $P(l, r)$  and solving for  $l$ , we obtain

$$l = \frac{\varepsilon\sqrt[M]{\gamma} - r}{1 - r(1 - \varepsilon)}. \quad (\text{B7})$$

In the denominator of this equality, we have a positive number (again, for our ranges of parameters and variables), so this does not change the sign of the solution for  $l$ . In the numerator, we subtract a positive number ( $r > 0$  in our range of interest) from  $\varepsilon\sqrt[M]{\gamma}$ . Thus, if we take the positive  $(M - 1)$ th root, we might get a positive or negative solution; if we take the negative root, then we obtain a negative solution even further away from 0.

Summarizing, Eq. (B6) might have either one positive and one negative, or two negative solutions for our range of parameters. This means that the derivative of the utility function changes sign at most once for  $l \geq 0$ ; thus, it has, at most, one extremum for positive  $l$ .

For  $l \rightarrow \infty$ , the utility function is dominated by

$$P(l, r)^M = [1 - r(1 - \varepsilon)]^M l^M, \quad (\text{B8})$$

which is positive and growing as a function of  $l$ . Thus, if there is one extremum for  $l > 0$ , it must be a minimum. This means that the maximum of the utility function along a line with fixed  $r$  and  $l \in [0, 1]$  must be found either at  $l = 0$  or  $l = 1$  [Fig. 8(b)].

This reasoning is valid for any constant  $r$  within our range of interest ( $r \in [0, 1]$ ) and also, symmetrically, for any constant  $l \in [0, 1]$  if we assume the utility function depends on  $r$  alone. Thus, the maxima of our utility function for  $(l, r) \in [0, 1] \times [0, 1]$  must be found around the contour and, specifically, either at  $(l, r) = (0, 0)$ ,  $(l, r) = (1, 0)$ ,  $(l, r) = (0, 1)$ , or  $(l, r) = (1, 1)$ . Because of the symmetry, if there is a maximum at  $(l, r) = (1, 0)$ ,

there is another one at  $(l, r) = (0, 1)$ . Thus, we only need to compare three points to solve our problem:

$$\begin{aligned} \rho^E(0, 0) &= 0, \\ \rho^E(1, 0) &= g^E(1 - \varepsilon)^M - c, \\ \rho^E(1, 1) &= g^E(1 - \varepsilon^2)^M - 2c - k. \end{aligned} \quad (\text{B9})$$

If the fully lateralized solution ever pays off,  $\rho^E(1, 0) > 0$ , we would obtain

$$c < g^E(1 - \varepsilon)^M \Rightarrow \varepsilon < 1 - \sqrt[M]{\frac{c}{g^E}}, \quad (\text{B10})$$

which traces a straight vertical line in the  $\varepsilon - k$  map [dotted black line in Fig. 8(c)].

If the bilateral solution ever pays off [ $\rho^E(1, 1) > 0$ ], we obtain

$$k < g^E(1 - \varepsilon^2)^M - 2c, \quad (\text{B11})$$

which is a nonlinear, monotonically decreasing function of  $\varepsilon$  [solid black curve in Fig. 8(c)].

Finally, we check when the bilateral solution is preferred to the lateralized one [ $\rho^E(1, 1) > \rho^E(1, 0)$ ]:

$$k < g^E(1 - \varepsilon)^M[(1 + \varepsilon)^M - 1] - c, \quad (\text{B12})$$

which is another nonlinear curve on the  $\varepsilon - k$  plane, with a unique maximum [dashed red line in Fig. 8(c)].

The interplay between these conditions yields a map that tells us whether the phenotype is too costly to emerge [large white area in Fig. 2(b)], or whether fully lateralized [gray area in Fig. 2(b)] or bilateral [black area in Fig. 2(b)] neural modules are preferred. We find that the conditions given by Eqs. (B11) and (B12) cross precisely at  $\varepsilon = 1 - \sqrt[M]{c/g^E}$ . Thus, while Eq. (B12) describes the bilateral region for

$\varepsilon < 1 - \sqrt[M]{c/g^E}$ , the rightmost part of the bilateral region is given by Eq. (B11).

### APPENDIX C: MAXIMA FOR EMERGING PHENOTYPES ALONG PREVIOUSLY EXISTING, SIMPLER ONES

In the previous appendix, we studied an emergent phenotype that contributed some fitness only if all building blocks (i.e., all individual tasks) were successfully implemented. However, since such emergent tasks evolve upon a previously existing substrate, it might be the case that each individual neural module contributes some fitness of its own irrespective of the emergent task—e.g., because they still carry out their ancient task during a fraction of the time. To model this, we add up the benefits reported both by the ancient and the emergent phenotypes:

$$G^E + G^S = g^E M(1 - \varepsilon)^M P(l, r)^M + g^S M(1 - \varepsilon) P(l, r), \quad (\text{C1})$$

where  $P(l, r)$  is as before. We assume that the costs still depend on the engagement of each neural unit, so they are as in Eq. (B1). We can write the following utility function:

$$\begin{aligned} \rho^{E+S}(l, r) &\equiv G^E/M + MG^S/M + C/M + K/M \\ &= g^E(1 - \varepsilon)^M P(l, r)^M + g^S(1 - \varepsilon) P(l, r) \\ &\quad - c(l + r) - klr. \end{aligned} \quad (\text{C2})$$

As usual, we need to find maxima of  $\rho^{E+S}(l, r)$  within  $(l, r) \in [0, 1] \times [0, 1]$ . Let us look again at derivatives and singular points along a straight line of constant, fixed  $r$  and as a function of  $l$ :

$$\begin{aligned} \rho_l^{E+S} &= g^E(1 - \varepsilon)^M M P(l, r)^{M-1} [1 - r(1 - \varepsilon)] \\ &\quad + g^S(1 - \varepsilon) [1 - r(1 - \varepsilon)] - c - kr. \end{aligned} \quad (\text{C3})$$

This becomes zero if

$$\rho_l^{E+S} = 0 \Leftrightarrow P(l, r) = \sqrt[M-1]{\gamma - \frac{g^S}{g^E M}}. \quad (\text{C4})$$

We can apply here the same argument as before [illustrated in Figs. 8(a) and 8(b)]. The left-hand side of this equation is a straight line as a function of  $l$ . On the right-hand side, we have, depending on whether  $M - 1$  is odd or even, either one or two constant values as a function of  $l$ . We are not interested in the negative root for even  $M - 1$ . The positive root for even  $M - 1$  or for odd  $M - 1$  might result in the solution of Eq. (C4) for positive  $l$  within  $l \in [0, 1]$ . If this singular point is an extremum of the utility function, it must be a minimum. The right-hand side of Eq. (C4) is the same as the right-hand side of Eq. (B6) with

a small correction,  $g^S/g^E M$ . This correction might turn the argument of the root negative; thus, there would not be any root in the real numbers for even  $M - 1$ . This does not affect our argument, as it just means that extrema do not exist for  $l \in [0, 1]$  in such a case.

