On Information in Evolutionary Processes

Enrico M. Bucci, PhD[1]

Abstract

Since the first attempts to introduce an information-theoretical formalism into the description of evolutionary processes, several authors have argued that such approaches are inappropriate because biological evolution does not unfold in a predefined space of possibilities.

To properly address that objection, we need to separate the semantics of the emergence of biological functions from the statistical structure of the selection processes.

Here, we propose a unified framework that treats Darwinian evolution as a population-level process and selection as a change in the frequency distributions of equivalence classes of the population components, independently of individual organisms, fitness assignments, or mechanistic implementations. The selection process is thus a population distribution restructuring, which can be quantitatively examined using divergence measures between pre- and post-selection distributions, independent of the selection process's adaptive nature.

In this framework, if the selection process is adaptive, then the statistical coupling between a set of relevant environmental variables/states and the population distribution is directly quantified by the mutual information connecting these two distributions, providing a non-semantic, operational measure of adaptation at the species level, conditional on a pre-specified infotype support and selective-domain definition. After selection, variety may be reintroduced into the population under scrutiny by several means: this is the process that opens and reshapes the space of selectable configurations, enabling the population to evolve iteratively through new selection and variation stages.

At the level of single organisms, fitness is formalized, within selective regimes in which persistence weights can be assigned to infotypes without explicit dependence on current population composition or density, as the function that projects the frequency of the corresponding selection equivalence class from the initial distribution to the final one during the selection process. Within this scope, fitness

---

[1] enrico.bucci@temple.edu - Sbarro Institute - College of Science and Technology, Department of Biology – Temple University - 1900 N 12th St - 19122 Philadelphia, PA - US

can be decomposed into components linked to individual information extraction from the environment, and non-informational contributions that modulate the persistence of the particular equivalence class considered. Regimes in which persistence depends explicitly on population composition or density require a state-dependent extension of the present framework and fall outside the scope of the present treatment.

We thus show how, using the Shannonian formalism, a unified, structurally constrained description of selection and evolution remains applicable across biological and non-biological domains, while explicitly delineating its scope and limits.

1. Introduction

Recent discussions on the role of information in evolutionary theory have raised a substantive objection to the use of informational formalisms as a general framework for selection. It has been argued that biological evolution does not unfold within a predefined space of possibilities, and that biological functions, affordances, and ecological roles emerge historically in ways that cannot be enumerated or classified ex ante. On this basis, informational measures defined over probability distributions on fixed domains—such as Shannon entropy or mutual information—are often taken to lack general scope, applying only to local or retrospective analyses rather than to selection as a general process[1], [2]. This objection does not concern the empirical adequacy of evolutionary theory, nor the explanatory success of specific adaptive accounts. It targets instead the legitimacy of treating selection itself as an informational phenomenon.

A source of ambiguity in this debate lies in the way evolutionary novelty is framed—questions about why a particular novelty emerges often conflate causal mechanisms with teleological interpretations. Within a scientific description of evolution, novelty is treated as a contingent outcome of physical processes, not as the realisation of a pre-existing functional possibility. The relevant explanatory task is therefore not to account for the direction of novelty, but to describe how, once present, it restructures the selective landscape. Clarifying this point is essential for assessing what an informational description of selection can and cannot claim to explain, and for identifying the appropriate level at which such a description operates.

A substantial literature has already employed information-theoretic quantities in evolutionary biology and related fields. Entropy, mutual information, and information gain have been used to characterise genetic adaptation, learning dynamics, neural plasticity, immune repertoire formation, and population responses to fluctuating environments [3]–[11]. These approaches demonstrate that Shannon-type measures can capture relevant statistical aspects of adaptive processes and provide quantitative insight into how systems respond to structured external conditions. Their scope, however, remains domain-specific mainly. They typically rely on modelling choices that fix the relevant state spaces, strategies, or environmental variables in advance, and they do not by themselves establish a general criterion for determining whether heterogeneous selective processes instantiate the same form of selection.

The central issue, therefore, is not the availability of informational tools, but the absence of an explicit and invariant object of selection that would make selective processes comparable in a principled way. In the absence of such an object, claims of continuity between different forms of selection remain largely analogical. Processes implemented through different physical mechanisms—such as differential reproduction, clonal expansion, synaptic stabilisation, or preferential retention—may all be described as selective. Yet, no formal criterion specifies whether they operate on the same quantity or merely share a superficial resemblance.

The framework developed in this work addresses this gap by making explicit both the object of selection and the structural constraints that any admissible selective process must satisfy. Its aim is not to introduce new empirical entities, evolutionary mechanisms, or biological principles. Instead, it provides a constrained formal specification of what selection operates on once a criterion of persistence is defined. In this sense, the contribution is methodological: it delineates the conditions under which heterogeneous processes can be treated as instances of the same selective phenomenon, independently of their physical realisation.

The present treatment concerns selective regimes in which persistence weights can be assigned to infotypes within a given selective episode without explicit dependence on current population composition or density. Frequency-dependent and density-dependent cases define an important class of state-dependent extensions of the framework and are left outside the scope of the present treatment.

A key step in this specification consists of separating the adaptive content of selection from other contributions to persistence. Selection is empirically identified through reproducible changes in population composition. These changes may arise for multiple reasons, including energetic costs, architectural constraints, ecological limitations, or demographic effects. Among these contributions, adaptive selection is identified narrowly as the ordering of selectable variants induced by their statistical coupling to a relevant selective domain. This ordering is quantified at the population level using mutual information. The selectable units entering this description are fixed beforehand by the adopted descriptive resolution and persistence criterion, so mutual information measures adaptive coupling within a specified regime rather than defining the units on which selection operates. In empirical applications, the descriptive alphabet and the resulting infotype partition must be specified independently of the observed outcomes of selection and independently of the selective-domain variables used to estimate mutual information, to avoid circular inference. Mutual information measures the extent to which persistent variants exploit correlations with the selective domain in an operational, non-semantic sense. Other contributions modulate persistence but do not define the adaptive object of selection itself.

This distinction prevents the systematic conflation of adaptation with persistence as such. Selection may favour or disfavour variants for many reasons, but only those differences that track exploitable correlations with the selective domain constitute adaptation in the informational sense adopted here. Making this separation explicit is essential both to avoid an indiscriminate extension of the language of information and to preserve its explanatory precision.

A further source of confusion in the existing literature concerns the level at which selection and adaptation are described. Fitness is often introduced as an individual property, and informational quantities are frequently interpreted as attributes of organisms or strategies. By contrast, the present framework treats selection and adaptation as population-level phenomena. Individuals and fitness are introduced only at a later stage, as derived constructs that account for how population restructuring is realised through organism-level processes. This separation of levels is required to prevent circularity and category errors in the interpretation of informational measures.

Concerns about the non-prestatability of biological functions bear on the semantic interpretation of evolutionary novelty, not on the statistical characterisation of selection once a selective criterion is operative. The emergence of new functions limits what can be predicted in advance, but it does not constrain what can be measured at the level of population dynamics. By explicitly separating semantic interpretation from statistical description, the present framework delineates the legitimate scope of informational analysis without assuming a predefined global space of functions.

The analysis that follows develops this framework in a stepwise manner. It begins by characterising Darwinian evolution as a population-level process of probabilistic restructuring, introduces divergence measures to quantify selection, and identifies mutual information as the population-level signature of adaptive coupling. Only after this structure is in place does it address organism-level fitness and individual information extraction, and finally clarifies how these levels are related without being conflated.

2. State of the art and conceptual positioning

The increasing use of information-theoretic concepts in evolutionary biology reflects a genuine intuition: selection, adaptation, and evolutionary change manifest themselves as structured regularities in populations, which appear naturally amenable to statistical description. Over the past decades, a diverse literature has explored this intuition from multiple angles, introducing entropy, mutual information, divergence measures, and related quantities into the analysis of biological evolution. At the same time, this literature remains fragmented, both conceptually and formally, and it has not converged on a shared understanding of what, exactly, information is supposed to measure in evolutionary processes.

A first line of critique, articulated most explicitly in discussions of open-ended evolution, targets the very legitimacy of informational descriptions. According to this view, biological evolution does not unfold within a predefined space of functions, traits, or affordances. New functions and phenotypes emerge historically, reshaping the space of possibilities itself. As a consequence, any formalism that presupposes a fixed state space—such as those underlying Shannon-type measures—is argued to be intrinsically inadequate for capturing evolution in its generality. This argument, developed in different forms by Longo, Montévil, and Kauffman [2] and, more recently, by Roli and collaborators [1], places a principled limit on global, semantic interpretations of information in biology. It highlights that biological functions cannot be treated as elements of a predefined alphabet whose meaning is fixed in advance.

At the same time, this critique operates primarily at the semantic and ontological level. It concerns the interpretation of biological novelty and the impossibility of anticipating future functions. It does not, however, deny that evolutionary processes exhibit a robust and reproducible statistical structure at the population level once a selective regime is in place. Empirically, selection is identified through systematic changes in the relative frequencies of variants under comparable conditions. This population-level regularity has been recognised since the origins of population genetics and has been made explicit in later work that treats selection as a transformation of probability distributions. In particular, Frank's information-theoretic and information-geometric analyses show that selection can be characterised abstractly as probabilistic reweighting, independently of the biological mechanisms that implement it [5]–[7]. This establishes a crucial distinction: while the space of biological meanings may be open-ended, the statistical structure of selection within a given regime is well defined and observable.

A large body of work has attempted to connect this statistical structure to informational quantities by focusing on adaptation to structured environments.

