Symmetry, gauge freedoms, and the interpretability of sequence-function relationships

Anna Posfai[1], **David M. McCandlish**[1], **and Justin B. Kinney**[1,†]

[1]Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 11724

This manuscript was compiled on May 13, 2024

Quantitative models of sequence-function relationships, which describe how biological sequences encode functional activities, are ubiquitous in modern biology. One important aspect of these models is that they commonly exhibit gauge freedoms, i.e., directions in parameter space that do not affect model predictions. In physics, gauge freedoms arise when physical theories are formulated in ways that respect fundamental symmetries. However, the connections that gauge freedoms in models of sequence-function relationships have to the symmetries of sequence space have yet to be systematically studied. Here we study the gauge freedoms of models that respect a specific symmetry of sequence space: the group of position-specific character permutations. We find that gauge freedoms arise when the transformations of model parameters that compensate for these symmetry transformations are described by redundant irreducible matrix representations. Based on this finding, we describe an "embedding distillation" procedure that enables analytic calculation of the dimension of the space of gauge freedoms, as well as efficient computation of a sparse basis for this space. Finally, we show that the ability to interpret model parameters as quantifying allelic effects places strong constraints on the form that models can take, and in particular show that all nontrivial equivariant models of allelic effects must exhibit gauge freedoms. Our work thus advances the understanding of the relationship between symmetries and gauge freedoms in quantitative models of sequence-function relationships.

sequence-function relationships | gauge freedoms | sequence space | permutation symmetry | representation theory

Introduction

Understanding the quantitative nature of sequence-function relationships is a major goal of modern biology (1). To study a specific sequence-function relationship of interest, researchers often propose a mathematical model, fit the parameters of the model to data, then biologically interpret the resulting parameter values. This interpretation step is often complicated, however, by gauge freedoms—directions in parameter space along which model parameters can be changed without altering model predictions. If any gauge freedoms are present in a model, the numerical values of individual model parameters cannot be meaningfully interpreted in the absence of additional constraints.

Researchers performing quantitative studies of sequence-function relationships routinely encounter gauge freedoms in their models. In practice, one of two methods is typically used to overcome the difficulties that such gauge freedoms can present. One method—called "gauge fixing"—removes gauge freedoms by introducing additional constraints on model parameters (2–18). Another method limits the mathematical models that one uses to models that do not have any gauge freedoms (19–24). But despite being frequently encountered in the course of research, the gauge freedoms present in models of sequence-function relationships have received little attention (though see e.g. 3, 5–7, 12, 25). In particular, the mathematical properties of these gauge freedoms have yet to be systematically studied.

In physics, by contrast, gauge freedoms are a topic of fundamental importance (26). Gauge freedoms are well-known to arise when a physical theory is expressed in a form that manifestly respects fundamental symmetries. For example, the classical theory of electricity and magnetism (E&M) is invariant to Lorentz transformations, i.e., changes in an observer's velocity (27). Lorentz invariance is obscured, however, when the equations of E&M are expressed directly in terms of electric and magnetic fields. To express E&M in a form that is manifestly Lorentz invariant, one must instead formulate the equations in terms of an electromagnetic four-potential. Doing this introduces gauge freedoms because the four-potential, unlike electric and magnetic fields, is neither directly measurable nor uniquely determined by the configuration of a physical system*. Nevertheless, working with the four-potential greatly simplifies the equations of E&M and often aids in both their solution and their physical interpretation.

Motivated by the connection between gauge freedoms and symmetries in physics, we investigated whether the gauge freedoms in mathematical models of sequence-function relationships have a connection to the symmetries of sequence space. Here we study the gauge freedoms of linear models that are equivariant under a specific symmetry group of sequence space—the group of position-specific character permutations. These models include many of the most commonly used models, including models with pairwise and/or higher-order interactions. Using techniques from the theory of matrix representations, we find that the gauge freedoms of these models arise when model parameters transform under redundant irreducible matrix representations of the symmetry group. Based on this finding, we describe an "embedding distillation" procedure that facilitates the analysis of the vector space formed by the gauge freedoms of a large class of commonly used models.

Finally, we investigate the connection between parameter interpretability and model transformation behavior. We show that the ability to interpret the individual parameters of an equivariant model as quantifying the effects of specific alleles requires that these parameters transform under a permutation representation of the symmetry group, rather than a more general matrix representation. A consequence is that

*Results in quantum physics, such as the Aharanov-Bohm effect (28, 29), suggest a reality to the four-potential beyond what can be inferred solely from classical E&M, though there are arguments against this interpretation (30).

Please provide details of author contributions here.

Please declare any competing interests here.

†Correspondence: jkinney@cshl.edu (JBK)

all nontrivial models that satisfy this interpretation criterion have gauge freedoms. This shows in particular that models that have gauge freedoms can have important advantages over mathematically equivalent models that do not have gauge freedoms. A companion paper (31) reports specific gauge-fixing strategies that can be applied to the most commonly used models that can be interpreted as quantifying allelic effects.

Background

We now establish definitions and notation used in Results. We also review basic results regarding gauge freedoms in mathematical models of sequence-function relationships. Our companion paper (31) provides an expanded discussion of these results together with corresponding proofs.

Sequence-function relationships. Let $\mathcal{A}$ denote an alphabet comprising α distinct characters. Let $\mathcal{S}$ denote the set of α^L sequences of length L built from these characters. A model of a sequence-function relationship, $f(s;\vec{\theta})$, is defined to be a function that maps each sequence $s \in \mathcal{S}$ to a complex number. The vector $\vec{\theta}$ denotes the parameters of the model and is assumed to comprise M complex numbers.

Linear models. Linear models of sequence-function relationships are linear in $\vec{\theta}$ and thus have the form

$$f(s;\vec{\theta}) = \vec{\theta}^{\dagger}\vec{x}(s) = \sum_{i=1}^{M} \theta_i x_i(s), \qquad [1]$$

where $\vec{x}(\cdot)$ is a vector of M distinct sequence features, and each feature $x_i(\cdot)$ is a function that maps sequences in $\mathcal{S}$ to the complex numbers. We refer to the space $\mathbb{C}^M$ in which these feature vectors live as feature space, and each specific feature vector $\vec{x}(s)$ as the embedding of sequence s.

Note that we let both sequence embeddings $\vec{x}$ and model parameters $\vec{\theta}$ be complex. By contrast, ref. (31) limited embeddings and parameters to the reals. We choose here to work in complex spaces because, in addition to the added generality of the results, the algebraic completeness of the complex numbers simplifies some of our proofs. All of our results, however, hold if the parameters and embeddings are restricted to the reals. See SI Sec. 10 for details.

Generalized one-hot (GO) models. GO models are linear models in which the sequence features indicate only the presence or absence of specific characters at specific positions (1). An example of a GO is the pairwise-interaction model, which has the form

$$f_{\text{pair}}(s) = \theta_0 x_0(s) + \sum_{l}\sum_{c} \theta_l^c x_l^c(s) + \sum_{l<l'}\sum_{c,c'} \theta_{ll'}^{cc'} x_{ll'}^{cc'}(s), \quad [2]$$

where $l, l' \in \{1,\ldots,L\}$ index positions within sequences s and $c, c' \in \mathcal{A}$ index characters at these positions. Pairwise-interaction models comprise three types of GO feature: the constant feature, $x_0(s)$, which equals one for every sequence s; additive features, $x_l^c(s)$, which equal one if $s_l = c$ and equal zero otherwise (where s_l denotes the character at position l in sequence s); and pairwise features, $x_{ll'}^{cc'}(s)$, which equal one if both $s_l = c$ and $s_{l'} = c'$, and which equal zero otherwise.

GO models are defined in a similar manner: as sums of terms that each have the form

$$\theta_{l_1 l_2 \ldots l_K}^{c_1 c_2 \ldots c_K} x_{l_1 l_2 \ldots l_K}^{c_1 c_2 \ldots c_K}(s). \qquad [3]$$

Here, $K \in \{0,\ldots,L\}$ is a term-specific number, $\{l_1, l_2, \ldots, l_K\}$ is a term-specific set of positions, and $\{c_1, c_2, \ldots, c_K\}$ is a term-specific set of characters at the corresponding positions. Each feature $x_{l_1 l_2 \ldots l_K}^{c_1 c_2 \ldots c_K}(s)$ is a K-order one-hot feature defined to be equal to one if $s_k = c_k$ for all $k \in \{1,\ldots,K\}$ and equal to zero otherwise. For example, the pairwise-interaction model is a GO model that contains a $K = 0$ term[†] as well as all possible terms of order $K = 1$ and $K = 2$.