All this reasoning implies that the utility function with fixed  $r$ , taking values over  $l \in [0, 1]$ , has its maximum either at  $l = 0$  or  $l = 1$ . Thus, again, within our region of interest, Eq. (C2) can only have its maxima at  $(l, r) = (0, 0)$ ,  $(l, r) = (1, 0)$ ,  $(l, r) = (0, 1)$ , or  $(l, r) = (1, 1)$ , and, again,  $(l, r) = (1, 0)$  and  $(l, r) = (0, 1)$  will be maxima simultaneously. We need to evaluate three points:

$$\begin{aligned} \rho^{E+S}(0, 0) &= 0, \\ \rho^{E+S}(1, 0) &= g^E(1 - \varepsilon)^M + g^S(1 - \varepsilon) - c, \\ \rho^{E+S}(1, 1) &= g^E(1 - \varepsilon^2)^M + g^S(1 - \varepsilon^2) - 2c - k. \end{aligned} \quad (\text{C5})$$

Imposing that the fully lateralized solution is larger than 0, we obtain

$$\rho^{E+S}(1, 0) > 0 \Leftrightarrow c < g^E(1 - \varepsilon)^M + g^S(1 - \varepsilon). \quad (\text{C6})$$

The right-hand side is dominated by the second term when  $M$  is large. In any case, we can ensure that this condition is met if we dismiss the first term and attend to the constraint:

$$\varepsilon < 1 - c/g^S. \quad (\text{C7})$$

Imposing that the bilateral solution has positive utility, we obtain

$$\rho^{E+S}(1, 1) > 0 \Leftrightarrow k < g^E(1 - \varepsilon^2)^M + g^S(1 - \varepsilon^2) - 2c. \quad (\text{C8})$$

This result is depicted as a solid black curve in Fig. 9. Finally, we check when the bilateral solution is preferred to the fully lateralized one,  $\rho^{E+S}(1, 1) > \rho^{E+S}(1, 0)$ :

$$k < g^E(1 - \varepsilon)^M [(1 + \varepsilon)^M - 1] + g^S \varepsilon(1 - \varepsilon) - c. \quad (\text{C9})$$

This result is shown as the dashed red curve in Fig. 9. As before, we can study when these two conditions cross. Imposing that Eq. (C8) equals Eq. (C9), we obtain

$$g^E x^{M-1} + g^S = \frac{c}{x}, \quad (\text{C10})$$

with the change of variable  $\varepsilon \rightarrow 1 - x \Rightarrow x = 1 - \varepsilon$ . Equation (C10) is a monotonically increasing polynomial that crosses a hyperbola only once, and this crossing happens for smaller  $x$  when  $M$  is larger. Thus, as more neural units are involved, both conditions cross for  $\varepsilon \rightarrow 1$ ; thus, Eq. (C8) is the most relevant for us for large  $M$ .

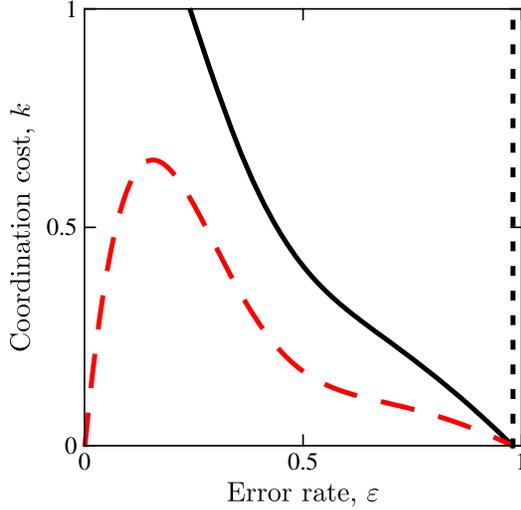


FIG. 9. Resulting conditions for optimal configurations in phenotypes that emerge along earlier, simpler tasks. The vertical dashed line indicates a lower bound for phenotype viability. For  $\varepsilon$  to the left of this line, the emergent phenotype (along with implementation of the earlier, simpler tasks) is viable. To the right of this line, the phenotype might not be viable [but we cannot rule it out since Eq. (C7) is an approximation]. The black solid curve indicates when bilateral configurations pay off. The red dashed curve delimits when bilateral configurations are preferred to lateralized ones. For large  $M$ , these curves do not cross, meaning that the red dashed curve is a more stringent condition for bilaterality. For smaller  $M$ , they might cross.

Again, the interplay between these conditions renders a phase space where we can determine which configuration of neural circuits (lateralized, bilateral, or none) is preferred depending on the model parameters (Fig. 3). All the conditions just derived are basically the same as in the previous section, plus some correction contributed by the nonemergent phenotypes.

#### APPENDIX D: SEGREGATING FUNCTIONS

The emergence of complex phenotypes upon a neural circuitry that continues to implement more ancient tasks opens up the possibility of segregating both computations, lateralizing each cognitive task to a hemisphere. We check when such a solution is favorable as compared to the fully lateralized and mirror-symmetric configurations described in the previous appendix.

If the emergent phenotype contributes some fitness  $\bar{g}^E$  when it is implemented, which happens a fraction  $\tau$  of the time, it results in  $g^E = \tau\bar{g}^E$  for Eq. (C1). Similarly, each ancient, simple task contributes a fitness  $\bar{g}^S$  during the remaining fraction  $1 - \tau$  of the time—resulting in  $g^S = (1 - \tau)\bar{g}^S$  for Eq. (C1). From Eq. (C5), the resulting utility functions for the fully lateralized and bilateral solutions read

$$\begin{aligned}\rho_\tau^{E+S}(1, 0) &= \tau\bar{g}^E(1 - \varepsilon)^M + (1 - \tau)\bar{g}^S(1 - \varepsilon) - c, \\ \rho_\tau^{E+S}(1, 1) &= \tau\bar{g}^E(1 - \varepsilon^2)^M + (1 - \tau)\bar{g}^S(1 - \varepsilon^2) - 2c - k.\end{aligned}\quad (\text{D1})$$

In these configurations, including the fully lateralized one, all functions are performed in the same circuit. Instead, segregating functionality allows two sets of circuits: one for the emergent phenotype and another one for the simpler tasks. We assume, incidentally, that the complex phenotype displaces simpler tasks in the unsegregated solution a fraction  $\tau$  of the time—hence their contribution  $(1 - \tau)$ . Now that phenotypes are segregated, simpler tasks can contribute fully again. With this in mind, we compute the costs and benefits of function segregation. We lateralize the complex phenotype to the left and the simpler tasks to the right, and we denote this as  $\langle L|$  and  $|R\rangle$ , respectively. Then,

$$\rho_\tau^{\langle L|R\rangle} = \tau\bar{g}^E(1 - \varepsilon)^M + \bar{g}^S(1 - \varepsilon) - (1 + \tau)c. \quad (\text{D2})$$

Here, we use  $P(1, 0) = 0 = P(0, 1)$ , and there are no coordination costs across hemispheres if the function is segregated.