Closely related formulations have also emerged in the statistical-physics literature on evolution, where selection is treated as a biasing process acting on ensembles of variants. In this context, evolutionary dynamics are analysed using tools developed initially for non-equilibrium systems, with selective bias expressed as a systematic distortion of probability distributions rather than as an intrinsic property of individual entities. Notably, Sella and Hirsh have shown that several features of evolutionary dynamics can be mapped onto ensemble-based descriptions, highlighting formal analogies between selection, statistical inference, and biased sampling processes [12]. These approaches do not ground selection in physical irreversibility. Still, they reinforce the idea that its observable signature lies in reproducible distributional restructuring, independently of the specific mechanisms by which persistence is realised.

In models of fluctuating environments, organisms are described as adopting phenotypic strategies conditional on environmental cues. Within this framework, the *value of information* is defined operationally as the increase in long-term population growth obtained by exploiting such cues. Classic results by Bergstrom and Lachmann [3] and subsequent refinements by Donaldson-Matasci and collaborators [4] show that, under unconstrained strategy spaces, this fitness gain coincides with the mutual information between environmental states and phenotypic responses. In contrast, under physiological or architectural constraints, mutual information provides an upper bound rather than an equality. These results establish an essential connection between information and growth, but they do so within explicitly defined decision problems, where strategies, cues, and optimisation criteria are fixed in advance.

Further developments have shown that the relation between information and fitness depends sensitively on how information is distributed across individuals. In scenarios where all individuals share the same signal, population growth advantages are well captured by mutual information bounds. In contrast, when individuals receive private and noisy signals, selection among individuals introduces additional contributions that can be expressed as divergence terms between forward and retrospective descriptions of the dynamics, as shown by Rivoire and Leibler [11] and by Kobayashi and Sughiyama [9]. These analyses clarify how selection itself introduces statistical bias. Still, they remain tied to specific sensing-and-growth architectures and do not isolate an object of selection that is invariant across different selective processes.

In parallel, information-theoretic measures have been used to characterise properties of fitness landscapes and evolutionary search in computational and algorithmic contexts. Work in evolutionary computation has shown that entropy- and mutual-information–based descriptors can diagnose problem difficulty and landscape structure in a mechanism-agnostic way. These contributions reinforce the idea that Shannonian quantities capture structured statistical dependence under selection-like updating. At the same time, they remain diagnostic tools: they characterise landscapes or problems, not selection itself as a general population-level phenomenon.

A distinct but related strand of the literature focuses on phenotypic novelty and evolvability. Wagner, in particular, has proposed quantifying the informational content or information cost associated with phenotypes, defined as the number of genotypic configurations that realise a given phenotype and the extent to which that phenotype is distinguished within genotype space [13]. In this perspective, the emergence of new phenotypes can be analysed statistically, and differences between de novo novelty and exaptative reuse can be related to the informational structure of genotype–phenotype maps. These contributions provide important insights into the statistical structure of phenotypes and their capacity to generate further variation. However, the informational quantities involved are marginal properties of genotype–phenotype mappings and do not by themselves define an object of selection. They characterise phenotypes as subsets of a space, rather than as the restructuring of populations under selective persistence.

Taken together, these lines of work demonstrate both the promise and the limitations of existing informational approaches. Information-theoretic quantities are successfully employed to analyse specific aspects of evolutionary processes: adaptation to cues, growth advantages, phenotypic specificity, landscape structure, and bias induced by selection. However, they are introduced in heterogeneous ways, at different levels of description, and often with implicit shifts between population-level and individual-level interpretations. As a result, key distinctions remain blurred. Measures of divergence between distributions are frequently conflated with measures of adaptation; mutual information is sometimes interpreted as a property of populations, phenotypes, organisms, or strategies; and fitness is sometimes treated as an informational quantity in its own right, rather than as a compound outcome influenced by informational and non-informational factors.

This fragmentation points to a deeper underdetermination. The existing literature does not provide a unified population-level definition of selection that cleanly separates: the *strength* of selection as population restructuring; *adaptation* as statistical coupling to an external domain; *evolutionary change* as the iterated effect of selection and variation; and the *value of information* at the level of individual organisms. Without such a separation, informational language remains context-dependent, making it impossible to compare selective processes across biological and non-biological domains in a principled way.

The scope of the present work is defined to address this gap. It does not attempt to assign semantic meaning to biological information, nor to predict evolutionary novelty or enumerate future functions. Instead, it aims to reconstruct the minimal population-level structure required to describe selection once persistence is empirically observed, and to assign precise, non-interchangeable roles to different informational quantities within that structure. By doing so, it seeks to integrate the partial insights of existing approaches while making explicit their limits, and to provide a framework in which selection, adaptation, and information can be related without conceptual conflation. This sets the stage for the population-level analysis developed in the following sections.

3. Darwinian evolution as population dynamics

In this section, Darwinian evolution is described entirely at the population level, as a process acting on probability distributions over selectable configurations.

Descriptions of evolutionary change in terms of probability distributions over variants have a long tradition in theoretical biology [14]–[16]. Classical formulations such as the replicator–mutator framework and its stochastic generalisations already treat populations as distributions evolving under selection and variation, with dynamics expressed directly at the level of frequencies rather than at the level of individual organisms. The present analysis operates at the same descriptive level, while deliberately abstracting from specific dynamical equations. Its aim is not to recover or modify existing models, but to isolate the population-level structure that remains invariant across different implementations of selective dynamics.

The goal is to characterise selection, adaptation, and evolutionary change without introducing organisms as agents or fitness as a primitive quantity. All concepts are defined operationally, starting from what is empirically observable: changes in population composition under reproducible conditions. This choice fixes the explanandum at the level of population restructuring itself, independently of any prior commitment to organisms, fitness concepts, or mechanistic narratives.

3.1 Selectable states and infotypes

Any probabilistic description of selection presupposes that what is selected can be distinguished, counted, and assigned a probability. In the existing literature, this requirement is often satisfied implicitly, allowing the object of selection to shift between individuals, traits, or strategies, with no explicit criterion fixing what counts as a selectable unit. Here, selection is empirically identified through systematic changes in relative frequencies: some configurations persist more often than others across comparable conditions. This constrains the objects that can enter the description. They must be physically instantiated configurations that can be repeatedly realised and whose relative abundance can be tracked at the population level.

We will refer to these selectable units as infotypes. The notion was previously introduced in the literature in a qualitative and conceptual form by Zu Castell and collaborators, to denote the informational basis underlying a system’s performance in its environment, extending beyond the classical genotype and potentially encompassing multiple physical supports [17]. A closely related qualitative notion was also developed independently by the author, motivated by different aims and without a quantitative or formal treatment [18].

Herein, an infotype is an equivalence class of physically instantiated selectable configurations defined at a chosen descriptive resolution for the selective process under consideration. A descriptive alphabet specifies which physical distinctions count as distinct configurations and thereby fixes the admissible support of the population-level analysis. Within that support, physical states are grouped into the same infotype when they are operationally indistinguishable with respect to the persistence criterion used to track selective reweighting across the episode. This step fixes the selectable units on which differential persistence is evaluated. Adaptive ordering relative to a selective domain, when present, is introduced later through the joint distribution between infotypes and domain states. In empirical applications, this descriptive specification must be fixed without reference to the realised selective outcomes and without using the selective-domain variables whose statistical coupling is subsequently quantified.

In biological evolution, infotypes may correspond to genotypes, alleles, regulatory architectures, phenotypic classes, or higher-order heritable configurations. In other selective processes, they may correspond to molecular conformations, immune receptors, behavioural patterns, or internal representational states. What matters is not their material nature but the fact that they define a set of alternatives among which differential persistence can be observed. It is important to note here that infotypes do not correspond to organisms. They are instead the set of characteristics relevant to the selective process, which are not necessarily fixed over time for a given organism but may change, e.g., as the organism develops or acquires new traits (such as knowledge of the environment) relevant to the selective process.

3.2 Populations as probability distributions

Let $\mathcal{I}$ denote the set of admissible infotypes for a given selective process. This set defines the support of the population-level description.

A population is identified with a probability distribution over infotypes,

$$p(i), i \in \mathcal{I}, \sum_i p(i) = 1.$$

This identification requires careful interpretation. The population is not identified with a specific collection of material individuals. Instead, it is identified with the distribution of infotypes instantiated by those individuals. Different material populations may realise the same distribution, and the same material population may be described by different distributions depending on the chosen descriptive resolution.

This is not a departure from standard evolutionary reasoning. In population genetics, the state of a population is given by allele frequencies, not by the enumeration of organisms. The present framework generalises this logic. Infotypes may group individuals into classes, exactly as genotypes or phenotypes do. Selection acts on these classes insofar as their relative frequencies change, and population change is detected as a change in the distribution over infotypes.

Because populations are identified with distributions, selection is necessarily described as a transformation of probability distributions. This is not a modelling convenience but a consequence of how selection is empirically recognised.

3.3 Selection as population restructuring

Rather than treating selection as a property of environments or attempting to derive it from fitness differences, the present framework reconstructs it directly from its observable effect on population composition. A selective episode is defined operationally as any reproducible process that transforms an initial population distribution $p(i)$ into a new distribution $p'(i)$ through differential persistence of infotypes. Empirically, this corresponds to systematic changes in relative frequencies under repeated realisations of comparable conditions.

Any such transformation can be written in the general form

$$p'(i) = \frac{p(i)\, w(i)}{\sum_j p(j)\, w(j)},$$

where $w(i) \geq 0$ is a persistence weight associated with infotype i. This expression introduces no assumptions beyond what is required to represent observed reweighting. The weights are defined only up to a multiplicative constant, reflecting the fact that only relative persistence matters.