Gauge freedoms. Gauge freedoms are transformations of model parameters that do not affect model predictions. Formally, a gauge freedom is any vector $\vec{g} \in \mathbb{C}^M$ that satisfies

$$f(s;\vec{\theta}) = f(s;\vec{\theta}+\vec{g}) \quad \text{for all} \quad s \in \mathcal{S}. \qquad [4]$$

For linear sequence-function relationships, the set of gauge freedoms G is a vector space in $\mathbb{C}^M$. G is the orthogonal complement of the space spanned by sequence embeddings, which we denote by span$\vec{x}$ (31). In what follows, we use γ to represent the dimension of G, and often refer to this quantity somewhat informally as the number of gauge freedoms.

Gauge freedoms arise from linear dependencies among sequence features. For example, one-hot pairwise-interaction models have $M = 1 + \alpha L + \binom{L}{2}\alpha^2$ parameters, but span $\vec{x}$ has only $1 + (\alpha-1)L + \binom{L}{2}(\alpha-1)^2$ dimensions due the presence of $L + \binom{L}{2}(2\alpha-1)$ constraints on the embedding. Specifically, $x_0(s) = \sum_{c'} x_l^{c'}(s)$ for all positions l (yielding 1 constraint per position), and both $x_l^c(s) = \sum_{c'} x_{ll'}^{cc'}(s)$ and $x_{l'}^c(s) = \sum_{c'} x_{ll'}^{c'c}(s)$ for all characters c and for all pairs of positions $l < l'$ [yielding $2\alpha - 1$ independent constraints per pair of positions (31)]. The one-hot pairwise interaction model therefore has $\gamma = L + \binom{L}{2}(2\alpha-1)$ gauge freedoms; See also (3, 5, 7, 10).

Fixing the gauge. Fixing the gauge is the process of removing gauge freedoms by restricting $\vec{\theta}$ to a subset of parameter space, Θ, called the gauge. Linear gauges are choices of Θ that are themselves vector spaces. One useful property of linear gauges is that gauge-fixing can be accomplished by projection. Specifically, for any linear gauge Θ, there exists a projection matrix P that projects each parameter vector $\vec{\theta} \in \mathbb{C}^M$ onto an equivalent parameter vector $\vec{\theta}_{\text{fixed}}$ that lies in Θ, i.e.,

$$\vec{\theta}_{\text{fixed}} = P\,\vec{\theta}. \qquad [5]$$

Given Θ, the projection matrix P is uniquely defined by the requirement that P is idempotent, the image P is Θ, and the kernel of P is G. Our companion paper (31) describes a parametric family of linear gauges (including an explicit formula for the projection matrix) that includes many of the most commonly used gauges as special cases.

Results

In what follows, we define the group of position-specific character permutations, as well as the linear models that are equivariant under this group. Next, we use methods from the

[†] Here and in what follows, $K = 0$ corresponds to feature x_0 and parameter θ_0.

theory of group representations (32) to identify all possible equivariant linear models. In the process, we also describe a procedure we call "embedding distillation" that allows one to compute the gauge freedoms of any equivariant linear model. After demonstrating embedding distillation on the one-hot pairwise-interaction model, we apply embedding distillation to GEO models and derive expressions for the number of gauge freedoms (i.e., the dimension of the space of gage freedoms) of a variety of commonly used models. Finally, we explore the relationship between model transformation behavior and parameter identifiability.

Position-specific character permutations. Let H^l_{CP} denote the group of permutations among the α possible characters at position l in a sequence. Note that H^l_{CP} is isomorphic to the symmetric group on α elements, S_α (32). The group of position-specific character permutations is given by the direct product of all H^l_{CP}, i.e.,

$$H_{\mathrm{PSCP}} = H^1_{\mathrm{CP}} \times \cdots \times H^L_{\mathrm{CP}}. \qquad [6]$$

Given any $h \in H_{\mathrm{PSCP}}$, the transformation of a sequence s by h is written as hs, and the transformation of sequence space $\mathcal{S}$ by h is written as $h\mathcal{S}$.

Equivariant embeddings and equivariant models. A representation R of a group H is a function that maps each $h \in H$ to a complex matrix $R(h)$ such that $R(h_1 h_2) = R(h_1)R(h_2)$ for all $h_1, h_2 \in H$. In what follows, we say that an embedding $\vec{x}$ is "equivariant" if and only if there is a representation R of H_{PSCP} such that

$$\vec{x}(hs) = R(h)\,\vec{x}(s) \qquad [7]$$

for all $h \in H_{\mathrm{PSCP}}$ and all $s \in \mathcal{S}$ (Fig. 1A). We also say that a linear model is equivariant if and only if it is defined with an equivariant embedding $\vec{x}$ in Eq. 1. For any equivariant model, the transformation of sequence space by any $h \in H_{\mathrm{PSCP}}$ can be compensated for by a corresponding transformation of model parameters. Specifically, the sequence-space transformation $\mathcal{S} \to h\mathcal{S}$, $h \in H_{\mathrm{PSCP}}$, is compensated for by the parameter transformation $\vec{\theta} \to R(h)^{-1\dagger}\vec{\theta}$, in the sense that $f\left(s; \vec{\theta}\right) = f\left(hs; R(h)^{-1\dagger}\vec{\theta}\right)$ for every $s \in \mathcal{S}$ and $\vec{\theta} \in \mathbb{C}^M$ (see SI Sec. 3.2). Using terminology from representation theory, every $R(h)$ is an $M \times M$ matrix where M is called the degree of R (denoted $\deg R$). Similarly $\vec{x}(s)$ is an M-dimensional vector, where m is called the degree of $\vec{x}$ (denoted $\deg \vec{x}$).

Maschke decomposition of equivariant embeddings. Every group representation is either reducible or irreducible. A representation is irreducible if and only if it has no proper invariant subspace. Maschke's theorem, a basic result in representation theory, says that all representations of finite groups are equivalent (i.e., equal up to a similarity transformation) to a direct sum of irreducible representations. Any representation R of H_{PSCP} can therefore be expressed as

$$R \simeq \bigoplus_{k=1}^{K} Q_k R_k, \qquad [8]$$

where $\simeq$ denotes equivalence, each R_k is an irreducible representation of H_{PSCP}, all R_k are pairwise inequivalent, and Q_k denotes the multiplicity of R_k in the direct sum.