First, we find that  $\rho_\tau^{\langle L|R\rangle} > \rho_\tau^{E+S}(1, 0)$  for any  $\varepsilon < 1 - c/\bar{g}^S$ . This includes any configuration in which lateralization is preferred, meaning that whenever lateralization is optimal, it is even more optimal to segregate the function and engage each hemisphere with a different task. This also includes configurations in which  $\rho_\tau^{E+S}(1, 0) < 0$  (i.e., in which a lateralized circuit fails to be cost-efficient enough as to be implemented). In that region, we try to determine if the segregated circuit is viable. We find that  $\rho_\tau^{\langle L|R\rangle} > 0$  for

$$(1 + \tau)c < \tau\bar{g}^E(1 - \varepsilon)^M + \bar{g}^S(1 - \varepsilon). \quad (\text{D3})$$

To roughly approximate this inequality, as above, we ignore the first term on the right-hand side (which becomes small for large  $M$ ), finding  $\varepsilon < 1 - (1 + \tau)c/\bar{g}^S$ .

Finally, we compare  $\rho_\tau^{\langle L|R\rangle}$  and  $\rho_\tau^{E+S}(1, 1)$  to find when the bilateral configuration remains optimal:

$$\begin{aligned}k &< \tau\bar{g}^E(1 - \varepsilon)^M[(1 + \varepsilon)^M - 1] \\ &+ \bar{g}^S(1 - \varepsilon)[\varepsilon - \tau - \tau\varepsilon] - c(1 - \tau).\end{aligned}\quad (\text{D4})$$

The interplay between these curves renders maps of optimality of bilateral versus lateralized and segregated configurations. It is noteworthy that a region of parameters exists [as given by Eq. (D4)] in which the bilateral configuration pays off. This means that the advantage of redundant computations can overcome the added fitness gain from keeping both phenotypes segregated.

## APPENDIX E: GENERALITY OF RESULTS

## 1. Heterogeneous modules—mixing error rates and costs

We have assumed that all modules recruited for an emergent phenotype are similar (i.e., same fallibility  $\varepsilon$ , and costs  $c$  and  $k$ ) and that they gain the same fitness when operating independently (same  $g^S$ ). This is unlikely in nature. If we allow heterogeneity (say, each module presents its own  $\varepsilon_i$ ,  $c_i$ , and  $k_i$  for  $i = 1, \dots, M$ ), we must also assume that modules might lateralize to different degrees independently (i.e., different  $l_i$  and  $r_i$  for each module). The emergent phenotype utility function reads

$$\rho^H(l, r) = g^E \prod_{i=1}^M (1 - \varepsilon_i) [l_i(1 - r_i) + (1 - l_i)r_i + (1 + \varepsilon_i)l_i r_i] - \sum_{i=1}^M [c_i(l_i + r_i) - k_i l_i r_i]. \quad (\text{E1})$$

Let us assume that all the activation levels  $l_j$  and  $r_j$  for  $j \neq i$  are constant. This gives us the utility function of a simple task [Eq. (4)] with the gain multiplied by some constant number and some other constant term (the fixed costs of all modules  $j \neq i$ ) subtracted. This is again a quadratic form with a saddle node; thus, maxima of  $\rho^H$  over  $(l_i, r_i)$  must lie in the outer circuit. This is true for any

module; thus, to maximize Eq. (E1), we need to test all  $(l_i, r_i) = (0, 0)$ ,  $(l_i, r_i) = (1, 0)$ , and  $(l_i, r_i) = (1, 1)$  combinations across modules (graded solutions can be dismissed), which amounts to  $3^M$  possible combinations. In any case, allowing heterogeneity in the model again results in all-or-nothing activation for each module.

Figures 10(a)–10(d) show the fraction of bilateral modules when some heterogeneity is added to the model. Results are stochastic—the  $3^M$  possibilities were tested with a greedy algorithm. However, convergence to each configuration seems robust. For  $M = 10$  and  $M = 100$ , we observe that the bilaterality region is not exhausted, suggesting that heterogeneity adds yet another pressure towards lateralization (likely following the least-efficient module).

## 2. More general dependence on module errors

Another assumption is that fitness gains are all or nothing. Either a computation is implemented perfectly (with probability  $1 - \varepsilon$  in each module) or it is not implemented at all. However, we could reinterpret  $\varepsilon$  as a distance to the target computation and assign some fitness as a function of this distance. Do our results still hold?

Let us assume that each individual unit contributes a fitness that is a generic, nonlinear function  $f_0(\varepsilon)$ . Similarly, if both units are engaged, they contribute another function  $f_{\text{syn}}(\varepsilon)$  that should capture the idea that synergy can result