At this level, selection is nothing more and nothing less than population restructuring. It is a mapping from one distribution to another induced by differential persistence. No reference to fitness, adaptation, or environment is required at this stage.

A generic consequence of selective reweighting is a change in population structure, but not necessarily a reduction in the dispersion of the population distribution itself. In the classical Darwinian adaptive case, selective reweighting induces statistical dependence between infotypes and the selective domain, even when the population distribution's entropy remains unchanged or increases.

Industrial melanism in *Biston betularia* provides a concrete illustration: after adaptive selection, the colour class becomes predictive of the local selective conditions, so the mutual information between colour and selective domain increases, while the marginal entropy of colour frequencies may remain similar or increase depending on whether the population is defined locally or pooled across environments.

What is generically produced by adaptive selection is not entropy reduction, but the emergence of a structured correlation between population composition and external variables.

3.4 Quantifying selection: divergence between distributions

Once selection is identified with a transformation of population structure, its strength must be characterised by a measure that compares distributions without encoding assumptions about function, adaptation, or optimisation. Because selection is defined as a transformation between distributions over the same support $\mathcal{I}$, its overall effect can be quantified by comparing pre- and post-selection distributions. The appropriate Shannonian quantity for this comparison is the Kullback–Leibler divergence,

$$D_{\mathrm{KL}}(p' \parallel p) = \sum_i p'(i) \log \frac{p'(i)}{p(i)}.$$

In this framework, KL divergence measures the strength of selection understood as population restructuring. It quantifies the extent to which the post-selection distribution departs from the pre-selection distribution.

Two clarifications are essential. First, KL divergence does not measure adaptation. Any selective process producing differential persistence yields a nonzero divergence, regardless of whether persistence depends on environmental structure, stochastic effects, or internal constraints. Second, KL divergence does not measure evolution. It characterises a single selective episode. Evolution requires the coupling of selection to variation across episodes.

KL divergence, therefore, plays a precise and limited role: it measures how strongly selection reshapes a population distribution within a fixed infotype domain.

3.5 Selection conditioned on external variables and adaptive selection

A recurrent source of confusion in the literature is the tendency to identify adaptation with persistence or with individual success, rather than with a population-level ordering relative to external conditions. Differential persistence may arise from intrinsic constraints, demographic effects, or stochastic processes. In such cases, selection produces population restructuring and a nonzero KL divergence without encoding any systematic relation between infotypes and external variables.

A distinct class of selective processes arises when persistence depends systematically on a relation between infotypes and a set of external variables. These variables do not constitute the environment as a whole. They define the selective domain: the subset of external degrees of freedom that actually enter the persistence criterion.

Let $\mathcal{E}$ denote the set of states of the selective domain. The identification of $\mathcal{E}$ is operational and process-relative. It includes exactly those variables whose states discriminate persistence outcomes.

When persistence depends on this relation, the appropriate description involves a joint distribution $p(i,e)$ over infotypes and selective-domain states. This joint distribution is defined on the infotype support already fixed by the descriptive alphabet and the persistence criterion. Mutual information therefore characterises adaptive ordering on a pre-specified support, rather than entering the definition of the infotypes themselves. If the descriptive alphabet or the infotype partition were instead tuned post hoc to maximise the observed dependence on the selective domain, mutual information could be upward biased; for this reason, interpretability requires a pre-specified support and a separately defined selective domain.

Adaptive selection is defined as selective restructuring ordered by the statistical dependence between infotypes and the selective domain. The relevant Shannonian quantity is the mutual information,

$$I(\mathcal{J};\mathcal{E}) = \sum_{i,e} p(i,e) \log \frac{p(i,e)}{p(i)p(e)}.$$

Mutual information quantifies the extent to which infotypes reduce uncertainty about the selective domain and vice versa. This quantity is defined entirely at the population level and refers to the ordering of infotypes across distributions, not to information processing or inference at the level of individual realisations. Adaptive selection is thus characterised as the case in which persistence weights order infotypes according to their informational coupling to $\mathcal{E}$. Strong selective restructuring may occur with negligible mutual information whenever persistence is driven by demographic bottlenecks, internal constraints, or stochastic filtering rather than by systematic coupling to the selective domain.

Crucially, this coupling need not be present prior to selection. Before a selective episode, the distribution of infotypes may be statistically independent of the selective domain, so that mutual information is approximately zero even in the presence of extensive variation. Adaptive selection consists precisely in transforming such an unstructured distribution into one in which infotypes become predictive of the selective domain. This increase in mutual information is independent of whether the entropy of the population distribution decreases, remains constant, or increases.

Mutual information here identifies the object of adaptation, not the total effect of selection. Adaptive selection is a special case of selection, not its definition.

3.6 Variation, regimes, and evolutionary change

The relation between selection as a filtering process and evolution as cumulative historical change is often left implicit, even though it is central to interpreting long-term dynamics. Selection alone cannot produce sustained evolutionary change. Repeated application of the same selective process to the same distribution leads to progressive concentration and eventual stasis. Evolution requires a second ingredient: variation.

Variation is defined by its population-level effect. A variation-generating process redistributes probability mass over the infotype space, modifying the distribution that enters subsequent selective episodes. From an informational standpoint, variation does not generate adaptive information. By redistributing probability mass independently of the selective domain, variation typically reduces or erases previously established statistical dependence between infotypes and external variables. Its evolutionary role is therefore not the introduction of information, but the reopening of the space of selectable configurations on which adaptive selection can subsequently operate. The physical mechanisms underlying variation—mutation, recombination, horizontal transfer, duplication, or reorganisation—are not specified at this level. What matters is how they alter population distributions.

Crucially, not all forms of variation act within the same informational regime. A regime is defined by a fixed infotype domain $\mathcal{I}$ and a fixed selective domain $\mathcal{E}$. Within a regime, population distributions before and after selection are defined on the same support, and quantities such as Kullback–Leibler divergence and mutual information are directly comparable across episodes.

Many forms of variation act within a regime. Point mutations, small recombinations, or incremental modifications introduce new variants by assigning a nonzero probability to infotypes already defined within $\mathcal{I}$. In this case, variation reshapes distributions without altering the domain itself.

Other forms of variation induce regime shifts by changing the effective infotype domain. Horizontal gene transfer provides a clear biological example. When a lineage acquires genetic material from a phylogenetically distinct source, the set of admissible heritable configurations is expanded. Novel combinations of genetic elements become available, and infotypes that were previously inaccessible enter the population with nonzero probability. At the same time, a substantial fraction of the original infotypes—corresponding to the pre-existing genomic background—is conserved across the transition.

In such cases, pre- and post-transition populations are defined on different infotype domains, and the selective ordering induced before and after the event does not refer to the same infotype structure. Informational quantities characterising selection within each phase, therefore, belong, in a strict sense, to different regimes.

This observation directly connects to the objection raised by Longo, Montévil, and Kauffman [2], namely that biological evolution cannot be described within a predefined, fixed space of possibilities, because the space of relevant functions, constraints, and affordances is itself historically generated. The notion of selective regimes provides a precise way to articulate this point without abandoning population-level description. Open-endedness concerns the historical emergence and redefinition of selectable configurations, not the statistical structure of selection once a regime is in place. By making regimes explicit, the framework distinguishes between phases in which informational quantities are well defined on a conserved support and transitions in which that support itself is reorganised. In this sense, non-prestatability is not an obstacle to informational analysis, but a condition that determines the domain within which such analysis is meaningful. Within the present framework, this objection is fully acknowledged at the level to which it applies. Regime shifts correspond precisely to situations in which the infotype domain is not conserved, and therefore no single, globally fixed support is assumed or required.

At the same time, regime shifts do not typically involve a complete loss of continuity. A non-empty subset of infotypes is usually conserved across the transition, providing an overlap between successive domains. This partial conservation makes it possible, once the transition has occurred, to construct an extended infotype support defined as the union of the pre- and post-regime domains. Population distributions before and after the transition can then be represented on this extended support by assigning zero probability to infotypes that are not admissible in a given phase.

In the formalism adopted here, divergence measures no longer describe a single selective process acting on a fixed domain. Instead, they quantify a composite transformation that includes selective reweighting within the conserved portion of the support, together with the expansion of the space of admissible configurations introduced by regime change. This reconstruction is necessarily retrospective and does not provide a basis for predicting the direction, content, or functional role of evolutionary novelty.

In this sense, the framework neither presupposes a globally fixed space of biological possibilities nor renounces informational description when that space changes. Kullback–Leibler divergence remains the appropriate descriptor of population restructuring within regimes. At the same time, extended descriptions allow regime transitions themselves to be situated within a coherent informational framework, without collapsing the emergence of novelty into ordinary selection. Evolution, in this broader sense, consists of sequences of selective regimes connected by variation-induced transitions, whose historical continuity can be analysed quantitatively while remaining compatible with the open-ended character of biological evolution.

3.7 Darwinian evolution as an iterated process across informational regimes

Within a fixed selective regime, Darwinian evolution unfolds as an iterated population-level process structured by the interaction between selection, variation, and the selective domain $\mathcal{E}$. Selection acts as a general filter on persistence, reweighting the population distribution according to differential persistence under given conditions. When differential persistence depends systematically on relations between population structure and the selective domain $\mathcal{E}$, selection operates in an adaptive mode.

Adaptive selection induces statistical dependence between the population distribution of infotypes and the selective domain $\mathcal{E}$. Through this dependence, mutual information between population structure and $\mathcal{E}$ increases, and population organisation acquires an adaptive ordering that reflects the structure of the prevailing selective conditions. Under a stable selective domain, repeated adaptive selection gives rise to extended phases of cumulative adaptive change, during which population structure is progressively refined relative to $\mathcal{E}$.