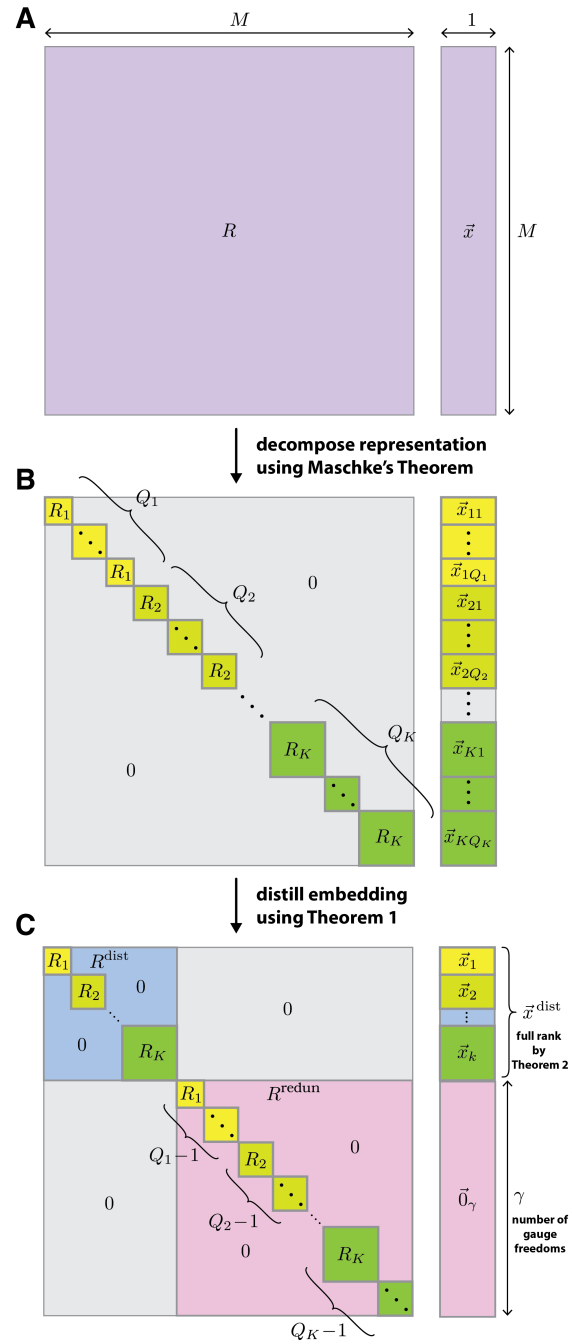

Fig. 1. Embedding distillation. (A) Given an M-dimensional embedding $\vec{x}$ that is equivariant under H_{PSCP}, let R be the representation of H_{PSCP} that acts on $\vec{x}$. (B) By Maschke's theorem, R can be decomposed into a direct sum of irreducible representations, R_k ($k \in \{1, \ldots, K\}$), each of which occurs with multiplicity Q_k (Eq. 8). Similarly, $\vec{x}$ can be decomposed into a direct sum of irreducible embeddings $\vec{x}_{kq}$ ($q \in \{1, \ldots, Q_k\}$), where each $\vec{x}_{kq}$ transforms under R_k (Eq. 9). (C) By Theorem 1, an additional similarity transformation can be performed that, for each value of k, zeroes out all but one $\vec{x}_{kq}$; the remaining $\vec{x}_{kq}$ is denoted by $\vec{x}_k$ (Eq. 11 and Eq. 12). Consequently, $\vec{x}$ decomposes into a direct sum of a distilled embedding, $\vec{x}^{\,\mathrm{dist}}$, and a zero vector, $\vec{0}_\gamma$, having dimension γ (Eq. 11). $\vec{x}^{\,\mathrm{dist}}$ is given by the direct sum of all $\vec{x}_k$ (Eq. 12) and is full rank by Theorem 2. The distilled representation, $R^{\,\mathrm{dist}}$, describes how $\vec{x}^{\,\mathrm{dist}}$ transforms and contains one copy of each R_k. The redundant representation, $R^{\,\mathrm{redun}}$, operates on $\vec{0}_\gamma$ and encapsulates the $Q_k - 1$ redundant copies of each R_k. γ, the degree of $R^{\,\mathrm{redun}}$, is equal to the number of gauge freedoms (Eq. 15).

In what follows, we say that a sequence embedding is irreducible if and only if it transforms under an irreducible representation of H_{PSCP}. One consequence of Eq. 8 is that any embedding $\vec{x}$ that transforms under R can be decomposed as

$$\vec{x} \simeq \bigoplus_{k=1}^{K} \bigoplus_{q=1}^{Q_k} \vec{x}_{kq}, \qquad [9]$$

where each $\vec{x}_{kq}$ is an irreducible embedding that transforms under R_k. This decomposition is illustrated in Fig. 1B. We assume in what follows that all $\vec{x}_{kq}$ are nonzero, but this assumption can be removed without fundamentally changing our results.[‡]

Distillation of equivariant embeddings. We now describe how equivariant models are analyzed via the distillation of their embeddings. In SI Sec. 5.1, we prove the following:

Theorem 1 *Any two nonzero sequence embeddings that transform under the same irreducible representation of H_{PSCP} are equal up to a constant of proportionality.*

Using Theorem 1, then performing additional similarity transformations to remove the constants of proportionality, we obtain,

$$\vec{x} \simeq \bigoplus_{k=1}^{K} Q_k \vec{x}_k, \qquad [10]$$

where $\vec{x}_k$ is any one of the $\vec{x}_{kq}$, and Q_k is the multiplicity of $\vec{x}_k$ in the direct sum. Additional similarity transformations can then be performed to zero out all except one copy of $\vec{x}_k$. We therefore find that there is an invertible "distillation matrix" T such that

$$T\vec{x} = \vec{x}^{\,\mathrm{dist}} \oplus \vec{0}_\gamma, \qquad [11]$$

where $\vec{0}_\gamma$ is a γ-dimensional vector of zeros, and

$$\vec{x}^{\,\mathrm{dist}} = \bigoplus_{k=1}^{K} \vec{x}_k, \qquad [12]$$

is the distilled embedding. Similarly, the matrix representation R decomposes as

$$TRT^{-1} = \vec{R}^{\,\mathrm{dist}} \oplus \vec{R}^{\,\mathrm{redun}} \qquad [13]$$

where the distilled representation, $R^{\mathrm{dist}} = \bigoplus_{k=1}^{K} R_k$, contains one copy of each R_k, and the redundant representation, $R^{\mathrm{redun}} = \bigoplus_{k=1}^{K} (Q_k - 1) R_k$, contains all of the other copies of each R_k that are present in R. These decompositions are illustrated in Fig. 1C.

Identification of gauge freedoms in equivariant models. To identify the gauge freedoms of any equivariant model, we use the fact that $\vec{x}^{\,\mathrm{dist}}$ is full rank. This is a consequence of the following Theorem, which is proven in SI Sec. 3.4:

Theorem 2 *For each $k \in \{1, \ldots, K\}$, let $\vec{x}_k$ be a nonzero embedding that transforms under an irreducible representation R_k of the group H_{PSCP}. Then the direct sum of all $\vec{x}_k$ is full rank if all R_k are pairwise inequivalent.*

[‡] See SI Sec. 5.2 for a statement of our main results when this assumption is removed.

Because $\vec{x}^{\,\mathrm{dist}}$ is full rank, $\vec{g}^{\,\dagger}\vec{x}(s) = 0$ for all $s \in \mathcal{S}$ if and only if

$$\vec{g} = T^{\dagger} \left[\vec{0}_{M-\gamma} \oplus \vec{g}_\gamma \right], \qquad [14]$$

for some γ-dimensional vector $\vec{g}_\gamma$. The space of gauge transformations, G, is therefore given by the set of vectors having the form in Eq. 14. In particular, the number of gauge freedoms is,

$$\gamma = \deg \vec{x} - \deg \vec{x}^{\,\mathrm{dist}} = \deg R^{\mathrm{redun}}. \qquad [15]$$

We thus see that the number of gauge freedoms is equal to the sum of the degrees of all redundant irreducible representations in R.

From Eq. 14, we also see that G is spanned by the last γ column vectors of $T^{\dagger}$. One can therefore compute a basis for G simply by computing T, and computing T only requires keeping track of the similarity transformations needed to express $\vec{x}$ in the distilled form shown in Eq. 11.

Identification of all equivariant embeddings. The specific structure of H_{PSCP} allows us to identify all possible inequivalent irreducible equivariant embeddings, $\vec{x}_k$. Because $\vec{x}_k$ is irreducible and H_{PSCP} is a product group, $\vec{x}_k$ can be expressed as

$$\vec{x}_k \simeq \bigotimes_{l=1}^{L} \vec{x}_l^{\,k}, \qquad [16]$$

where each $\vec{x}_l^{\,k}$ is an embedding that depends only on the character at position l and that transforms under an irreducible representation of H_{CP}^{l}. Moreover, because H_{CP}^{l} is isomorphic to S_α and S_α supports only two inequivalent embeddings (see SI Sec. 4.3 for proof), there are only two inequivalent choices for each $\vec{x}_l^{\,k}$: the trivial embedding and the simplex embedding. The trivial embedding, denoted $\vec{x}^{\,\mathrm{triv}}$, maps every sequence to a one-dimensional vector and transforms under what is called the "trivial representation" of S_α. The simplex embedding, denoted $\vec{x}_l^{\,\mathrm{sim}}$, maps sequences to the α vertices of an $\alpha - 1$ dimensional simplex and transforms under what is called the "standard representation" of S_α. One example of the simplex embedding is the tetrahedral embedding of DNA and RNA (20, 22). Note: to lessen the notational burden in what follows, we avoid writing $\vec{x}^{\,\mathrm{triv}}$ within tensor products over positions l, and only show factors that contribute nontrivially to these products.