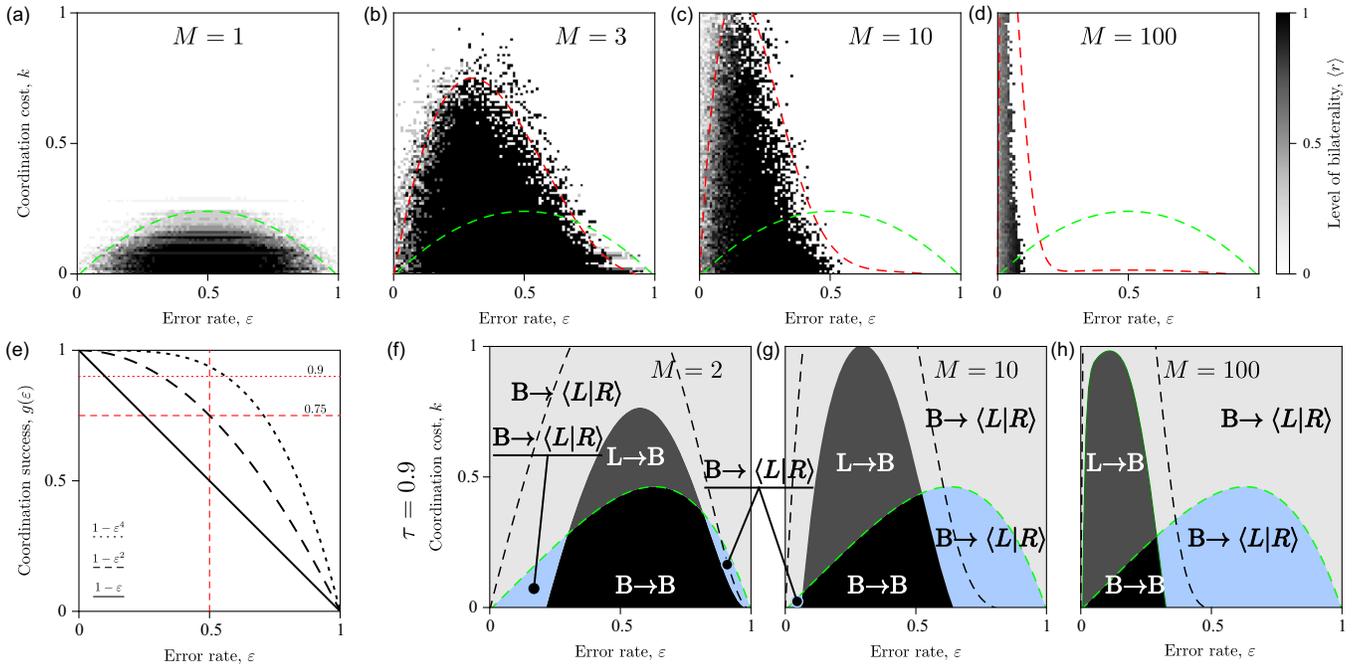


FIG. 10. Optimality in more general cases. (a)–(d) Fraction of bilateral modules in a heterogeneous situation for increasing complexity. For each  $(\varepsilon, k)$ , each module's  $g^S$ ,  $\varepsilon$ , and  $k$  are drawn from a Gaussian centered on  $g^S = 1$ ,  $\varepsilon$ , and  $k$  with standard deviation 0.05. Running costs are kept at  $c = 0.01$ . (e) Illustration of different synergistic functions. Horizontal red lines mark  $f_{\text{syn}}(\varepsilon) = 0.75$  and 0.9. The vertical red line marks units with 0.5 fallibility, which achieve large synergy even with the least synergistic function  $f_{\text{syn}} = 1 - \varepsilon^2$ . (f)–(h) Morphospaces for segregated functions with synergy function  $f_{\text{syn}} = 1 - \varepsilon^4$ .

in higher accuracy. It makes sense that  $f_0(\varepsilon)$  and  $f_{\text{syn}}(\varepsilon)$  are monotonically decreasing with  $\varepsilon$ .

On the other hand, we can impose generic costs  $h_c(\varepsilon)$  and  $h_k(\varepsilon)$ . We can demand that these costs grow with accuracy—thus,  $h_c(\varepsilon)$  and  $h_k(\varepsilon)$  should also be positive, monotonically decreasing with  $\varepsilon$ , and nonlinear, in general.

For simple tasks, the utility function reads

$$\rho^S(l, r) = g^S[f_0(\varepsilon)l(1-r) + f_0(\varepsilon)(1-l)r + g_{\text{syn}}(\varepsilon)lr] - ch_c(\varepsilon)(l+r) - kh_k(\varepsilon)lr, \quad (\text{E2})$$

with derivatives

$$\begin{aligned} \rho_l^S &= g^S[f_0(\varepsilon)(1-2r) + f_{\text{syn}}(\varepsilon)r] - ch_c(\varepsilon) - kh_k(\varepsilon)r, \\ \rho_r^S &= g^S[f_0(\varepsilon)(1-2l) + f_{\text{syn}}(\varepsilon)l] - ch_c(\varepsilon) - kh_k(\varepsilon)l, \end{aligned} \quad (\text{E3})$$

and second derivatives

$$\begin{aligned} \rho_{ll}^S &= 0 = \rho_{rr}^S, \\ \rho_{lr}^S &= g^S[f_{\text{syn}}(\varepsilon) - 2f_0(\varepsilon)] - kh_k(\varepsilon). \end{aligned} \quad (\text{E4})$$

However complicated  $f_0(\varepsilon)$ ,  $f_{\text{syn}}(\varepsilon)$ , or  $h_k(\varepsilon)$  might be, for constant values of  $\varepsilon$ , the utility function is always a quadratic form, and its discriminant is always negative. This implies that  $\rho(l, r)$  is always a saddle, and maxima will lie along the enclosing circuit. We again dismiss graded solutions, which only happen in a negligible region of parameter space (green curves throughout all our morphospaces), and, when they show up, they are indeed as optimal as all-or-nothing configurations.

For emergent phenotypes, we have

$$\rho^E(l, r) = g^E P(l, r)^M + g^S P(l, r) - ch_c(\varepsilon) - kh_k(\varepsilon)lr, \quad (\text{E5})$$

where

$$P(l, r) = f_0(\varepsilon)l(1-r) + f_0(\varepsilon)(1-l)r + f_{\text{syn}}(\varepsilon)lr. \quad (\text{E6})$$

As before, this polynomial is a straight line of  $l$  (respectively,  $r$ ) if we take a fixed, constant  $r$  (respectively,  $l$ ). Let us look at the utility function along a line of fixed, constant  $r$  and as a function of  $l$ ; we take

$$\begin{aligned} \rho_l^E &= g^E M P^{M-1}(l, r)[(1-2r)f_0(\varepsilon) + rf_{\text{syn}}(\varepsilon)] \\ &\quad + g^S[(1-2r)f_0(\varepsilon) + rf_{\text{syn}}(\varepsilon)] \\ &\quad - ch_c(\varepsilon) - kh_k(\varepsilon)r. \end{aligned} \quad (\text{E7})$$

Thus, we obtain

$$\begin{aligned} \rho_l^E = 0 &\Leftrightarrow P(l, r) = {}^{M-1}\sqrt{\gamma}, \\ \gamma &\equiv \frac{ch_c(\varepsilon) + kh_k(\varepsilon)r - g^S[(1-2r)f_0(\varepsilon) + rf_{\text{syn}}(\varepsilon)]}{g^E M[(1-2r)f_0(\varepsilon) + rf_{\text{syn}}(\varepsilon)]}. \end{aligned} \quad (\text{E8})$$

Again, we have solutions to this equation when a straight line as a function of  $l$  crosses either one or two constant values [as illustrated in Fig. 8(a)]. Introducing the actual form of  $P(l, r)$ , we obtain

$$l = \frac{{}^{M-1}\sqrt{\gamma} - rf_0(\varepsilon)}{(1-2r)f_0(\varepsilon) + rf_{\text{syn}}(\varepsilon)}. \quad (\text{E9})$$

In the numerator, we have either the positive or negative  $(M-1)$ th root of  $\gamma$  minus  $rf_0(\varepsilon)$ , a positive term. If  $M-1$  is odd, we obtain either one negative or one positive value in the numerator. If  $M-1$  is even and  $rf_0(\varepsilon) < \|{}^{M-1}\sqrt{\gamma}\|$ , we obtain one positive and one negative value. If  $M-1$  is even and  $rf_0(\varepsilon) > \|{}^{M-1}\sqrt{\gamma}\|$ , we obtain two negative values in the numerator. If this is ever the case, and the denominator is negative, the two negative values in the numerator could turn into two singular points within  $l \in [0, 1]$ , and one of them might be a maximum. In any other situation, again, maxima of  $\rho^E(l, r)$  will lie either at  $(l, r) = (0, 0)$ ,  $(l, r) = (1, 0)$ ,  $(l, r) = (0, 1)$ , or  $(l, r) = (1, 1)$ .