Selection also produces population restructuring through processes that are not adaptively ordered with respect to $\mathcal{E}$. In these cases, differential persistence alters population composition without establishing a systematic informational relation between population structure and the selective domain. Such events modify boundary conditions for subsequent evolutionary dynamics while remaining informationally unstructured relative to $\mathcal{E}$.

Evolutionary dynamics extend beyond adaptive convergence through population-level processes that generate variation whose statistical structure is independent of the selective domain that shaped the preceding adaptive ordering. This form of variation reshapes the population distribution, reduces the mutual information accumulated under adaptive selection, and restores degrees of freedom in population structure. Subsequent selective episodes, when adaptive, operate on the reorganised distribution and establish new informational couplings relative to $\mathcal{E}$.

A further source of population restructuring arises from changes in the selective domain $\mathcal{E}$ itself. When the conditions defining persistence are altered, the equivalence relations through which organisms (herein intended as selectable configurations) are grouped are redefined. As a consequence, the infotypes associated with the population change even when the underlying organisms remain the same. Population structure is thereby reorganised through a transformation of the informational framework within which selection operates.

Darwinian evolution proceeds as a cycle of informational regimes. Adaptive selection increases mutual information between population structure and a given selective domain $\mathcal{E}$. Variation that reshapes population structure independently of that domain modifies the informational coupling established by prior adaptation. Changes in the selective domain redefine the informational categories relative to which persistence is evaluated. Through the iterated interaction of these processes, populations undergo cumulative change across regimes.

On longer timescales, extended phases of adaptive refinement under stable selective domains alternate with transitions in which population structure is reorganised by stochastic selective events or by shifts in the selective domain. This interplay between relative stability and episodic reorganisation provides a population-level informational interpretation of punctuated equilibria.

This completes the population-level description of Darwinian evolution. Selection appears as a general filtering process on persistence, adaptation as the emergence of statistical dependence between population structure and a selective domain, and evolution as the iterated interaction of adaptive selection, variation that reshapes population structure independently of the current domain, stochastic selective events, and changes in the selective domain itself. Organisms and fitness will be introduced next as derived constructs that realise these population-level processes.

4. Fitness as a function of infotypes

Selection has been described as a transformation of a population distribution $p(i)$ over a set of infotypes $\mathcal{I}$, induced by differential persistence across a selective episode. Within this population-level description, fitness is not taken as an explanatory starting point, but emerges as a functional assignment required to account for how differential persistence induces that transformation. The transformation can be represented as a reweighting,

$$p'(i) = \frac{p(i)\, w_i}{\sum_{j \in \mathcal{J}} p(j)\ w_j},$$

where w_i quantifies the persistence contribution of infotype i in the episode considered. This section introduces fitness as the functional object that generates these weights, derives the mathematical properties implied by the reweighting formalism, and then specifies how an information-dependent contribution is identified by clarifying which populations enter the definition of . The treatment concerns regimes in which the persistence contribution associated with an infotype can be specified for the selective episode under consideration without explicit dependence on current population composition or density. Frequency-dependent and density-dependent cases require a state-dependent extension and are not developed here. The Darwinian biological case is recovered as a particular identification of infotypes with organisms, with persistence realised through reproduction.

4.1 Fitness as a mapping that generates persistence weights

Fitness is introduced as a function

$$f : \mathcal{J} \to \mathbb{R}_{\geq 0},$$

defined operationally through its role in population restructuring. For a fixed selective regime - specified by the set of infotypes $\mathcal{J}$, by the criterion used to assess persistence, and by the restriction that persistence weights are assigned without explicit dependence on the current population state - the persistence weight associated with infotype i is the value returned by f when evaluated at i. This functional assignment is therefore expressed as the mapping

$$f\colon\, i \mapsto w_i.$$

Fitness is distinct from the weights: it is the mapping that assigns a weight to each infotype. This definition does not commit to any particular causal mechanism of persistence. Different physical or biological processes may underlie the same functional assignment, as long as they induce the same population-level transformation. Of course, as already specified, this formulation does not cover regimes in which the persistence contribution of an infotype depends explicitly on current population composition, population density, or other state variables generated by the evolving population itself. Those cases require a state-dependent persistence mapping and are left for future development.

4.2 Properties of the fitness function implied by the reweighting formalism

Several structural properties of f follow directly from the reweighting equation and from minimal probabilistic consistency, once the weights are understood as the outputs of the mapping $f: i \mapsto w_i$.

Non-negativity and the role of zero.
For all admissible population distributions $p(i) \geq 0$, the updated distribution $p'(i)$ must remain non-negative. Since the reweighting uses the weights $w_i$multiplicatively, probabilistic consistency requires

$$w_i \geq 0 \ \forall i \in \mathcal{I},$$

which in turn constrains the range of $f$to $\mathbb{R}_{\geq 0}$. Negative values would violate the probabilistic interpretation of $p'(i)$. A vanishing weight $w_i = 0$ implies $p'(i) = 0$ independently of $p(i)$, corresponding to categorical loss of support within the episode. Within a fixed selective regime, differential persistence is therefore represented by relative reweighting among infotypes with strictly positive weights, while zero marks boundary events such as extinction within the regime or loss of support associated with regime change.

Scale invariance.
Multiplying all outputs of the fitness mapping by a positive constant leaves the population update unchanged. If a rescaled mapping $\tilde{f}$is defined by $\tilde{f}: i \mapsto \widetilde{w}_i$with $\widetilde{w}_i = c\, w_i$for some $c > 0$, then

$$p'(i) = \frac{p(i)\, \widetilde{w}_i}{\sum_j p(j)\, \widetilde{w}_j} = \frac{p(i)\, c w_i}{\sum_j p(j)\, c\, w_j} = \frac{p(i)\, w_i}{\sum_j p(j)\, w_j}.$$

Fitness is therefore defined only up to a global multiplicative factor, and only relative values are meaningful.

Identification through ratios.
From the update rule,

$$\frac{p'(i)}{p'(k)} = \frac{p(i)\,w_i}{p(k)\,w_k},$$

so weight ratios are determined by relative changes in population frequencies:

$$\frac{w_i}{w_k} = \frac{p'(i)/p(i)}{p'(k)/p(k)}.$$

This expresses the operational character of fitness: the mapping f is empirically constrained through the ratios of its outputs inferred from population restructuring.

Multiplicative composition across episodes.
Across successive selective episodes within the same regime and within the state-independent scope specified above, persistence contributions compound multiplicatively. If two episodes induce weight assignments $\left\{w_i^{(1)}\right\}$and $\left\{w_i^{(2)}\right\}$, the combined effect is represented, up to scale, by the product weights $\left\{w_i^{(1)} w_i^{(2)}\right\}$. Correspondingly, the composed fitness assignment has outputs proportional to that product. Taking logarithms on the active support yields additivity of $\log w_i$, which linearises the composition of selective effects. As previously specified, regimes in which persistence contributions vary with population composition or density across episodes require a distinct state-dependent treatment.

4.3 Information-dependent contributions to fitness

At this point, fitness is still treated as a general persistence mapping; the question addressed here is whether some of its components can be isolated based on their dependence on structured statistical relations rather than on their physical or biological origin. Among the possible contributors to persistence, a special role is played by factors that depend on structured covariation with a selective domain $\mathcal{E}$. When such covariation is present, informational quantities defined in Shannon's sense become relevant.

In this setting, the informational quantities under consideration contribute directly to the persistence weight assigned by fitness. During an organism's lifetime, changes in how effectively it exploits structured regularities of the selective domain modify this informational contribution, thereby altering the resulting persistence weight. Within the descriptive alphabet adopted for the selective episode, such changes may shift the organismal realisation from one infotype to another, because the response profile relevant to persistence has changed.

In this situation, the decomposability of the weight assignment supports a factorised representation,

$$w_i^{\text{tot}} = w_i^{\text{info}}\, w_i^{\text{rest}},$$

where w_i^{info}collects the contribution arising from information-dependent effects and w_i^{rest}aggregates all remaining factors. This representation does not presuppose that information dominates persistence; it isolates one multiplicative component whose dependence on a Shannonian quantity can be analysed explicitly. At the same time, the factorisation is introduced at the structural level of the framework and does not by itself imply unique empirical identification of the informational component or of the residual contribution from population reweighting alone.

4.4 Populations involved in the definition of the informational term

The informational component $w_i^{\text{info}}$is defined with respect to a population that is distinct from the population of infotypes over which evolutionary dynamics is described. The distinction concerns which random variable is paired with the selective domain, not the selective domain itself.

At the population-evolutionary level, adaptive selection is characterised by the mutual information between infotypes and the selective domain $\mathcal{E}$. In this case, the two terms entering the joint distribution are:

- $\mathcal{I}$, representing the population distribution over infotypes;
- $\mathcal{E}$, representing the distribution of environmental states relevant to selection.

Here and below, when no ambiguity arises, the same notation is used for the variable and for its support.

By contrast, the informational contribution to persistence at the level of fitness is defined with respect to a different population. Still, it is constructed with respect to the same selective domain $\mathcal{E}$. What changes is the other term entering the joint distribution.

Let R denote the set of admissible responses associated with the realisation of an infotype. Consider a joint distribution $p(r, e)$ over responses $r \in R$ and environmental states $e \in \mathcal{E}$. The mutual information

$$J \equiv I(\mathcal{R}; \mathcal{E}) = \sum_{r,e} p(r,e) \log \frac{p(r,e)}{p(r)p(e)}$$

quantifies the statistical coupling between the organism's possible responses and the same selective domain $\mathcal{E}$.