We now identify all equivariant embeddings $\vec{x}$. Because there are 2 inequivalent choices for each $\vec{x}_l^{\,k}$ ($\vec{x}^{\,\mathrm{triv}}$ or $\vec{x}_l^{\,\mathrm{sim}}$), there are 2^L inequivalent choices for $\vec{x}_k$, and thus $\binom{2^L}{K}$ inequivalent choices for the set $\{\vec{x}_k\}_{k=1}^{K}$. Letting K in Eq. 12 range from 0 to 2^L, we find that there are $\sum_{K=0}^{2^L} \binom{2^L}{K} = 2^{2^L}$ inequivalent choices for $\vec{x}^{\,\mathrm{dist}}$. Every equivariant embedding $\vec{x}$ can therefore be expressed, using one of these 2^{2^L} inequivalent distilled embeddings $\vec{x}^{\,\mathrm{dist}}$ together with a zero vector $\vec{0}_\gamma$ and an invertible matrix T. Conversely, choosing any of the 2^{2^L} inequivalent distilled embeddings $\vec{x}^{\,\mathrm{dist}}$, any non-negative integer γ, and any invertible matrix T of the appropriate size will yield an equivariant embedding $\vec{x}$ via Eq. 11. We therefore find that, modulo the choice of the similarity matrix T and number of gauge freedoms γ, there are 2^{2^L} distinct choices for $\vec{x}$.

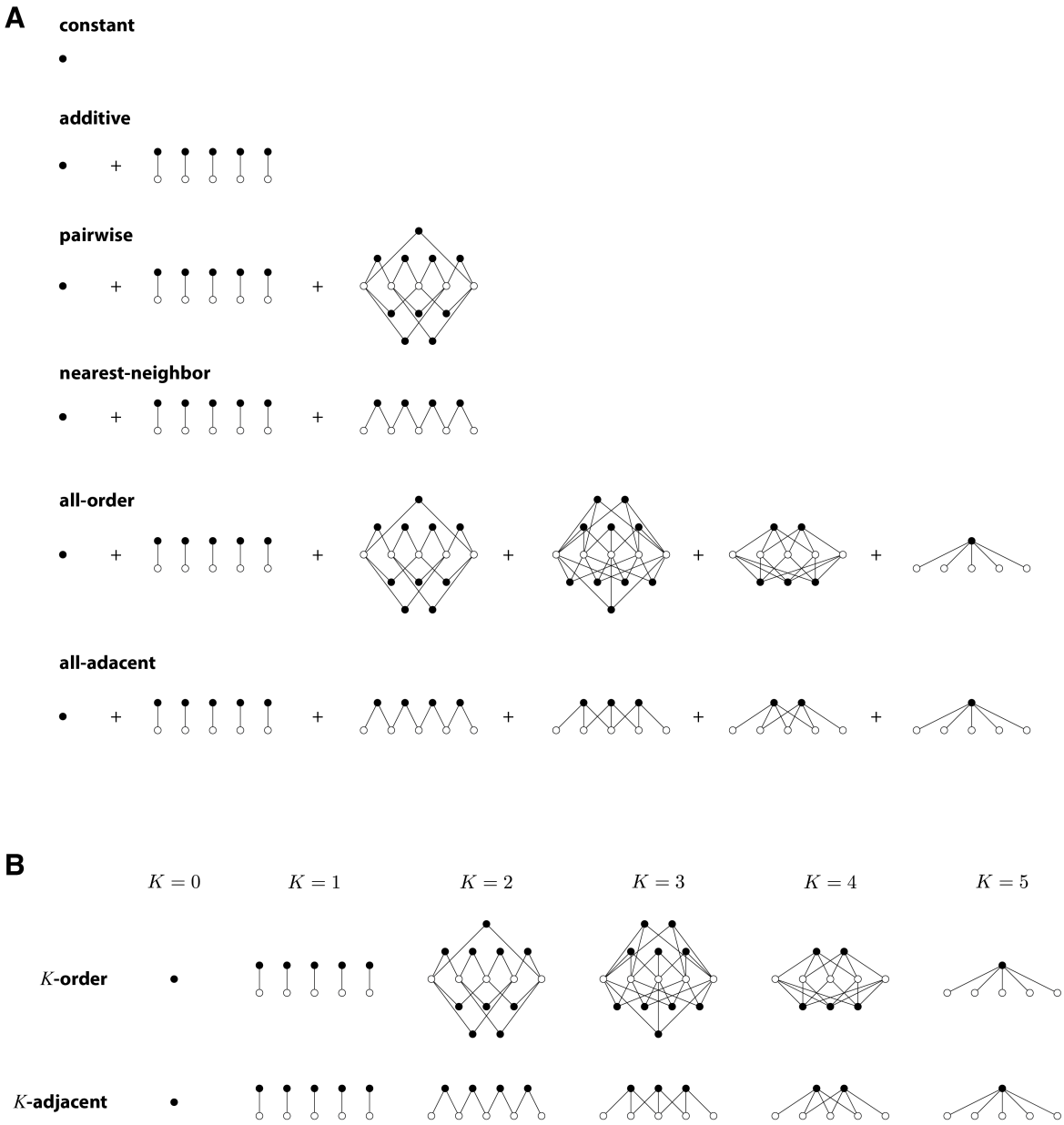

Fig. 2. Structure of GEO models. (A,B) Models analyzed in Table 1, illustrated for $L = 5$. Open circles represent sequence positions. Closed circles represent sets of parameters that are closed under the action of H_{PSCP}, as in Eq. 23. Edges indicate position indices shared by all the parameters within each closed set. (A) Structure of specific models of interest. (B) Structure of K-order models and K-adjacent models for various interaction orders K.

Analytical analysis of pairwise-interaction models. We now demonstrate the embedding distillation procedure on the pairwise-interaction model. First we specify the pairwise-interaction embedding, $\vec{x}_{\,\mathrm{pair}}$, as a direct sum of direct products of simpler embeddings:

$$\vec{x}_{\,\mathrm{pair}} = \vec{x}^{\,\mathrm{triv}} \oplus \left\{ \bigoplus_{l} \vec{x}_l^{\,\mathrm{ohe}} \right\} \oplus \left\{ \bigoplus_{l<l'} \vec{x}_l^{\,\mathrm{ohe}} \otimes \vec{x}_{l'}^{\,\mathrm{ohe}} \right\}, \qquad [17]$$

where $\vec{x}_l^{\,\mathrm{ohe}}$ is a position-specific one-hot embedding of dimension α given by

$$\vec{x}_l^{\,\mathrm{ohe}}(s) = \begin{bmatrix} x_l^{c_1}(s) \\ \vdots \\ x_l^{c_\alpha}(s) \end{bmatrix} \qquad [18]$$

for all sequences s, where $c_1, \ldots, c_\alpha$ denote the elements of $\mathcal{A}$. The number of model parameters is equal to the dimension of $\vec{x}_{\,\mathrm{pair}}$, which is seen from Eq. 17 to be $\deg \vec{x}_{\,\mathrm{pair}} = 1 + L\alpha + \binom{L}{2}\alpha^2$.

The gauge freedoms of pairwise-interaction models arise because $\vec{x}_{\,\mathrm{pair}}$ is not full rank. The reduced rank of $\vec{x}_{\,\mathrm{pair}}$ is a consequence of the fact that $\vec{x}_l^{\,\mathrm{ohe}}$ is reducible. To derive a

model type	interaction orders	no. parameters (M)	no. gauge freedoms (γ)
constant	0	1	0
additive	$0, 1$	$1 + L\alpha$	L
pairwise	$0, 1, 2$	$1 + L\alpha + \binom{L}{2}\alpha^2$	$L + \binom{L}{2}(2\alpha - 1)$
nearest-neighbor	$0, 1, 2$	$1 + L\alpha + (L-1)\alpha^2$	$L + (L-1)(2\alpha - 1)$
all-order	$0, 1, \ldots, L$	$(\alpha+1)^L$	$(\alpha+1)^L - \alpha^L$
all-adjacent	$0, 1, \ldots, L$	$1 + \frac{\alpha}{(\alpha-1)^2}\left[\alpha^{L+1} - (L+1)\alpha + L\right]$	$1 + \frac{\alpha}{(\alpha-1)^2}\left[2\alpha^L - \alpha^{L-1} - (L+1)\alpha + L\right]$
K-order	K	$\binom{L}{K}\alpha^K$	$\binom{L}{K}\alpha^K - \sum_{k=0}^{K}\binom{L}{k}(\alpha-1)^k$
hierarchical K-order	$0, 1, \ldots, K$	$\sum_{k=0}^{K}\binom{L}{k}\alpha^k$	$\sum_{n=0}^{K}\binom{L}{k}\left[\alpha^k - (\alpha-1)^k\right]$
K-adjacent†	K	$(L-K+1)\alpha^K$	$(L-K)\alpha^{K-1}$
hierarchical K-adjacent†	$0, 1, \ldots, K$	$1 + \sum_{k=1}^{K}(L-k+1)\alpha^k$	$(L-K)\alpha^{K-1} + 1 + \sum_{k=1}^{K-1}(L-k+1)\alpha^k$