We obtain a negative number in the denominator when

$$(1-2r)f_0(\varepsilon) + rf_{\text{syn}}(\varepsilon) < 0 \Leftrightarrow \frac{f_{\text{syn}}(\varepsilon)}{f_0(\varepsilon)} > 2 - \frac{1}{r}. \quad (\text{E10})$$

In the simple model studied above, we have chosen  $f_0(\varepsilon) = 1 - \varepsilon$  and  $f_{\text{syn}}(\varepsilon) = 1 - \varepsilon^2$ . We obtain

$$\frac{f_{\text{syn}}(\varepsilon)}{f_0(\varepsilon)} = \frac{(1-\varepsilon)(1+\varepsilon)}{1-\varepsilon} = 1 + \varepsilon. \quad (\text{E11})$$

This number is always larger than 1 and, within our range of interest ( $r \in [0, 1]$ ),  $2 - 1/r$  is always smaller or equal than 1. In other words, our chosen functions fulfill the condition without graded activities of the circuits, as we already know.

Figure 10(e) compares  $f_0(\varepsilon) = 1 - \varepsilon$  (straight line) to the computational efficiency of synergy functions  $f_{\text{syn}}(\varepsilon) = 1 - \varepsilon^2$  (also used above) and  $f_{\text{syn}}(\varepsilon) = 1 - \varepsilon^4$ . Figures 10(f)–10(h) show optimality constraints for this much more synergistic version of the model. Again, we observe how increased complexity leads to optimality pressures towards lateralization. This morphospace is for segregated functions. Comparing Figures 10(f)–10(h) to Fig. 4, we see that the interesting  $\mathbf{L} \rightarrow \mathbf{B}$  pathway becomes much more abundant again with larger within-module synergy.

### 3. Nonlinear dependencies on unit engagement

Alternatively, we could also conceive of modules whose contribution to fitness grows nonlinearly with the time that they are engaged. A circuit can become more effective the longer it is working on a task—e.g., because it learns or simply because more computational resources have a nonlinear effect on performance. Assume more general functions:  $f_0(\varepsilon, l)$  or  $f_0(\varepsilon, r)$  (depending on whether  $L$  or  $R$  is independently engaged) and  $f_{\text{syn}}(\varepsilon, l, r)$ . Assume that they increase monotonically with  $l$  and  $r$ . A good Ansatz could be  $f_0 = f_0(\varepsilon)\sigma(l)$  or  $f_0 = f_0(\varepsilon)\sigma(r)$  and  $f_{\text{syn}} = f_{\text{syn}}(\varepsilon)\sigma(l)\sigma(r)$ , where  $\sigma(\cdot)$  are monotonically increasing sigmoids, or perhaps  $\sigma(l) = l^\beta$ . This does not matter as long as  $f_0(\varepsilon, l)$ ,  $f_0(\varepsilon, r)$ , and  $f_{\text{syn}}(\varepsilon, l, r)$  increase monotonically with  $l$  and  $r$ . Then, these functions are bound by maxima,  $f_0^+(\varepsilon) \geq f_0(\varepsilon, 1)$  and  $f_{\text{syn}}^+(\varepsilon) \geq f_{\text{syn}}(\varepsilon, 1, 1)$ , which do not depend on  $l$  and  $r$ . Solving for these bounds, we get the same results as before: all-or-nothing engagement, morphospaces, etc.

For  $(l, r) = (0, 0)$ ,  $(1, 0)$ , or  $(1, 1)$ , solving for  $f_0(\varepsilon, l)$  or  $f_0(\varepsilon, r)$  and  $f_{\text{syn}}(\varepsilon, l, r)$  is the same as solving for  $f_0^+(\varepsilon)$  and  $f_{\text{syn}}^+(\varepsilon)$ . For any other values  $(l, r) \in (0, 1) \times (0, 1)$ , the fitness gains are lower if we use  $f_0(\varepsilon, l)$  or  $f_0(\varepsilon, r)$  and  $f_{\text{syn}}(\varepsilon, l, r)$  than if we use  $f_0^+(\varepsilon)$  and  $f_{\text{syn}}^+(\varepsilon)$ , while costs remain the same. Thus, the maxima for  $f_0^+(\varepsilon)$  and  $f_{\text{syn}}^+(\varepsilon)$  are maxima also for  $f_0(\varepsilon, l)$  or  $f_0(\varepsilon, r)$  and  $f_{\text{syn}}(\varepsilon, l, r)$ . In other words, optimal solutions for these kinds of nonlinearities are the same as for the scenarios studied throughout the paper, substituting the corresponding bounds  $f_0^+(\varepsilon)$  and  $f_{\text{syn}}^+(\varepsilon)$ .

### 4. Synergies across modules

We assume that all modules operate independently. A compelling alternative is that one module might increase the performance of another (e.g., because it provides a substrate to delegate operations or because it preprocesses inputs and clears noise), even if no new phenotype emerges. Indeed, this could depict early stages in which preadaptations start working together before true novelty arises.

To study how different modules enhance each other, we would need to define a network providing their synergies. A simple example shows how this situation does not alter our results. Assume two modules  $m_1$  and  $m_2$ , each devoted to a simple task. Let the performance of  $m_1$  increase whenever  $m_2$  operates successfully. The fitness gain by module  $m_1$  now reads

$$G_1^S = [g^S + g^E P_2(l, r; \varepsilon)] \cdot P_1(l, r; \varepsilon), \quad (\text{E12})$$

where  $P_i \equiv (1 - \varepsilon)[l(1 - r) + (1 - l)r + (1 + \varepsilon)lr]$  is the likelihood that module  $m_i$  performs correctly (note that we could assume heterogeneous parameters or other, nonlinear dependencies, as above). The parameter  $g^E$  here measures

the contribution of module  $m_2$  to enhancing  $m_1$ . We use the same superscript as for emergent phenotypes to denote the similarity between all cases.

Equation (E12) is the same as Eq. (A1), with a correction from the enhancement that  $m_2$  contributes to  $m_1$ . Costs of running  $m_2$  are accounted for by the corresponding term. When plugging Eq. (E12) into a utility function, we obtain a situation mathematically similar to that in Sec. II C 1 [see Eq. (C1)]. This means that synergies across modules can be seamlessly incorporated into our model—indeed, the increase in diversity of modules (i.e., a larger number of different computing parts available) can also result in optimality pressures towards lateralization *before* true phenotypic novelty arises. We can add the effects of emergent phenotypes by including additional terms. The difference lies in the values of  $g^E$ , which should be larger in truly new, emergent phenotypes. Considering an interpretation in terms of preadaptations coming together, we open up the model to capturing parsimonious evolutionary pathways.