Thus, both informational quantities considered in the framework involve the same selective domain, but differ in the population on which the corresponding joint distribution is defined: infotypes at the population level, responses at the level relevant for persistence within an infotype realisation. This distinction ensures that the two uses of mutual information are formally homogeneous, since both are Shannonian quantities defined on joint distributions, while remaining conceptually and functionally non-equivalent.

4.5 Mathematical constraints on the informational mapping

Once an informational contribution has been isolated as one factor among others in persistence, its admissible functional form is no longer arbitrary but constrained by the same consistency requirements that apply to persistence weights as such. Because w_i^{info}enters multiplicatively in persistence weights, the mapping $\Phi(J)$ is constrained by the same structural considerations as persistence contributions more generally.

On the active support, $\Phi(J) > 0$. A convenient gauge choice fixes the baseline so that the absence of informational coupling corresponds to neutrality,

$$\Phi(0) = 1.$$

Within a regime where greater statistical coupling corresponds to greater informational contribution, the mapping preserves ordering: larger J does not decrease $\Phi(J)$.

When informational contributions accumulate additively across successive episodes within the same regime, within the state-independent scope specified above,

$$J_{12} = J_1 + J_2,$$

consistency with the multiplicative compounding of persistence requires

$$\Phi(J_1 + J_2) = \Phi(J_1)\,\Phi(J_2).$$

Under standard regularity assumptions, this functional equation yields

$$\Phi(J) = \exp(\alpha J),$$

with $\alpha \geq 0$ setting the scale of informational contribution in that regime; a detailed derivation is provided in **Supplementary Material 2**. What is established here is the admissible functional form of the informational persistence component under the structural assumptions of the framework. This result does not by itself provide unique empirical identification of or of the decomposition between informational and residual persistence factors from population reweighting alone.

4.6. Linking individual information extraction and population-level adaptation

The analysis developed in this section connects two uses of Shannon mutual information that involve the same selective domain $\mathcal{E}$ but are defined on different variables and at different levels of description. The mutual information defined on the joint distribution over infotypes and states of the selective domain $\mathcal{E}$, introduced in Section 3, characterises adaptive selection at the population level. The present section has introduced a distinct informational quantity, J, defined on the joint distribution over environmental states and the responses produced by organisms realising a given infotype, and has shown how this quantity contributes to fitness.

The two informational quantities operate at different levels and play different roles. The quantity J, defined on the joint distribution over responses and states of the selective domain $\mathcal{E}$, captures how, for organisms instantiating a given infotype, the statistical coupling between environmental states and their responses contributes to persistence. It is evaluated on the population of organismal realisations associated with that infotype and quantifies the functional value of information within a given selective regime. By contrast, the mutual information defined on the joint distribution over infotypes and states of the selective domain $\mathcal{E}$characterises how the population distribution over infotypes reflects regularities in the selective domain as a consequence of differential persistence.

The connection between these levels is mediated entirely by fitness. At the level of organismal realisations, informational coupling between responses and the selective domain affects fitness by modulating persistence. When fitness depends on this informational term, the resulting persistence weights differ across infotypes. Because population updating is defined on the distribution over infotypes fixed for the regime, such differences translate into differential representation in the subsequent population distribution.

Through this mechanism, information extracted at the level of organismal realisations influences population restructuring without altering the population-level nature of selection. Individual organisms realise an infotype at a given time and may be reassigned to a different infotype when changes in the organismal state alter the values of the descriptors relevant under the chosen alphabet across the selective episode. Selection, however, acts on the population distribution over infotypes, not on individual trajectories.

No inverse inference is implied. The presence of mutual information defined on the joint distribution over infotypes and states of the selective domain at the population level does not require that individual organisms optimise J, nor that information extraction be adaptive when considered in isolation. Population-level mutual information arises from the aggregation of differential persistence across many organismal realisations, mediated by fitness, rather than from optimisation at the level of individuals.

The same Shannonian formalism thus operates coherently across levels. J quantifies the functional contribution of information to persistence within an infotype. Mutual information defined on the joint distribution over infotypes and states of the selective domain $\mathcal{E}$ quantifies adaptive population restructuring. Fitness provides the sole link between these two descriptions, translating individual-level informational effects into population-level evolutionary outcomes.

5. Recapitulating the Darwinian biological case

In the Darwinian biological setting, the abstract structure introduced above admits a direct and concrete interpretation. Infotypes can be identified with equivalence classes of heritable biological configurations whose persistence across time is realised through reproduction. Depending on the descriptive alphabet adopted, these configurations may correspond to genotypes, regulatory states, or equivalence classes of organisms sharing transmissible traits. This choice affects the resolution of the description, but not the population-level formalism: in all cases, infotypes are the distinguishable states whose frequencies are tracked and reweighted under selection.

Under this identification, the fitness assignment $f: i \mapsto w_i$ acquires a clear operational meaning. Fitness corresponds to a biological persistence weight, implemented through differential reproductive success. Reproduction is the physical mechanism by which persistence is realised in this domain, but it does not alter the general structure of selection as population-level reweighting. What matters for the formalism is that heritable configurations contribute unequally to future populations, thereby inducing a systematic transformation of the distribution over infotypes.

This perspective clarifies the role of reproduction in Darwinian evolution. Reproduction does not define selection; it implements persistence. As a consequence, only configurations that are stably transmitted can participate in cumulative evolutionary dynamics. From an informational standpoint, reproduction acts as a constraint on variability: it determines which redistributions of probability mass can accumulate across selective episodes.

Adaptive selection in the Darwinian regime corresponds to the case in which persistence weights depend systematically on the relation between infotypes and a selective domain D. The selective domain is not the environment as a whole, but the subset of external variables that effectively enter the criterion of differential reproductive success. Its structure is shaped by organismal physiology, developmental constraints, and ecological interactions, and is therefore neither exhaustive nor neutral.

Within this setting, adaptive evolution manifests as a population-level ordering of infotypes according to their mutual information with D. Infotypes that are more strongly coupled, in a statistical sense, to the relevant states of the selective domain acquire higher persistence weights and become enriched in the population. This ordering is not imposed at the level of individual organisms, nor does it require explicit representation of the environment. It emerges from differential survival and reproduction, and is defined entirely at the population level.

The Darwinian case also makes explicit the distinction between informational and non-informational contributions to fitness introduced earlier. Total reproductive fitness reflects multiple factors, only one of which is adaptive in the informational sense. Energetic costs, developmental trade-offs, demographic effects, and ecological constraints modulate persistence without contributing to the ordering of infotypes by their coupling to D. As a result, strong selection, measured as substantial restructuring of population distributions, need not coincide with strong adaptive ordering, and vice versa.

Variation redistributes probability mass over the space of heritable infotypes through mutation, recombination, and related processes. In typical Darwinian regimes, this redistribution occurs within a conserved infotype space, so that population distributions before and after selection are defined on the same support. Under these conditions, the informational measures introduced above provide well-defined descriptors of cumulative selective change.

In this sense, the Darwinian biological case constitutes a canonical instantiation of the general framework developed in this work. It realises persistence through reproduction, adaptive ordering through coupling to a biologically constrained selective domain, and cumulative change through repeated cycles of selection and heritable variation, without requiring any modification of the population-level informational structure. A worked field-style illustration of this construction, based on Darwin’s finches treated as replicated selective systems, is provided in **Supplementary Material 1**.

6. Beyond the Darwinian case: changing regimes, supports, and persistence mechanisms

The population-level description developed in the previous sections deliberately abstracts from the biological specifics of Darwinian evolution. Selection has been characterised as a probabilistic restructuring of a population distribution over infotypes, driven by differential persistence under a specified criterion. Adaptation has been identified as a particular ordering of infotypes induced by their statistical coupling to a selective domain, quantified by mutual information. Evolution, in turn, has been defined as the iterated interaction of selection and variation within and across regimes.

This formulation allows the analysis of selective processes that depart from the canonical Darwinian case while preserving the same formal structure. What varies across such processes is not the logic of selection, but the identification of infotypes, the mechanism by which persistence is realised, the nature of the selective domain, and the structure of the variability-generating process. Together, these elements define a *selective regime*.

6.1 Regimes and invariants of the formalism

Four components specify a selective regime:

1. a support of infotypes, defining the admissible selectable configurations;
2. a persistence criterion, determining how infotypes are differentially retained;
3. a selective domain, consisting of the external variables that condition persistence;
4. a mechanism of variation, reshaping the population distribution between selective episodes.

Within a fixed regime, population distributions before and after selection are defined on the same support, and informational quantities such as Kullback–Leibler divergence and mutual information are directly comparable. Across regime changes, by contrast, these quantities remain well-defined within each regime but lose their direct comparability across regimes, because the underlying infotype domain has changed.

What remains invariant is the probabilistic structure of selection itself: a population of physically instantiated states, a reproducible reweighting induced by differential persistence, and a redistribution of probability mass over a space of alternatives. This invariant structure allows the exact Shannonian quantities to be used across regimes, while strictly limiting the scope of each application.

6.2 Adaptive but non-evolutionary selection: the immune repertoire

The adaptive immune system provides an instructive example. Here, the infotypes are not organisms or genotypes, but molecular configurations of antigen receptors. Each lymphocyte clone instantiates an infotype defined by the binding properties of its receptor, and the immune repertoire at a given time is described by a population distribution over such configurations.

Persistence is realised through clonal expansion and survival. Exposure to an antigen constitutes a selective episode that reproducibly reweights the population: receptor configurations that bind the antigen more effectively acquire higher persistence weights. The selective domain is defined by antigenic structure, and adaptive selection is present in the precise sense described earlier. At the population level, the post-selection distribution exhibits a positive mutual information between receptor configurations and antigenic states.