Table 1. Number of parameters and gauge freedoms of various GEO models. Columns show model type, the orders of interaction included in each model, the number of parameters of each model, and the number of gauge freedoms of each model. See SI Sec. 6 for derivations of these results. GEO, generalized equivariant one-hot. †Assumes $K \geq 1$.

distilled version of $\vec{x}_{\text{pair}}$ that is full rank, we reexpress $\vec{x}_l^{\text{ohe}}$ as a direct sum of irreducible embeddings using

$$\vec{x}_l^{\text{ohe}} \simeq \vec{x}^{\text{triv}} \oplus \vec{x}_l^{\text{sim}}; \qquad [19]$$

see SI Sec. 2.4 for details. Plugging Eq. 19 into Eq. 17, expanding the direct product, and grouping like terms, we get

$$\vec{x}_{\text{pair}} \simeq \left[1 + L + \binom{L}{2}\right]\vec{x}^{\text{triv}} \oplus \left\{\bigoplus_l L\,\vec{x}_l^{\text{sim}}\right\} \oplus \left\{\bigoplus_{l<l'} \vec{x}_l^{\text{sim}} \otimes \vec{x}_{l'}^{\text{sim}}\right\}, \qquad [20]$$

where the scalar coefficients denote the multiplicity of each term in the direct sum. Because $\vec{x}^{\text{triv}}$, $\vec{x}_l^{\text{sim}}$, and $\vec{x}_l^{\text{sim}} \otimes \vec{x}_l^{\text{sim}}$ are irreducible and pairwise inequivalent, the distillation of $\vec{x}_{\text{pair}}$ is seen from Eq. 20 to be

$$\vec{x}_{\text{pair}}^{\text{dist}} = \vec{x}^{\text{triv}} \oplus \left\{\bigoplus_l \vec{x}_l^{\text{sim}}\right\} \oplus \left\{\bigoplus_{l<l'} \vec{x}_l^{\text{sim}} \otimes \vec{x}_{l'}^{\text{sim}}\right\}. \qquad [21]$$

From this we observe that $\deg \vec{x}_{\text{pair}}^{\text{dist}} = 1 + L(\alpha-1) + \binom{L}{2}(\alpha-1)^2$. The number of gauge freedoms then follows from Eq. 15:

$$\gamma = L + \binom{L}{2}(2\alpha - 1), \qquad [22]$$

which matches the well-known result (3).

Generalized equivariant one-hot (GEO) models. For a GO model to be equivariant, it is sufficient for the model to be expressible as a sum of equivariant terms, each term of the form

$$\sum_{c_1 \in \mathcal{A}} \cdots \sum_{c_K \in \mathcal{A}} \theta_{l_1 l_2 \ldots l_K}^{c_1 c_2 \ldots c_K} x_{l_1 l_2 \ldots l_K}^{c_1 c_2 \ldots c_K}(s), \qquad [23]$$

for some term-specific choice of K and term-specific set of positions $\{l_1, \ldots, l_K\}$. Observe that GEO models differ from GO models in that, for every set of positions used to define a term, a GEO model sums over all possible characters at all positions in the set, whereas a GO model need not include terms for every possible choice of characters. An example of GO models that are not GEO models are those based on wild-type embeddings, i.e., embeddings that exclude features that involve character-positions pairs that occur in a chosen "wild-type" sequence.

The embeddings of GEO models all have the following form. Let A_j denote a set of sequence positions, and let $\{A_j\}_{j=1}^J$ denote the sets of positions used to construct an GEO model with sequence embedding $\vec{x}$. By analogy to Eq. 17, $\vec{x}$ can then be written as

$$\vec{x} = \bigoplus_{j=1}^{J} \bigotimes_{l \in A_j} \vec{x}_l^{\text{ohe}}. \qquad [24]$$

Because each direct product in Eq. 24 yields an embedding of dimension $\alpha^{|A_j|}$, the full dimension of $\vec{x}$ (and thus the number of model parameters) is

$$\deg \vec{x} = \sum_{j=1}^{J} \alpha^{|A_j|}. \qquad [25]$$

Analytical analysis of GEO models. Now we derive the corresponding distilled embedding. Using Eq. 19 to decompose each $\vec{x}_l^{\text{ohe}}$ in terms of $\vec{x}^{\text{triv}}$ and $\vec{x}_l^{\text{sim}}$, then expanding each tensor product and grouping the resulting terms, we find that $\vec{x}$ is given by Eq. 10 where

$$\vec{x}_k = \bigotimes_{l \in B_k} \vec{x}_l^{\text{sim}}, \qquad [26]$$

where each B_k ($k \in \{1, \ldots, K\}$) denotes a subset of positions that occurs among at least one of the A_j, and Q_k denotes the number of A_j in which B_k occurs.[§] By inspection we see that each $\vec{x}_k$ in Eq. 26 has dimension $(\alpha-1)^{|B_k|}$. Therefore, the dimension of $\vec{x}$ can alternatively be written as

$$\deg \vec{x} = \sum_{k=1}^{K} Q_k (\alpha - 1)^{|B_k|}. \qquad [27]$$

Every $\vec{x}_k$ is irreducible because every $\vec{x}_l^{\text{sim}}$ is irreducible. Consequently, the distilled embedding $\vec{x}^{\text{dist}}$ is given by Eq. 12 and has dimension

$$\deg \vec{x}^{\text{dist}} = \sum_{k=1}^{K} (\alpha - 1)^{|B_k|}. \qquad [28]$$

[§]Formally, $\{B_k\}_{k=1}^K = \bigcup_{j=1}^J P(A_j)$ where $P(\cdot)$ denotes the powerset, and $Q_k = \sum_{j=1}^J 1_{P(A_j)}(B_k)$ where $1_{P(A)}(\cdot)$ is the indicator function for membership in $P(A)$.