## APPENDIX F: ABOUT THE CONNECTION BETWEEN MATHEMATICS AND ACTUAL BIOLOGY

Our modules, tasks, and phenotypes are mathematical contraptions used to perform calculations, conferring the great generality of our results. Our optimality constraints must apply whenever real-life components fit our framework—whether they are neural circuits, microchip components, etc. Note that our modules contain two, exact copies of the same circuit—whether they are two similar neurons with the same input or output targets, or two overgrown cortical columns conforming mirror-symmetric barrels in the mice cortex. They might also consist of other modules nested in a hierarchy. If so, fallibility  $\varepsilon$  and costs  $c$  and  $k$  would become compounded, mapping recursively into our model—as in a renormalization scheme. Thus, it is not important to us at which point of a hierarchy we start. Working renormalization-like flows in our model might be an interesting follow-up, but here we assume that each module is irreducible, performing a task much simpler than that of the emerging phenotype. Do real brains resemble our modules, tasks, and emerging phenotypes in any way? If so, the derived optimality constraints must apply. The answer is a resounding yes, with examples abounding across scales. The set of regions involved in human language offers a case in which we know some of the subtasks precisely. Figure 11(a) illustrates this hierarchical, modular, extended language network, which has recently been worked out by Fedorenko and colleagues [10–13, 15].

A final mathematical aspect worth discussing is how the studied phenomenology results from the interplay between symmetries in the brain versus those of a cognitive task [Fig. 11(b)]. Brains might have evolved for motor control. In Bilateria, this results in a perfect correspondence between mirror symmetry in the brain and in the parts

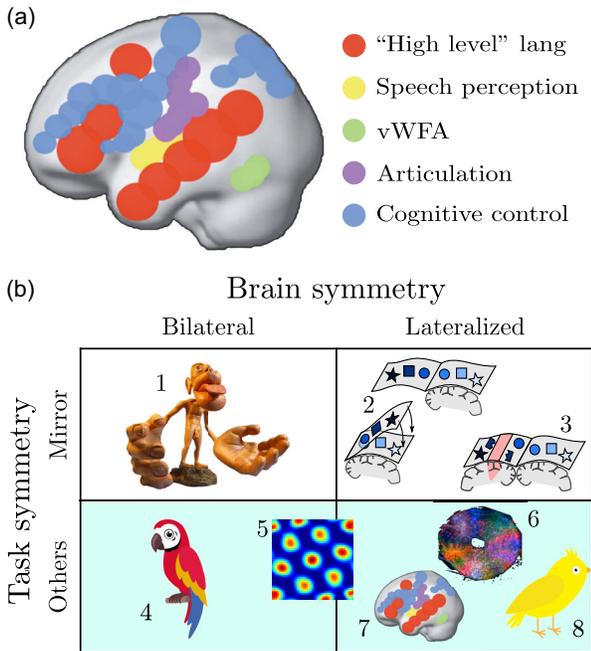


FIG. 11. Considerations about the model mathematics and biology. (a) Example of modules implementing subtasks within an emergent phenotype—human language. Some modules (e.g., “high level” language) are further divided into Broca, Wernicke, and others, each solving specific, indispensable subtasks. Plot adapted from [12]. (b) Symmetry constraints of our brains and of cognitive tasks. Brain bilaterality can be present (left) or broken (right). Cognitive tasks might be fundamentally mirror symmetry (top; e.g., body control), or they may impose other symmetries. 1: Control of a mirror-symmetric body with a bilateral brain. 2: Control of a mirror-symmetric body with one hemisphere (adapted from Ref. [40]). 3: Mirror-symmetric topographic map affected by a stroke (adapted from Ref. [68]). 4: Parrots with bilateral control circuits for speech production [69]. 5: Grid cells representing space in the entorhinal cortex. They are found bilaterally in mice [70] but lateralized in humans [71]. 6: Head-direction representation in the fly demanding a toroid, which the fly fits within its mirror-symmetric brain [72]. 7: Asymmetric network for human language. 8: Song production in songbirds, which is fully lateralized [69].

that need to be motor controlled [Fig. 11(b), item 1]. This paper does not investigate that specific situation. Rather, we study the possible fate of mirror symmetry when it is still available in the brain, but a cognitive task does not demand it [Fig. 11(b), lower row]. Producing symbolic sequences does not require mirror symmetry, as the cases of human language (item 7) and singing birds (item 8) show, but parrots retain it (item 4). The same happens to space representation (items 5 and 6), which engages mirror-symmetric circuits in mice but not humans, and which occupies a toroid topology in flies. Elsewhere, we have studied the opposite situation in which a brain has lost its mirror symmetry, but a task still demands it [40,68]. At the moment, we do not know, empirically, how much each of

these scenarios might owe to the optimality pressures explored in this work. Rather, examples in Fig. 11(b) intend to illustrate a rich niche for mathematical neuroscience in the study of how symmetries inherited from our bodies’ biology and evolutionary history interact with other abstract constraints (perhaps other symmetries as well) imposed by cognitive tasks.

[1] A. M. Galaburda, M. LeMay, T. L. Kemper, and N. Geschwind, *Right-Left Asymmetries in the Brain*, *Science* **199**, 852 (1978).

[2] A. W. Toga and P. M. Thompson, *Mapping Brain Asymmetry*, *Nat. Rev. Neurosci.* **4**, 37 (2003).

[3] K. Hugdahl, *Symmetry and Asymmetry in the Human Brain*, *Eur. Rev.* **13**, 119 (2005).

[4] P. Y. Hervé, L. Zago, L. Petit, B. Mazoyer, and N. Tzourio-Mazoyer, *Revisiting Human Hemispheric Specialization with Neuroimaging*, *Trends Cognit. Sci.* **17**, 69 (2013).

[5] X. Z. Kong, S. R. Mathias, T. Guadalupe, D. C. Glahn, B. Franke, F. Crivello, N. Tzourio-Mazoyer, S. E. Fisher, P. M. Thompson, and C. Francks (ENIGMA Laterality Working Group), *Mapping Cortical Brain Asymmetry in 17,141 Healthy Individuals Worldwide via the ENIGMA Consortium*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5154 (2018).

[6] X. Z. Kong, M. C. Postema, T. Guadalupe, C. de Kovel, P. S. Boedhoe, M. Hoogman, S. R. Mathias, D. Van Rooij, D. Schijven, D. C. Glahn, and S. E. Medland, *Mapping Brain Asymmetry in Health and Disease through the ENIGMA Consortium*, *Hum. Brain Mapp.* **43**, 167 (2020).