Despite this, the immune system does not exhibit evolution in the sense defined in this work. The crucial reason lies in the structure of variation. The enormous diversity of receptor configurations is generated primarily once, during early immune development, through recombination processes largely independent of prior selective outcomes. After each selective episode, the system is driven back toward a baseline distribution that is statistically similar across episodes, except for a limited contribution from memory cells.

In informational terms, variation acts to restore an approximately fixed prior distribution. Selection is adaptive and reproducible, but cyclic rather than cumulative. The immune response, therefore, exemplifies adaptive selection without evolution: a regime in which mutual information with the selective domain is generated and exploited, yet does not accumulate across episodes.

6.3 Regime shifts and changes in the support of infotypes

Other departures from the Darwinian case involve changes in the support of the infotype itself. Horizontal gene transfer provides a paradigmatic biological example. When genetic material is acquired from an external lineage, the space of admissible heritable configurations is altered. New combinations become possible, existing configurations acquire new functional roles, and the relation between infotypes and the selective domain is reorganised.

Such events cannot be treated as ordinary variation within a fixed regime. They induce a *regime shift* in which the infotype domain itself is redefined. Although pre- and post-transition configurations can often be embedded in a common extended space, the selective ordering is no longer defined with respect to the same infotype structure. Informational quantities characterising selection before and after the transition belong to different regimes and cannot be directly compared.

Genome duplication, large-scale rearrangements, and regulatory reorganisation have similar effects. Evolution, in this broader sense, proceeds as a sequence of regimes: extended periods of cumulative change within stable supports, punctuated by transitions that redefine what counts as a selectable configuration.

6.4 Persistence without reproduction: cognitive and epistemic systems

The formalism also applies to selective processes in which persistence is not realised through reproduction. Consider a cognitive or epistemic system in which infotypes correspond to stable informational structures: internal representations, models, or hypotheses that can be distinguished, counted, and tracked at the population level. The population distribution now represents the relative availability or effective presence of such structures within a cognitive system or a community.

Persistence is realised through continued usability rather than numerical amplification. An infotype persists insofar as it continues to function as an operative structure: it organises experience, supports reliable expectations, and integrates coherently with other structures. The selective domain comprises salient facts of the physical world that must be accounted for across repeated interactions.

In this context, adaptive selection occurs when persistence depends systematically on the statistical coupling between infotypes and the selective domain. Infotypes that reduce uncertainty about relevant aspects of the world acquire higher persistence weights and become more prevalent in the population of available structures. Mutual information again quantifies this adaptive ordering at the population level.

Variation in this regime is strongly history-dependent. New infotypes are generated by recombination, abstraction, and extension of previously stabilised structures. The variability-generating process therefore does not restore a fixed baseline distribution, but reshapes the population in a way that depends on past selective outcomes. Under these conditions, selection and variation jointly produce cumulative reorganisation. In the informational sense defined here, the process qualifies as evolutionary, despite the absence of biological reproduction.

6.5 Scope and limits of the introduced informational framework

These examples illustrate how changing the support of infotypes, the mechanism of persistence, the selective domain, and the structure of variation yields selective and evolutionary processes that differ profoundly in their physical realisation, while remaining describable within a single formal framework. What unifies them is not a shared mechanism, but a shared probabilistic structure.

At the same time, the framework has explicit limits. The present formalisation of fitness and of its informational component is restricted to regimes in which persistence weights can be assigned to infotypes without explicit state dependence on current population composition or density. Frequency-dependent and density-dependent cases require an extended treatment. Informational quantities are defined only relative to a specified regime. They do not predict the emergence of new infotype supports, nor do they assign semantic meaning to novelty. They characterise selection and adaptation *once a regime is operative*, not the historical process by which regimes arise.

Within these limits, the framework provides a principled way to compare selective processes across biological and non-biological domains, to distinguish adaptive selection from cumulative evolution, and to identify which features of a system are responsible for long-term change. By making regimes explicit, it avoids the common error of treating informational measures as universally comparable or semantically loaded, while preserving their explanatory power where they are structurally warranted.

7. Conclusions

The problem posed in the opening sections of this work concerned the status of informational descriptions in evolutionary theory. The state of the art makes clear that information-theoretic quantities have been applied in many evolutionary contexts, yet often with shifting meanings, overlapping scopes, and implicit assumptions about fitness, adaptation, or individual agency. The question motivating this manuscript was therefore not whether information can be used in evolutionary models, but how such use can be stabilised so that different informational quantities retain distinct and non-interchangeable roles.

The population-level framework developed in the preceding sections provides a precise way to address this question. By identifying the population distribution over selectable configurations as the primary object of description, the analysis fixes the level at which informational quantities apply. Selection is expressed as a reproducible transformation of this distribution, making probabilistic reweighting the basic observable of selective dynamics. Within this setting, the use of divergence measures follows directly from the structure of the problem and characterises the extent of population restructuring induced by selection.

This characterisation allows selective change to be discussed independently of its causes. Differential persistence may arise from environmental ordering, internal constraints, stochastic effects, or combinations thereof. The magnitude of restructuring is captured by divergence between distributions, while the nature of that restructuring requires additional specification. This separation responds directly to a recurrent ambiguity in the literature, where selective strength and adaptive significance are often treated as inseparable.

Adaptation enters the description through the explicit introduction of a selective domain. When population restructuring is systematically ordered with respect to this domain, population structure reflects regularities of the domain itself. The mutual information between the population distribution of infotypes and the selective domain provides a quantitative descriptor of this relation. In this formulation, adaptation is neither inferred from fitness values nor attributed to individual success. It appears as a statistical property of population structure conditioned on external variables.

The same informational formalism also accommodates the role of information at the level relevant for persistence. Infotypes may differ in how their possible realisations couple to the selective domain, and this coupling can affect persistence through fitness. Mutual information defined on the distribution of possible responses associated with an infotype quantifies this contribution. Although formally identical to the population-level quantity, it refers to a different population and fulfils a different role. The framework keeps these quantities distinct by construction, preventing the common conflation of individual information use with population-level adaptation.

Fitness mediates between these levels. It is introduced as a functional mapping from infotypes to persistence weights, inferred from observed changes in population distributions. Informational contributions enter fitness insofar as they affect persistence within a given regime, alongside non-informational factors. This placement clarifies how information can influence evolutionary outcomes without being identified with fitness itself or with adaptive ordering. The organism-level treatment developed here concerns regimes in which persistence weights can be assigned to infotypes without explicit dependence on current population composition or density. Frequency-dependent and density-dependent cases require a state-dependent extension of the present framework.

The treatment of variation and regimes further refines the interpretation of informational measures over time. Variation reshapes population distributions and prevents convergence under selection, enabling cumulative change within regimes where the support of infotypes is conserved. In such regimes, informational quantities remain comparable across selective episodes. When the support itself is reorganised, as in horizontal gene transfer or large-scale regulatory changes, the framework identifies a transition between regimes. Once a regime transition has occurred, the partial conservation of selectable configurations allows pre- and post-transition populations to be embedded retrospectively in an extended support, making informational comparisons possible at the level of historical transformation without presupposing a globally fixed space of biological possibilities. Informational descriptors remain well defined within each regime, while their loss of direct comparability across regimes marks a change in the structure of the selective problem.

Within this general setting, Darwinian evolution appears as a specific class of selective processes. It corresponds to regimes in which selectable configurations are heritable, variation operates within a conserved domain, and persistence is realised through reproduction. These conditions generate a cyclic interaction between selection and variation that allows cumulative evolutionary change. Other selective processes share only part of this structure and are therefore described by the same formalism without exhibiting Darwinian evolution in this sense.

Taken together, the framework developed in this work provides a way to reorganise informational approaches to evolution by making explicit the populations, selective domains, and persistence mechanisms to which informational measures are applied. In this way, apparent ambiguities surrounding the use of information in evolutionary contexts can be traced to shifts in level or object, rather than to intrinsic limitations of informational formalisms.

Under this reading, informational quantities do not directly define adaptation, nor do they presuppose predefined spaces of functions or affordances. They characterise how selective domains structure persistence across levels, with fitness as the sole mediator between functional effects at the level of realisations and adaptive restructuring at the population level. This resolves the original objection motivating the analysis without weakening the Darwinian interpretation of selection.

Bibliography

[1] A. Roli, S. Patra, and S. Kauffman, “The creation of information: how evolution generates novel improvisations in the biosphere,” *Interface Focus*, vol. 15, no. 6, Dec. 2025.

[2] G. Longo, M. Montévil, and S. Kauffman, “No entailing laws, but enablement in the evolution of the biosphere,” in *GECCO’12 - Proceedings of the 14th International Conference on Genetic and Evolutionary Computation Companion*, 2012, vol. 279, no. 1743, pp. 1379–1391.

[3] C. T. Bergstrom and M. Lachmann, “Shannon information and biological fitness,” *Proc. IEEE*, vol. 92, no. 3, pp. 515–528, 2004.

[4] M. C. Donaldson-Matasci, C. T. Bergstrom, and M. Lachmann, “The fitness value of information,” *Oikos*, vol. 119, no. 2, pp. 219–230, 2010.

[5] S. A. Frank, “The common patterns of nature,” *Journal of Evolutionary Biology*, vol. 22, no. 8. pp. 1563–1585, 2009.

[6] S. A. Frank, “Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory,” *J. Evol. Biol.*, vol. 25, no. 12, pp. 2377–2396, 2012.

[7] S. A. Frank, “Natural selection. VI. Partitioning the information in fitness and characters by path analysis,” *J. Evol. Biol.*, vol. 26, no. 3, pp. 457–471, 2013.