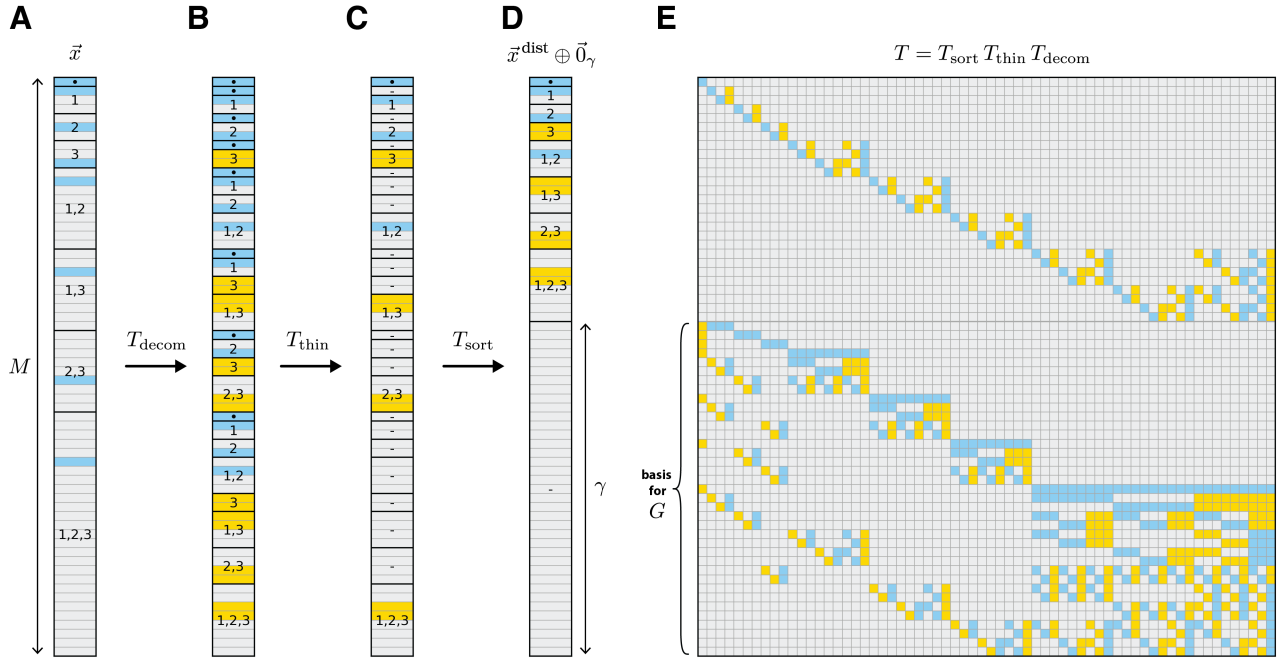

Fig. 3. Illustrated distillation computation. (A) Embedding $\vec{x}$ of sequence $s = \mathrm{ABC}$ for an all-order interaction model based on the alphabet $\mathcal{A} = \{\mathrm{A},\ \mathrm{B},\ \mathrm{C}\}$. Embedding has degree $M = 64$. (B) Result of multiplication by the decomposition matrix, T_{decom}. (C) Result of subsequent multiplication by the thinning matrix T_{thin}. (D) Result of subsequent multiplication by the sorting matrix T_{sort}, which yields $\vec{x}^{\,\mathrm{dist}} \oplus \vec{0}_\gamma$ with $\gamma = 37$ gauge freedoms. In B-D, dots indicate $\vec{x}^{\,\mathrm{triv}}$, dashes indicate zero vectors, and numbers indicate $\vec{x}_l^{\,\mathrm{sim}}$ or Kronecker products thereof for specified positions l. (E) Distillation matrix T that implements the full distillation procedure in A-D. Last γ rows of T provide a sparse basis for the gauge space, G. In A-E, vector and matrix elements are colored using: blue, +1; yellow, -1; gray, 0.

Using Eq. 15, the number of gauge freedoms of the embedding $\vec{x}$ is thus seen to be

$$\gamma = \sum_{k=1}^{K} (Q_k - 1)(\alpha - 1)^{|B_k|}. \qquad [29]$$

This result provides a way to analytically compute the number of gauge freedoms of any GEO model. Table 1 reports the number of gauge freedoms thus computed for a variety of such models. SI Sec. 6 provides expanded descriptions for each model, as well as detailed computations of the results in Table 1.

We note that the only GEO models that have no gauge freedoms are those that have embeddings built from only one tensor product in Eq. 24. To see this, observe from Eq. 29 that $\gamma = 0$ if and only if none of the Q_k are greater than 1. This requires that none of the B_k are subsets of two or more A_j. But the empty set, $\emptyset$, is a subset of every A_j, which means that $Q_k = J$ whenever $B_k = \emptyset$. Gauge freedoms will therefore be present unless $J = 1$, i.e. the direct sum in Eq. 24 includes only one term.

Computational analysis of GEO models. To derive a basis for the space of gauge freedoms, we must choose a specific realization of the irreducible embeddings $\vec{x}^{\,\mathrm{triv}}$ and $\vec{x}^{\,\mathrm{sim}}$. In what follows we choose $\vec{x}^{\,\mathrm{triv}}(s) = [1]$ and

$$\vec{x}_l^{\,\mathrm{sim}}(s) = \begin{cases} \begin{bmatrix} x_l^{c_1}(s) \\ \vdots \\ x_l^{c_{\alpha-1}}(s) \end{bmatrix} & \text{if } s_l \neq c_\alpha, \\ \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix} & \text{if } s_l = c_\alpha, \end{cases} \qquad [30]$$

for all sequences s, where $c_1, \ldots, c_\alpha$ represent an ordering of the characters in $\mathcal{A}$. With these choices in hand, Eq. 19 can be written as an equality:

$$T^{(1)} \vec{x}^{\,\mathrm{ohe}} = \vec{x}^{\,\mathrm{triv}} \oplus \vec{x}^{\,\mathrm{sim}}, \qquad [31]$$

where $T^{(1)}$ is an $\alpha \times \alpha$ matrix given by

$$T^{(1)} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & -1 \\ 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \end{bmatrix}. \qquad [32]$$

Using $T^{(1)}$ one can compute the distillation matrix for any GEO model as the product of three matrices:

$$T = T_{\mathrm{sort}}\, T_{\mathrm{thin}}\, T_{\mathrm{decom}}. \qquad [33]$$

The effects of these three matrices are illustrated in Fig. 3. The "decomposition matrix", T_{decom}, decomposes the one-hot

embedding $\vec{x}$ (Fig. 3A) into a direct sum of irreducible embeddings (Fig. 3B). The "thinning matrix", T_{thin}, then zeros out all except the first copy of each irreducible embedding (Fig. 3C). The "sorting matrix", T_{sort}, then rearranges the direct sum so that the remaining nonzero embeddings come first (Fig. 3D). SI Sec. 8 provides explicit algorithms for constructing T_{decom}, T_{thin}, and T_{sort}, as well as the inverse of each of these three matrices, for a large class of GEO models. Each of these six matrices has only $O(L)$ nonzero elements, and the algorithm for constructing each matrix has $O(L)$ computational complexity. The resulting distillation matrix T, as well as its inverse, are also sparse. Moreover, every nonzero element of T is +1 or -1 (Fig. 3E). Because the last γ columns of $T^\dagger$ provide a basis for G, we thus obtain a basis for the gauge space consisting of sparse vectors whose only nonzero elements are +1 and -1.

This result can also be used to efficiently fix the gauge of any GEO. Define the projection matrix

$$P = T^\dagger\big|_{M-\gamma}\, T^{-1\dagger}, \qquad [34]$$

where $|_{M-\gamma}$ denotes that the last γ columns of a matrix have been set to zero. P projects parameter vectors $\vec{\theta}$ onto the spaced spanned by the first $M-\gamma$ columns of $T^\dagger$. Moreover, by expanding P as

$$P = T^\dagger_{\text{decom}}\, T^\dagger_{\text{thin}}\, T^\dagger_{\text{sort}}\;\; I\big|_{M-\gamma}\, T^{-1\dagger}_{\text{sort}}\, T^{-1\dagger}_{\text{thin}}\, T^{-1\dagger}_{\text{decom}}, \qquad [35]$$

and applying each matrix factor to $\vec{\theta}$ individually, this projection can be performed in $O(L)$ time. Projection by P therefore provides and efficient way to remove gauge freedoms by projecting model parameters into a linear gauge. We note, however, that the resulting linear gauge is not one of the parametric gauges discussed in our companion paper (31).

Interpretability of pairwise-interaction models. The ability to interpret the parameters of equivariant models as quantifying allelic effects is closely related to how those parameters transform under H_{PSCP}. To illustrate this point, we consider two equivariant models: a pairwise-interaction GEO model with embedding $\vec{x}_{\text{pair}}$, and the corresponding distilled model with embedding $\vec{x}^{\text{dist}}_{\text{pair}}$, both operating on sequences built from a three-character alphabet, $\mathcal{A} = \{\text{A}, \text{B}, \text{C}\}$. The two embeddings encode the same set of interactions but do so in different ways: $\vec{x}_{\text{pair}}$ is built from the single-position one-hot encodings $\vec{x}^{\text{ohe}}_l$, whereas $\vec{x}^{\text{dist}}_{\text{pair}}$ is built from the single-position simplex encodings $\vec{x}^{\text{sim}}_l$ (Fig. 4A). And as we show above, the GEO model has gauge freedoms whereas the distilled model does not.

We now focus on how the features and parameters of these two models are affected by the transformation $h \in H_{\text{PSCP}}$ that exchanges the characters A and C at all positions l. For the GEO model, h induces a permutation of embedding coordinates (Fig. 4B) and thus of model parameters. Consequently, h preserves the set of values taken by the GEO parameters; it simply permutes which parameters have which values. This mirrors the action of h on the alleles that drive model predictions: h permutes sequences and thus the one-position and two-position alleles they contain, but does not alter the full set of alleles present among the full set of sequences. And in fact we see that individual parameter values track their corresponding alleles: θ^{A}_l and θ^{C}_l switch values, $\theta^{\text{AA}}_{ll'}$ and $\theta^{\text{CC}}_{ll'}$ switch values, etc.. The transformation behavior of the GEO model is therefore consistent with individual parameters quantifying the effects of individual alleles.