[7] L. F. Seoane, *Fate of Duplicated Neural Structures*, *Entropy* **22**, 928 (2020).

[8] N. Geschwind, *Language and the Brain*, *Sci. Am.* **226**, 76 (1972).

[9] M. Catani, D. K. Jones, and D. H. Ffytche, *Perisylvian Language Networks of the Human Brain*, *Ann. Neurol.* **57**, 8 (2005).

[10] E. Fedorenko and N. Kanwisher, *Neuroimaging of Language: Why Hasn’t a Clearer Picture Emerged?*, *Lang. Linguist.* **3**, 839 (2009).

[11] E. Fedorenko, A. Nieto-Castanon, and N. Kanwisher, *Lexical and Syntactic Representations in the Brain: An fMRI Investigation with Multi-Voxel Pattern Analyses*, *Neuropsychologia* **50**, 499 (2012).

[12] E. Fedorenko and S. L. Thompson-Schill, *Reworking the Language Network*, *Trends Cognit. Sci.* **18**, 120 (2014).

[13] I. Blank, Z. Balewski, K. Mahowald, and E. Fedorenko, *Syntactic Processing Is Distributed Across the Language System*, *NeuroImage* **127**, 307 (2016).

[14] R. C. Berwick and N. Chomsky, *Why Only Us: Language and Evolution* (MIT Press, Cambridge, USA, 2016).

[15] S. Malik-Moraleda, D. Ayyash, J. Gallée, J. Affourtit, M. Hoffmann, Z. Minerof, O. Jouravlev, and E. Fedorenko, *An Investigation Across 45 Languages and 12 Language Families Reveals a Universal Language Network*, *Nat. Neurosci.* **25**, 1014 (2022).

[16] H. D. Brown and S. M. Kosslyn, *Hemispheric Differences in Visual Object Processing: Structural versus Allocation Theories*, in *Brain Asymmetry*, edited by R. J. Davidson

- and K. Hugdahl (MIT Press, Cambridge, MA, 1995), pp. 77–97.
- [17] J. B. Hellige, *Hemispheric Asymmetry for Components of Visual Information Processing*, in *Brain Asymmetry*, edited by R. J. Davidson and K. Hugdahl (MIT Press, Cambridge, MA, 1995), pp. 99–121.
- [18] A. M. Galaburda, *Anatomic Basis of Cerebral Dominance*, in *Brain Asymmetry*, edited by R. J. Davidson and K. Hugdahl (MIT Press, Cambridge, MA, 1995), pp. 51–73.
- [19] L. J. Rogers, P. Zucca, and G. Vallortigara, *Advantages of Having a Lateralized Brain*, *Proc. R. Soc. B* **271**, S420 (2004).
- [20] M. E. Halpern, O. Güntürkün, W. D. Hopkins, and L. J. Rogers, *Lateralization of the Vertebrate Brain: Taking the Side of Model Systems*, *J. Neurosci.* **25**, 10351 (2005).
- [21] L. J. Rogers, *Cognitive and Social Advantages of a Lateralized Brain*, in *Behavioral and Morphological Asymmetries in Vertebrates*, edited by Y. B. Malashichev and A. Wallace Deckel (Landes Bioscience, Georgetown, 2006), pp. 129–139, <https://tinyurl.com/mvpdev85>.
- [22] A. Harrington, *Unfinished Business: Models of Laterality in the Nineteenth Century*, in *Brain Asymmetry*, edited by R. J. Davidson and K. Hugdahl (MIT Press, Cambridge, MA, 1995), pp. 3–27.
- [23] M. C. Corballis, *The Evolution and Genetics of Cerebral Asymmetry*, *Phil. Trans. R. Soc. B* **364**, 867 (2009).
- [24] P. F. MacNeilage, L. J. Rogers, and G. Vallortigara, *Evolutionary Origins of Your Right and Left Brain*, *Sci. Am.* **301**, 60 (2009).
- [25] A. Pascual, K. L. Huang, J. Neveu, and T. Pr eat, *Brain Asymmetry and Long-Term Memory*, *Nature (London)* **427**, 605 (2004).
- [26] G. Vallortigara, L. J. Rogers, and A. Bisazza, *Possible Evolutionary Origins of Cognitive Brain Lateralization*, *Brain Res. Rev.* **30**, 164 (1999).
- [27] J. L. Ringo, R. W. Doty, S. Demeter, and P. Y. Simard, *Time Is of the Essence: A Conjecture that Hemispheric Specialization Arises from Interhemispheric Conduction Delay*, *Cereb. Cortex* **4**, 331 (1994).
- [28] M. C. Corballis, *The Evolution of Lateralized Brain Circuits*, *Front. Psychol.* **8**, 1021 (2017).
- [29] A. Kolchinsky and D. H. Wolpert, *Thermodynamic Costs of Turing Machines*, *Phys. Rev. Res.* **2**, 033312 (2020).
- [30] D. H. Wolpert and A. Kolchinsky, *Thermodynamics of Computing with Circuits*, *New J. Phys.* **22**, 063047 (2020).
- [31] A. Kolchinsky and D. H. Wolpert, *Work, Entropy Production, and Thermodynamics of Information under Protocol Constraints*, *Phys. Rev. X* **11**, 041024 (2021).
- [32] J. von Neumann, *Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components*, *Automata Studies* **34**, 43 (1956).
- [33] E. F. Moore and C. E. Shannon, *Reliable Circuits Using Less Reliable Relays*, *J. Franklin Inst.* **262**, 191 (1956).
- [34] S. Winograd and J. D. Cowan, *Reliable Computation in the Presence of Noise* (MIT Press, Cambridge, MA, 1963).
- [35] J. Liang, S. J. Wang, and C. Zhou, *Less Is More: Wiring-Economical Modular Networks Support Self-Sustained Firing-Economical Neural Avalanches for Efficient Processing*, *Natl. Sci. Rev.* **9**, nwab102 (2022).
- [36] G. M. Edelman, *Neural Darwinism: The Theory of Neuronal Group Selection* (Basic Books, New York, 1987).
- [37] A. Oll -Vila, L. F. Seoane, and R. Sol , *Ageing, Computation and the Evolution of Neural Regeneration Processes*, *J. R. Soc. Interface* **17**, 0181 (2020).
- [38] M. Chakraborty and E. D. Jarvis, *Brain Evolution by Brain Pathway Duplication*, *Phil. Trans. R. Soc. B* **370**, 0056 (2015).
- [39] O. A. Olulade, A. Seydell-Greenwald, C. E. Chambers, P. E. Turkeltaub, A. W. Dromerick, M. M. Berl, W. D. Gaillard, and E. L. Newport, *The Neural Basis of Language Development: Changes in Lateralization over Age*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 23477 (2020).
- [40] L. F. Seoane and R. Sol , *Modeling Brain Reorganization after Hemispherectomy*, [10.1101/2020.12.25.424412](https://doi.org/10.1101/2020.12.25.424412) (2020).
- [41] S. Koelsch, *Neural Substrates of Processing Syntax and Semantics in Music*, *Curr. Opin. Neurobiol.* **15**, 207 (2009).
- [42] D. M. Raup, *Geometric Analysis of Shell Coiling: General Problems*, *J. Paleontol.* **40**, 1178 (1966).
- [43] J. Tyszka, *Morphospace of Foraminiferal Shells: Results from the Moving Reference Model*, *Lethaia* **39**, 1 (2006).
- [44] K. J. Niklas, *Computer Models of Early Land Plant Evolution*, *Annu. Rev. Earth Planet Sci.* **32**, 47 (2004).
- [45] B. Corominas-Murtra, J. Go ni, R. V. Sol , and C. Rodr guez-Caso, *On the Origins of Hierarchy in Complex Networks*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13316 (2013).
- [46] J. Go ni, A. Avena-Koenigsberger, N. Velez de Mendizabal, M. P. van den Heuvel, R. F. Betzel, and O. Sporns, *Exploring the Morphospace of Communication Efficiency in Complex Networks*, *PLoS One* **8**, e58070 (2013).
- [47] A. Avena-Koenigsberger, J. Go ni, R. Sol , and O. Sporns, *Network Morphospace*, *J. R. Soc. Interface* **12**, 0881 (2014).
- [48] X. D. Arsiwalla, R. Sol , C. Moulin-Frier, I. Herreros, M. Sanchez-Fibla, and P. Verschure, *The Morphospace of Consciousness*, [arXiv:1705.11190](https://arxiv.org/abs/1705.11190).
- [49] L. F. Seoane and R. Sol , *The Morphospace of Language Networks*, *Sci. Rep.* **8**, 1 (2018).
- [50] R. Sol  and L. F. Seoane, *Evolution of Brains and Computers: The Roads Not Taken*, *Entropy* **24**, 665 (2022).
- [51] L. F. Seoane, *Evolutionary Aspects of Reservoir Computing*, *Phil. Trans. R. Soc. B* **374**, 0377 (2019).
- [52] D. Duong-Tran, K. Abbas, E. Amico, B. Corominas-Murtra, M. Dziedzic, D. Kareken, M. Ventresca, and J. Go ni, *A Morphospace of Functional Configuration to Assess Configurational Breadth Based on Brain Functional Networks*, *Netw. Neurosci.* **5**, 666 (2021).
- [53] D. Zhou, C. Lebel, A. Evans, and C. Beaulieu, *Cortical Thickness Asymmetry from Childhood to Older Adulthood*, *NeuroImage* **83**, 66 (2013).
- [54] K. J. Plessen, K. Hugdahl, R. Bansal, X. Hao, and B. S. Peterson, *Sex, Age, and Cognitive Correlates of Asymmetries in Thickness of the Cortical Mantle Across the Life Span*, *J. Neurosci.* **34**, 6294 (2014).
- [55] X. Kang, T. J. Herron, M. Ettlinger, and D. L. Woods, *Hemispheric Asymmetries in Cortical and Subcortical Anatomy*, *Laterality* **20**, 658 (2015).