[8] M. Hilbert, “Linking Information, Knowledge and Evolutionary Growth: A Multilevel Interplay between Natural Selection and Informed Intervention,” *SSRN Electron. J.*, 2014.

[9] T. J. Kobayashi and Y. Sughiyama, “Fluctuation Relations of Fitness and Information in Population Dynamics,” *Phys. Rev. Lett.*, vol. 115, no. 23, p. 238102, 2015.

[10] V. Mustonen and M. Lässig, “Fitness flux and ubiquity of adaptive evolution,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 9, pp. 4248–4253, 2010.

[11] O. Rivoire and S. Leibler, “The Value of Information for Populations in Varying Environments,” *J. Stat. Phys.*, vol. 142, no. 6, pp. 1124–1166, 2011.

[12] G. Sella and A. E. Hirsh, “The application of statistical physics to evolutionary biology,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 27, pp. 9541–9546, 2005.

[13] A. Wagner, “Information theory, evolutionary innovations and evolvability,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 372, no. 1735, p. 20160416, Dec. 2017.

[14] M. Eigen, “Self-organization of matter and the evolution of biological macromolecules,” *Naturwissenschaften*, vol. 58, no. 10, pp. 465–523, 1971.

[15] M. A. Nowak, *Evolutionary dynamics: exploring the equations of life*, vol. 44, no. 08. Harvard University Press, 2007.

[16] N. Champagnat, R. Ferrière, and S. Méléard, “Individual-Based Probabilistic Models of Adaptive Evolution and Various Scaling Approximations,” in *Progress in Probability*, vol. 59, 2008, pp. 75–113.

[17] W. Zu Castell, U. Lüttge, and R. Matyssek, “Gaia-A holobiont-like system emerging from interaction,” in *Emergence and Modularity in Life Sciences*, Cham: Springer International Publishing, 2019, pp. 255–279.

[18] E. Bucci, *Geni, Memi e Bit*. Mondadori Università, 2024.

Supplementary material 1

An ideal field experiment: Darwin’s finches as replicated selective systems

The Galápagos finches constitute a paradigmatic evolutionary case because the empirical logic underlying their interpretation is the same one Charles Darwin employed qualitatively. Darwin’s argument proceeds by comparison: finches collected on different islands are clearly related, yet their beaks differ in consistent and repeatable ways, and these differences covary with what each island offers as food. Islands, therefore, function as natural replicates of the same evolutionary process, in which the organismal background is broadly comparable while the environmental composition differs. The construction developed here follows the same logic, making explicit—step by step—the quantitative objects that remain implicit in the classical narrative and maintaining a strict conceptual distinction between Darwinian (reproductive) fitness and persistence-based informational effects.

Each island is treated as a replicate $r \in \{1, \dots, R\}$. Replicate r yields two empirical descriptions, collected independently: a distribution of beak morphologies in the local population and a distribution of food resources in the local environment. What varies across islands is not a single environmental label but the entire distribution over an environmental domain E. This mirrors Darwin’s own descriptions, which repeatedly refer to mixtures of resources rather than to a unique food type characterising each island.

The construction begins with beak morphology. On each island r, a sample of N_r individuals is collected, and beak depth, width, and length are measured. These are precisely the morphological features Darwin relied upon when distinguishing robust, slender, elongated, or short beaks across islands, here rendered in quantitative form. All measurements from all islands are pooled and summarised by a principal component analysis. The first principal component defines a one-dimensional coordinate, z, that captures the dominant axis of beak variation across the archipelago, the quantitative analogue of Darwin’s recurring contrast between robust and slender beaks. The continuous coordinate z is discretised into K globally fixed bins, defining the set of distinguishable morphological states, or infotypes,

$$T = \{t_1, \dots, t_K\}.$$

This construction is chosen for generality rather than necessity. In cases where the phenotype of interest can be directly reduced to a small number of known genes and allelic variants, the definition of the infotype is simpler: the alphabet T can be taken directly as the set of genotypes or alleles, without introducing morphometric axes or discretisation schemes. In the present case, however, the morphometric approach is deliberately adopted. It is historically faithful to Darwin's own practice, which relied entirely on comparative morphology, and it remains the most general strategy available. It applies equally when the genetic basis of a trait is unknown, polygenic, developmentally mediated, or environmentally modulated. For these reasons, morphometric descriptors are, both historically and in contemporary practice, the primary empirical basis for reconstructing phylogeny and documenting evolutionary divergence.

For each replicate r, the empirical distribution of infotypes is estimated as

$$P_r(T = t_k) = \frac{n_r(t_k)}{N_r},$$

where $n_r(t_k)$ is the number of sampled individuals on island r whose beak falls in class , thist_k. This distribution describes the population composition observed in that replicate. At this stage, it encodes neither function nor reproductive success, but only the relative presence of distinguishable phenotypic configurations.

Independently, the environmental domain is constructed from resource measurements. On each island r, seeds are sampled using standardised ecological procedures. For each seed, physical properties such as size and mechanical resistance are measured. These measurements define an empirical resource space. This space is discretised into M globally fixed resource classes,

$$\mathcal{E} = \{e_1, \dots, e_M\},$$

each corresponding to a physically distinguishable type of food item. For replicate r, the empirical distribution of resources over the environmental domain is estimated as

$$P_r(E = e_j) = \frac{m_r(e_j)}{\sum_{\ell=1}^{M} m_r\,(e_\ell)},$$

where $m_r(e_j)$ denotes the abundance or biomass of resources in class $e_j$observed on island r. This construction is the quantitative counterpart of Darwin's repeated references to local diets and feeding conditions, expressed as a distribution rather than as a qualitative label.

The dataset thus consists of a collection of replicate-level pairs $\{(P_r(T), P_r(E))\}_{r=1}^{R}$. To move from replicate-level objects to system-level distributions, observational weights $\alpha_r \geq 0$ are introduced, with $\sum_r \alpha_r = 1$. These weights encode each replicate's contribution to the overall estimate. They are determined from directly observable features of the sampling design, such as the number of individuals measured or the extent of ecological sampling. When sampling effort is homogeneous across islands, $\alpha_r = 1/R$; when it is not, the weights correct for that heterogeneity. Using these weights, the aggregated marginals are defined as

$$P(T) = \sum_{r=1}^{R} \alpha_r \; P_r(T), P(E) = \sum_{r=1}^{R} \alpha_r \; P_r(E),$$

and the aggregated joint distribution as

$$P(T, E) = \sum_{r=1}^{R} \alpha_r \; P_r(T) \, P_r(E).$$

This joint distribution formalises the empirical observation that, across islands, particular population compositions systematically co-occur with environmental compositions, in precisely the comparative sense Darwin employed.

From the joint distribution, conditional distributions are obtained in the standard way:

$$P(T \mid E = e_j) = \frac{P(T, E = e_j)}{P(E = e_j)}.$$

The mutual information between population structure and environmental structure quantifies selective pressure at the level of the replicated system,

$$I(T; E) = \sum_{k=1}^{K} \sum_{j=1}^{M} P(\, t_k, e_j) \log \frac{P(t_k, e_j)}{P(t_k) \, P(e_j)}.$$

This scalar quantity measures the degree to which the distribution of beak forms departs from statistical independence with respect to the distribution of food resources. In Darwin’s terms, it expresses how much information about beak morphology is gained by knowing the dietary context. The same quantity admits the decomposition

$$I(T;E) = \sum_{j=1}^{M} P(e_j)\, D_{\mathrm{KL}}\big(P(T \mid E = e_j) \;\|\; P(T)\big),$$

with

$$D_{\mathrm{KL}}\big(P(T \mid E = e_j) \;\|\; P(T)\big) = \sum_{k=1}^{K} P(t_k \mid e_j) \log \frac{P(t_k \mid e_j)}{P(t_k)}.$$

This form makes explicit that selective pressure corresponds, in informational terms, to systematic distortions of the baseline population distribution $P(T)$ associated with different components of the environmental domain.

Within the same structure, the persistence-based informational reweighting of an infotype $t_k$conditional on an environmental state $e_j$is defined as

$$w_{\mathrm{info}}(t_k \mid e_j) = \frac{P(t_k \mid e_j)}{P(t_k)}.$$

This quantity measures the relative persistence of infotypes in association with specific environmental components. It does not represent Darwinian fitness, which is reproductive and refers to differential survival and reproduction across generations. Here, persistence is assessed at the level of population composition observed across replicated environments, abstracting from the biological mechanisms—reproductive or otherwise—that underlie that composition.

When this protocol is applied to the finches, the outcome reproduces Darwin's conclusions in explicit quantitative form. Replicates whose resource distributions $P_r(E)$ are skewed toward large and mechanically resistant seeds increase the joint mass $P(t_k, e_j)$ for robust beak classes t_k paired with hard-seed resource states e_j. This produces larger Kullback–Leibler contributions for those components and yields a positive mutual information $I(T;E)$ across the archipelago. Replicates dominated by finer resources generate complementary reweightings. The familiar Darwinian picture—closely related populations diverging across islands in association with local food conditions—thus emerges as an explicit set of empirical distributions and information-theoretic quantities.

What has been achieved is a translation of a paradigmatic historical argument into a form that is directly testable and comparable. The classical narrative is preserved, but expressed in terms of distributions, joint probabilities, and divergences that can be estimated from field data. The distinction between reproductive fitness and informational persistence clarifies roles rather than multiplying entities: reproductive fitness remains the causal engine of evolutionary change. In contrast, persistence captures the statistical imprint that this engine leaves on population structure when observed across heterogeneous environments. In this sense, the formalism is not an optional reformulation but a natural consequence of treating Darwin's comparative reasoning as a statement about measurable associations between populations and environments.