For the distilled model, however, h induces a non-permutation transformation of embedding coordinates (Fig. 4C) and thus of model parameters. Using the embedding shown in Fig. 4A, one finds that the value of θ^1_l transforms to $-\theta^1_l + \theta^2_l$, the value of $\theta^{11}_{ll'}$ transforms to $\theta^{11}_{ll'} - \theta^{12}_{ll'} - \theta^{21}_{ll'} + \theta^{22}_{ll'}$, etc.. The transformation h therefore changes the full set of values taken by the distilled model parameters. Consequently, the individual parameters of this model cannot be interpreted as quantifying the effects of individual alleles.

Nontrivial equivariant allelic models have gauge freedoms. To clarify the connection between the interpretation and transformation behavior of model parameters, we now formalize the notion of an allele, and allelic effect, and related concepts. We define an allele a to be a pattern of characters that is either present or absent in every sequence. The corresponding allelic set $\mathcal{S}_a$ is defined to be the set of sequences that have allele a, and the corresponding allelic feature x_a is defined be the indicator function for membership in $\mathcal{S}_a$. An allelic model is defined to be a linear sequence-function relationship in which every feature is an allelic feature. The effect of allele a is defined, in the context of a specific allelic model, to be the parameter θ_a that multiplies the allelic feature x_a.

Requiring an allelic model to be equivariant puts strong constraints on which alleles it can describe, and on how the corresponding allelic features and allelic effects transform. Given a specific allele a, the action of H_{PSCP} on a generates a set of alleles $\mathcal{O}$, which we call an allelic orbit. If the allelic model is equivariant, the allelic sets $\mathcal{S}_{a'}$ corresponding to all $a' \in \mathcal{O}$ will tile sequence space without overlaps. This requirement greatly constraints the set of possible alleles such a model can describe. Moreover, the model must include one feature $x_{a'}$ for every allele $a' \in \mathcal{O}$. These features will then transform among themselves according to a permutation representation. See SI Sec. 9 for details.

An equivariant allelic model must therefore contain features that can be partitioned into a set of complete allelic orbits. The features of the model will then transform under a direct sum of permutation representations, one for each allelic orbit. Because every permutation representation contains the trivial representation in its Maschke decomposition, the allelic model will have at least as many gauge freedoms as the number of allelic orbits minus one. Perhaps more intuitively, the sum of all allelic features corresponding to each orbit is equal to one for all sequences. Therefore, each orbit's features are sufficient to represent a constant function on sequence space. Including features from multiple orbits therefore overparameterizes the model and introduces gauge freedoms. We emphasize, however, that additional gauge freedoms can be present as well, so this result only provides a lower bound on γ.

It is readily seen that all GEO models are allelic models. In a GEO model, each allelic orbit corresponds to a position set A_j in Eq. 24, and the number of allelic orbits is given by J. Our lower-bound on the number of gauge freedoms recapitulates the finding above that only GEO models with $J = 1$ have no gauge freedoms. We also show in SI Sec. 9 that, given a model defined by a direct sum of direct products of single-position embeddings, the corresponding GEO model has the smallest number of gauge freedoms possible.

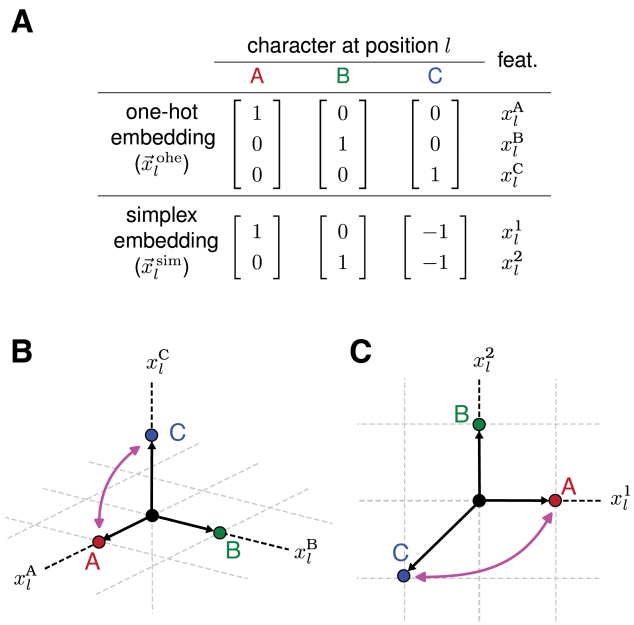

Fig. 4. Transformation behavior of two single-position embeddings. (A) Two single-position embeddings, $\vec{x}_l^{\mathrm{ohe}}$ and $\vec{x}_l^{\mathrm{sim}}$, for the three-character alphabet $\mathcal{A} = \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$. The specific features corresponding to each element of $\vec{x}_l^{\mathrm{ohe}}$ and $\vec{x}_l^{\mathrm{sim}}$ are also shown. (B) The three-dimensional one-hot embedding, $\vec{x}_l^{\mathrm{ohe}}(c)$, for each $c \in \mathcal{A}$. (C) The two-dimensional simplex embedding, $\vec{x}_l^{\mathrm{sim}}(c)$, for each $c \in \mathcal{A}$. Pink arrows indicate the transformation of each embedding vector induced by permuting characters A and C at position l.

We therefore see that there is an incompatibility between two distinct notions of parameter interpretability. In all except a limited class of models, the ability to interpret parameters as quantifying allelic effects is incompatible with the ability to interpret parameter values in the absence of gauge-fixing constraints. The only exceptions to this rule are single-orbit allelic models, but these models are trivial in the following sense:[¶] each sequence has only one allele, the effect of which is the model's prediction for the sequence. In a single-orbit allelic model, each sequence has only one allele–and thus one feature and one parameter–that contributes to its activity. The parameters are therefore essentially just a catalog of allelic effects. By contrast, the reason researchers quantitatively model sequence-function relationships in the first place is to deconvolve the influence of multiple co-occurring alleles. We conclude that, among nontrivial equivariant models (i.e., models that support co-occurring alleles), the ability to interpret model parameters as quantifying allelic effects requires that the model have gauge freedoms.

Discussion

Motivated by the connection between gauge freedoms and symmetries in physics, we investigated the relationship between gauge freedoms and symmetries in quantitative models of sequence-function relationships. We found that, for models that are equivariant under the group of position-specific character permutations (denoted H_{PSCP}), gauge freedoms arise due to model parameters transforming according to redundant irreducible matrix representations of H_{PSCP}. From a practical standpoint, this result facilitates the analytic calculation of the dimension of the space of gauge freedoms in a large class of commonly used models, as well as the efficient computation of a sparse basis for this space. From a conceptual standpoint, the results link the gauge freedoms of models of sequence-function relationships to the transformation behavior of these models under a specific symmetry group of sequence space.

We also investigated the link between parameter transformation behavior and parameter interpretability. In doing so, we identified a tension between two different notions of parameter interpretability: in all nontrivial equivariant models, the ability to interpret the values of model parameters in the absence of gauge-fixing constraints is incompatible with the ability to interpret parameters as quantifying allelic effects. Consequently, models that do have gauge freedoms (including nontrivial additive models, pairwise-interaction models, etc.) have important advantages over equally expressive models that do not have gauge freedoms.

We now return to the analogy with theoretical physics. In classical field theories like E&M, there are specific symmetries that are well-established by experiment and that any mathematical formulation of the theory must be consistent with. This does not, however, mean that the equations of the theory must transform in a simple way under those symmetries. Mathematically formulating physical theories so that the equations themselves manifestly respect the symmetries of the theory generally requires over-parameterizing the equations, thereby introducing gauge freedoms. Physicists often find it worthwhile to do this, as having fundamental symmetries be reflected in one's equations can greatly facilitate the interpretation and application of those equations. Solving such equations, however, requires fixing the gauge—introducing additional constraints that make the solution of the equations unique.