- [56] S. F. Witelson, *The Brain Connection: The Corpus Callosum Is Larger in Left-Handers*, *Science* **229**, 665 (1985).
- [57] H. Liepmann, *Die linke Hemisphäre und das Handeln. I*, in *Drei Aufsätze aus dem Apraxiegebiet* (Karger Publishers, Berlin, 1905), pp. 17–50.
- [58] D. Kliemann, R. Adolphs, L. K. Paul, J. M. Tyszka, and D. Tranel, *Reorganization of the Social Brain in Individuals with Only One Intact Cerebral Hemisphere*, *Brain Sci.* **11**, 965 (2021).
- [59] G. Schlaug, L. Jäncke, Y. Huang, J. F. Staiger, and H. Steinmetz, *Increased Corpus Callosum Size in Musicians*, *Neuropsychologia* **33**, 1047 (1995).
- [60] H. Steinmetz, *Structure, Function and Cerebral Asymmetry: in vivo Morphometry of the Planum Temporale*, *Neurosci. Biobehav. Rev.* **20**, 587 (1996).
- [61] J. P. Keenan, V. Thangaraj, A. R. Halpern, and G. Schlaug, *Absolute Pitch and Planum Temporale*, *NeuroImage* **14**, 1402 (2001).
- [62] H. Markram, *The Blue Brain Project*, *Nat. Rev. Neurosci.* **7**, 153 (2006).
- [63] I. Hurley, M. E. Hale, and V. E. Prince, *Duplication Events and the Evolution of Segmental Identity*, *Evol. Dev.* **7**, 556 (2005).
- [64] T. H. Oakley and A. S. Rivera, *Genomics and the Evolutionary Origins of Nervous System Complexity*, *Curr. Opin. Genet. Dev.* **18**, 479 (2008).
- [65] E. Schneidman, W. Bialek, and M. J. Berry, *Synergy, Redundancy, and Independence in Population Codes*, *J. Neurosci.* **23**, 11539 (2003).
- [66] X. Pitkow and D. E. Angelaki, *Inference in the Brain: Statistics Flowing in Redundant Population Codes*, *Neuron* **94**, 943 (2017).
- [67] J. L. Puchalla, E. Schneidman, R. A. Harris, and M. J. Berry, *Redundancy in the Population Code of the Retina*, *Neuron* **46**, 493 (2005).
- [68] A. Carballo-Castro and L. F. Seoane, *Phase Transitions in a Simple Model of Focal Stroke Imitate Recovery and Suggest Neurorehabilitation Strategies* (2022), [10.1101/2022.12.14.520421](https://doi.org/10.1101/2022.12.14.520421).
- [69] M. Hiscock and M. Kinsbourne, *Phylogeny and Ontogeny of Cerebral Lateralization*, in *Brain Asymmetry*, edited by R. J. Davidson and K. Hugdahl (MIT Press, Cambridge, MA, 1995), pp. 535–578.
- [70] T. Hafting, M. Fyhn, S. Molden, M. B. Moser, and E. I. Moser, *Microstructure of a Spatial Map in the Entorhinal Cortex*, *Nature (London)* **436**, 801 (2005).
- [71] C. F. Doeller, C. Barry, and N. Burgess, *Evidence for Grid Cells in a Human Memory Network*, *Nature (London)* **463**, 657 (2010).
- [72] S. S. Kim, H. Rouault, S. Druckmann, and V. Jayaraman, *Ring Attractor Dynamics in the Drosophila Central Brain*, *Science* **356**, 849 (2017).