The symmetric use of mutual information as Darwinian inference

A defining property of mutual information is its exact symmetry,

$$I(T;E) = I(E;T),$$

which follows directly from its definition,

$$I(T;E) = \sum_{t,e} P(t,e) \log \frac{P(t,e)}{P(t)\,P(e)}.$$

The expression depends only on the joint distribution and on the marginals and is therefore invariant under exchange of the two variables. In intuitive terms, mutual information quantifies how strongly two descriptions of the same system constrain one another. Because the constraint is a property of their joint regularity, it does not privilege either description: the same structure that allows one to infer phenotypes from environments allows one, with equal informational content, to infer environments from phenotypes.

This symmetry closely corresponds to the inductive logic that characterised Charles Darwin's evolutionary reasoning. Darwin's analysis of the Galápagos finches did not proceed by controlled experiments, but by systematic comparison across islands. He repeatedly observed that particular beak forms recurred in association with specific kinds of food resources, and that similar ecological conditions on different islands tended to be accompanied by similar morphological solutions. From these repeated associations, he inferred both directions at once: that environments shape morphology, and that morphology reveals something reliable about the environment in which it evolved.

Within the present framework, once the joint distribution $P(T, E)$ has been constructed across replicated islands, the symmetric inferential direction is expressed by the conditional distribution

$$P(E \mid T = t_k) = \frac{P(T = t_k, E)}{P(T = t_k)}.$$

For a given infotype t_k, this distribution specifies how the environmental domain is structured across those islands where that morphology is observed. Empirically, it answers the same question Darwin addressed qualitatively: given a particular form of beak, what kinds of feeding conditions tend to accompany it across the archipelago?

Darwin's notebooks and later discussions show that this was not an incidental inference. Robust, deep beaks were repeatedly taken as evidence of feeding on large or mechanically resistant seeds; slender or elongated beaks were associated with finer food sources. These conclusions were not drawn from a single island, but from the recurrence of the same association across multiple islands, which serve as natural replicates. The conditional distribution $P(E \mid T = t_k)$ formalizes exactly this inductive step: it aggregates those recurrences into a probabilistic description of the environments compatible with a given phenotype.

The informational weight of this inductive inference is made explicit by the Kullback–Leibler decomposition associated with the symmetric form of mutual information,

$$I(T; E) = \sum_{k=1}^{K} P(t_k)\, D_{\mathrm{KL}}\left(P(E \mid T = t_k) \;\|\; P(E)\right),$$

with

$$D_{\mathrm{KL}}\left(P(E \mid T = t_k) \;\|\; P(E)\right) = \sum_{j=1}^{M} P(e_j \mid t_k) \log \frac{P(e_j \mid t_k)}{P(e_j)}.$$

Each infotype contributes according to how selectively it constrains the environmental distribution. This mirrors Darwin's practice of treating some morphological traits as more informative than others. Beak forms that appeared across a wide range of islands and feeding conditions carried little inductive weight.

Beak forms that recurred only in association with specific food types carried strong inferential force.

In the finches, robust beak morphologies exemplify the latter case. Their occurrence is concentrated on islands where large and mechanically resistant seeds dominate. The corresponding conditional distributions $P(E \mid T)$ are sharply structured and depart strongly from the baseline environmental distribution $P(E)$. Such morphologies function as reliable indicators of environmental structure, in precisely the sense Darwin exploited when reasoning from form to function and from morphology to niche.

The same inferential structure underlies a much broader class of scientific practices, most prominently palaeontology. In that domain, direct observation of the environment is unavailable. Inference proceeds almost entirely from phenotype to environment: tooth shape is used to reconstruct diet, limb proportions to infer locomotion and substrate use, and skeletal robustness to infer mechanical loads and habitual behaviour. These inferences rely on repeated associations established across living or well-characterised systems and projected onto extinct ones. What is inferred is not a single environmental state, but a constrained distribution over plausible environments, precisely of the form $P(E \mid T)$.

The mutual-information framework clarifies why such reconstructions are possible and why they differ in strength across traits. Some phenotypic features are weakly informative because they are compatible with many ecological contexts; their associated conditional distributions are broad and close to the baseline. Other features sharply constrain the set of compatible environments, thereby supporting strong ecological inference. The Kullback–Leibler contribution associated with each infotype provides a quantitative expression of this familiar distinction.

Seen in this light, the symmetry of mutual information captures a general mode of scientific induction. Darwin’s comparative reasoning, paleontological reconstruction, functional morphology, and ecological inference all rely on the same structure: stable, repeated associations between forms and conditions, accumulated across replicates. The formalism does not replace these practices but makes explicit the informational content they exploit. It shows that inference from morphology to environment is not a heuristic shortcut but the symmetric, quantitatively well-defined counterpart of selection analysis itself.

Supplementary material 2

Justification of the functional form of the informational persistence component

In the main text, the informational component of persistence, w_{info}, is introduced as the factor that implements adaptive ordering by coupling infotypes to a selective domain through mutual information. Its exponential dependence on informational gain is stated there as a structural consequence of the framework. For clarity and continuity of exposition, the main text presents this result without a detailed derivation.

The purpose of the present supplementary material is to make explicit the constraints that uniquely fix the admissible functional form of w_{info}. This analysis does not introduce additional biological assumptions, nor does it rely on specific mechanistic models of selection. Instead, it articulates the minimal structural requirements that any informational persistence mapping must satisfy to be coherent with: (i) the interpretation of fitness as a relative persistence weight inferred from population reweighting, (ii) the multiplicative composition of persistence effects, and (iii) the additive structure of informational measures such as mutual information.

The starting point is the mapping

$$w(x) = f(J(x)),$$

which links the informational gain $J(x)$associated with infotype $x$to its persistence weight. If the function $f$were left completely arbitrary, the formalism would lose substantive content: any observed population restructuring could always be justified by choosing an ad hoc functional form. It is therefore necessary to constrain the structure of $f$explicitly, not to determine its exact form in every possible case, but to restrict it to forms that are coherent with the meaning of the quantities involved.

The first constraint is conceptual, yet non-trivial: the function must be monotonically increasing. If an infotype realises a more effective informational coupling with the relevant selective domain, it cannot be systematically penalised by the selective process; otherwise, the notion of informational selection itself would lose meaning. Informational gain must therefore translate into non-decreasing persistence.

A second constraint concerns the absolute scale of fitness. Persistence weights are inferred from population reweighting and are defined only up to a global multiplicative constant. Multiplying all fitness values by a positive constant does not alter the post-selection distribution, since the constant cancels out upon normalisation. What matters is not the absolute value of w, but the relative comparison between persistence weights of different infotypes. Within the present framework, fitness is therefore not an absolute quantity, but a comparative measure of persistence.

A third constraint follows from the composition of selective effects. Any real selective process involves multiple contributions to persistence, whether interpreted as successive episodes or as independent components acting within the same episode. At the level of persistence weights, these contributions compose multiplicatively: being favoured twice corresponds to a total weight equal to the product of the individual weights. It follows that the structurally natural quantity is not w itself, but $\log w$, since logarithms transform products into sums. This observation places the present treatment within a well-established tradition. In models of multiplicative growth and in Malthusian fitness theory, it is the logarithm of fitness—not its absolute value—that has dynamical significance, composing additively and determining long-term growth rates. The central role of $\log w$is therefore not specific to the informational framework, but a general property of selective processes in which effects accumulate multiplicatively.

The argument becomes decisive once the structure of information is taken into account. Informational measures relevant to adaptive selection, such as mutual information, are additive across independent contributions. If a total informational gain can be decomposed as

$$J = J_1 + J_2,$$

Coherence requires that the overall selective effect correspond to the composition of the effects associated with the individual contributions. Translated into fitness terms, this amounts to imposing the functional condition

$$f(J_1 + J_2) = f(J_1)\, f(J_2).$$

This condition is not a biological assumption, but a requirement of structural coherence between the additivity of information and the multiplicative composition of persistence weights. Together with the boundary condition $f(0) = 1$, expressing informational neutrality in the absence of coupling, and with minimal regularity assumptions excluding pathological solutions, this functional equation strongly constrains the admissible form of f.

Defining $\phi(J) = \log\ f(J)$, the condition reduces to Cauchy’s additive functional equation

$$\phi(J_1 + J_2) = \phi(J_1) + \phi(J_2), \phi(0) = 0.$$

Under standard regularity conditions (such as continuity or measurability), the unique solution is linear,

$$\phi(J) = \beta J,$$

for some real constant β. Consequently, the informational persistence component must take the exponential form

$$f(J) = \exp(\beta J),$$

or, when information is measured in bits,

$$f(J) = 2^{\beta J}.$$

The parameter β quantifies the strength with which informational gain is translated into persistence weight. It is not a universal constant: its value depends on the selective domain, on operational and physical constraints, and on the specific selective regime under consideration. What is fixed by the present analysis is not the numerical value of β, but the functional dependence of informational persistence on informational gain. The present derivation is therefore structural: it establishes the admissible form of an informational persistence component once such a component is defined within the framework. It does not, from population reweighting alone, provide unique empirical identification of β or of the decomposition of total persistence into informational and residual components, which may require additional constraints or independent observations.

This derivation shows that the exponential form of w_{info} is not a modelling choice introduced for convenience, but a necessary consequence of the structural requirements imposed by the probabilistic description of selection and by the additive nature of informational quantities. It therefore provides a rigorous foundation for the form of the informational persistence component used in the main text, while remaining independent of any particular biological implementation. This result concerns the admissible functional form of the informational component and does not by itself establish unique empirical recoverability of β from observational population restructuring alone.