Unlike in physics, there is no experimentally established requirement that models of sequence-function relationships be equivariant under symmetries of sequence space. The specific mathematical form one uses for such models is subjective, and different models are commonly used in different contexts. Citing the ambiguities caused by gauge freedoms, some have argued for restricting one's choice of model to those that have no gauge freedoms. Nevertheless, models that have gauge freedoms remain dominant in the literature. We suggest that a major reason for this may be that researchers prefer to use models that both (a) reflect symmetries of sequence space and (b) have parameters that can be interpreted as allelic effects. As we showed, these criteria require the use of over-parameterized models. And in this way, the origin of gauge freedoms in models of sequence-function relationships does mirror the origin of gauge freedoms in physical theories.

There is still much to understand about the relationship between models of sequence-function relationships, the symmetries of these models, and how these modes can be biological interpreted. This paper and its companion (31) have only addressed gauge freedoms and symmetries in linear models of sequence-function relationships. Some work has explored the gauge freedoms and symmetries of nonlinear models of sequence-function relationships (33, 34), but important questions remain. The sloppy modes (35, 36) present in sequence-function relationships are also important to understand but, to our knowledge, these have yet to be systematically studied. Addressing these problems is becoming increasingly urgent,

[¶] This is the same sense in which the "trivial gauge" described in (31) is trivial.

not just because of the rapidly expanding use of quantitative models of sequence-function relationships, but also because of the emerging use of surrogate models for interpreting sequence-function relationships described by genomic deep neural networks (37).

Materials and Methods

See Supplemental Information for full derivations of the mathematical results. Python code implementing the embedding distillation algorithm described the section "Computational analysis of GEO models", as well as used for generating Fig. 3, is available at https://github.com/jbkinney/23_posfai.

ACKNOWLEDGMENTS. We thank Peter Koo and Vijay Balasubramanian for helpful discussions. This work was supported by NIH grant R35 GM133613 (AP, DMM), NIH grant R35 GM133777 (AP, JBK), NIH grant R01 HG011787 (JBK), the Alfred P. Sloan foundation (DMM), and additional funding from the Simons Center for Quantitative Biology at CSHL (DMM, JBK).

1. JB Kinney, DM McCandlish, Massively parallel assays and quantitative sequence-function relationships. *Annu. Rev. Genomics Hum. Genet.* **20**, 99–127 (2019) Wrote.
2. JB Kinney, G Tkacik, CG Callan, Precise physical models of protein-DNA interaction from high-throughput data. *Proc. Natl. Acad. Sci.* **104**, 501–506 (2007) Wrote.
3. M Weigt, RA White, H Szurmant, JA Hoch, T Hwa, Identification of direct residue contacts in protein-protein interaction by message passing. *Proc. Natl. Acad. Sci.* **106**, 67–72 (2009).
4. DS Marks, et al., Protein 3D Structure Computed from Evolutionary Sequence Variation. *PLoS ONE* **6**, e28766 (2011).
5. M Ekeberg, C L ovkvist, Y Lan, M Weigt, E Aurell, Improved contact prediction in proteins: Using pseudolikelihoods to infer Potts models. *Phys. Rev. E* **87**, 012707 (2013).
6. M Ekeberg, T Hartonen, E Aurell, Fast pseudolikelihood maximization for direct-coupling analysis of protein structure from many homologous amino-acid sequences. *J. Comput. Phys.* **276**, 341–356 (2014).
7. RR Stein, DS Marks, C Sander, Inferring Pairwise Interactions from Biological Data Using Maximum-Entropy Probability Models. *PLoS Comput. Biol.* **11**, e1004182 (2015).
8. JP Barton, ED Leonardis, A Coucke, S Cocco, ACE: adaptive cluster expansion for maximum entropy graphical model inference. *Bioinformatics* **32**, 3089–3097 (2016).
9. A Haldane, WF Flynn, P He, RM Levy, Coevolutionary Landscape of Kinase Family Proteins: Sequence Probabilities and Functional Motifs. *Biophys. J.* **114**, 21–31 (2018).
10. S Cocco, C Feinauer, M Figliuzzi, R Monasson, M Weigt, Inverse statistical physics of protein sequences: a key issues review. *Reports on Prog. Phys.* **81**, 032601 (2018).
11. A Haldane, RM Levy, Influence of multiple-sequence-alignment depth on Potts statistical models of protein covariation. *Phys. Rev. E* **99**, 032405 (2019).
12. HT Rube, et al., Probing molecular specificity with deep sequencing and biophysically interpretable machine learning. *bioRxiv* p. 2021.06.30.450414 (2021).
13. S Zamuner, PDL Rios, Interpretable Neural Networks based classifiers for categorical inputs. *arXiv* (2021).
14. C Feinauer, B Meynard-Piganeau, C Lucibello, Interpretable pairwise distillations for generative protein sequence models. *PLoS Comput. Biol.* **18**, e1010219 (2022).
15. A Gerardos, N Dietler, AF Bitbol, Correlations from structure and phylogeny combine constructively in the inference of protein partners from sequences. *PLoS Comput. Biol.* **18**, e1010147 (2022).
16. C Hsu, H Nisonoff, C Fannjiang, J Listgarten, Learning protein fitness models from evolutionary and assay-labeled data. *Nat. Biotechnol.* **40**, 1114–1122 (2022).
17. C Feinauer, E Borgonovo, Mean Dimension of Generative Models for Protein Sequences. *bioRxiv* p. 2022.12.12.520028 (2022).
18. HT Rube, et al., Prediction of protein-ligand binding affinity from sequencing data with interpretable machine learning. *Nat. Biotechnol.* **40**, 1520–1527 (2022).
19. ED Weinberger, Fourier and taylor series on fitness landscapes. *Biol. cybernetics* **65**, 321–330 (1991).
20. CT Zhang, R Zhang, Analysis of distribution of bases in the coding sequences by a diagrammatic technique. *Nucleic acids research* **19**, 6313–7 (1991).
21. PF Stadler, Spectral landscape theory in *Evolutionary Dynamics: Exploring the Interplay of Selection, Accident, Neutrality and Function*, eds. J Crutchfield, P Schuster. (Oxford Univ. Press, Oxford), pp. 231 – 271 (2003).
22. GD Stormo, Maximally efficient modeling of DNA sequence motifs at all levels of complexity. *Genetics* **187**, 1219 – 1224 (2011-04).
23. FJ Poelwijk, V Krishna, R Ranganathan, The Context-Dependence of Mutations: A Linkage of Formalisms. *PLOS Comput. Biol.* **12**, e1004771 (2016).
24. DH Brookes, A Aghazadeh, J Listgarten, On the sparsity of fitness functions and implications for learning. *Proc. Natl. Acad. Sci.* **119**, e2109649118 (2022).
25. A Tareen, et al., MAVE-NN: learning genotype-phenotype maps from multiplex assays of variant effect. *Genome Biol.* **23**, 98 (2022).
26. JD Jackson, LB Okun, Historical roots of gauge invariance. *Rev. Mod. Phys.* **73**, 663–680 (2001).
27. JD Jackson, *Classical electrodynamics*. (John Wiley & Sons), (1998).
28. Y Aharonov, D Bohm, Significance of electromagnetic potentials in the quantum theory. *Phys. review* **115**, 485 (1959).
29. M Peshkin, A Tonomura, *The Aharonov-Bohm Effect*. (Springer Verlag), (2005).
30. L Vaidman, Role of potentials in the aharonov-bohm effect. *Phys. Rev. A* **86**, 040101 (2012).
31. A Posfai, J Zhou, DM McCandlish, JB Kinney, Gauge fixing for sequence-function relationships. *In prep.* (2024).
32. BE Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*, Graduate Texts in Mathematics. (Springer), 2 edition, (2001) Read in early 2022.
33. JB Kinney, GS Atwal, Parametric Inference in the Large Data Limit Using Maximally Informative Models. *Neural computation* **26**, 637–653 (2014-04) Wrote.
34. GS Atwal, JB Kinney, Learning Quantitative Sequence–Function Relationships from Massively Parallel Experiments. *J. Stat. Phys.* **162**, 1203–1243 (2016) Wrote.
35. BB Machta, R Chachra, MK Transtrum, JP Sethna, Parameter space compression underlies emergent theories and predictive models. *Science* **342**, 604 – 607 (2013).
36. MK Transtrum, et al., Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *The J. Chem. Phys.* **143**, 010901 – 14 (2015).
37. E Seitz, DM McCandlish, JB Kinney, PK Koo, Interpreting cis-regulatory mechanisms from genomic deep neural networks using surrogate models. *bioRxiv* (2023).