Proofreading Is Too Noisy For Effective Ligand Discrimination

Duncan Kirby\textsuperscript{a} and Anton Zilman\textsuperscript{a,b}

\textsuperscript{a}Department of Physics, University of Toronto, Toronto, Ontario, M5S 1A7, Canada; \textsuperscript{b}Institute for Bioengineering, University of Toronto, Toronto, Ontario, M5S 1A7, Canada

This manuscript was compiled on January 13, 2023

Kinetic proofreading (KPR) has been used as a paradigmatic explanation for the high specificity of important biological processes including ligand discrimination by cellular receptors. Kinetic proofreading enhances the difference in the mean receptor occupancy between different ligands, thus potentially enabling better discrimination. On the other hand, proofreading also attenuates the signal, increasing the relative magnitude of noise in the downstream signal. This can interfere with reliable ligand discrimination. To understand the effect of noise on ligand discrimination beyond the comparison of the mean signals, we formulate the task of ligand discrimination as a problem of statistical estimation of the molecular affinity of ligands. Our analysis reveals that proofreading typically worsens ligand resolution which decreases with the number of proofreading steps under most commonly considered conditions. This contrasts with the usual notion that kinetic proofreading universally improves ligand discrimination with additional proofreading steps. Our results are consistent across a variety of different proofreading schemes, suggesting that they are inherent to the KPR mechanism itself rather than any particular model of molecular noise. Based on our results, we suggest alternative roles for kinetic proofreading schemes such as multiplexing and combinatorial encoding in multi-ligand/multi-output pathways.

1. Introduction

Cells send and receive signals in the form of molecular ligands that bind corresponding cell surface receptors. Ligand-receptor binding activates intracellular cascades of molecular modifications that eventually lead to cellular response. The ability of the cell to respond distinctly to different ligands depends on the ability of the receptor to discriminate between different ligands. In the simplest — and very common — scenario, discrimination relies on the differences in the receptor binding affinities between different ligands: cognate ligands bind the receptor more strongly (and for a longer time) than non-cognate ligands. At equilibrium, the ratio of the receptor occupancies by different ligands present at the same concentration is expected to be proportional to $e^{\Delta G/k_BT}$ where $\Delta G$ is the difference in binding free energy between the ligands (1–3). However, the fidelity of many important biological processes has been observed to be higher than what can be expected from the measured differences in the equilibrium binding free energies between ligands (1, 4–9). In one prominent examples, T cells appear to be able to distinguish between self-peptides and foreign antigens which may differ by as little as a single amino acid mutation (as little as a two-fold difference in affinity) (3, 4, 10–14). Such extreme specificity of the cellular response to cognate ligands is frequently explained in terms of a mechanism known as kinetic proofreading (KPR) (3, 7, 13–16).

The KPR mechanism enhances occupancy differences between ligands beyond the equilibrium factor $e^{\Delta G/k_BT}$ via introduction of one or more intermediate bound states between the initial binding and the final state where the output signal is produced (see Fig. 1A). The delay introduced by these additional states increases the probability of the ligand to dissociate from the receptor before the signal is produced, so that the weakly binding ligands rarely stay bound long enough to reach the signaling state and produce an output. Magnifying the differences between the ligands comes at the expense of energy consumption (typically in the form of ATP hydrolysis) to ensure irreversibly unbinding of the ligand from the proofreading states of the receptor (3, 16–18).

Kinetic proofreading was originally proposed by Hopfield and Ninio as a mechanism to explain the fidelity of DNA replication and protein biosynthesis and is believed to describe the proofreading step of the endonuclease and tRNA proofreading by the ribosome (1, 2, 19). McKeithan subsequently adapted multi-step KPR as an explanatory mechanism behind the remarkable specificity of the T cell receptor (TCR) in recognizing foreign versus self antigens (3). Subsequently, proofreading models and their modifications have been applied to explain additional features of the T cell response such as the speed and antagonism, while other works explored energetic and time constraints of generalized proofreading schemes (15–17, 20–23). KPR models have also been applied to other receptor signaling systems including the FcεRI-IgE system, RecA assembly, and various intracellular signal cascades (7, 8, 13, 14, 16, 18, 24–31). This body of work established the common view of the generalized sensing problem: discrimination of two different ligand types (such as self- vs. non-self peptides for T cells) based on a signal produced by the ligand binding to the receptor.

In the context of cellular sensing, receptor resolution for different ligands is commonly measured by the ratio of the corresponding average occupancy probabilities of the signaling state of the receptor, or by the ratio of physically related quantities such as the average time spent in the signaling state, or the average amount of the produced downstream signaling output (1–3, 13, 21). In this paper, we will refer to this metric as the Hopfield-Ninio (HN) metric.

The HN metric characterizes the resolution in terms of ensemble average properties, and thus does not take into account the stochasticiy in the molecular processes of the ligand binding and unbinding, proofreading steps, and the subsequent signaling reactions. However, as known in the context of concentration sensing, this stochasticity can have substantial effects on the sensing accuracy (32–39). As illustrated in Fig. 1B, the ability of a cell to resolve two different ligands depends not only on the separation of the mean outputs of

\[ e^{\Delta G/k_BT} \]

A.Z. and D.K. devised the project. D.K. performed the research. A.Z. and D.K. wrote the paper.

The authors declare no competing interests.

\textsuperscript{2}To whom correspondence should be addressed. E-mail: zilman@physics.utoronto.ca
Fig. 1. Ligand discrimination performance depends on molecular noise. A) Schematic diagram of a 1-step KPR receptor. The receptor in the unbound state can bind a ligand at a rate $k_{on}$, where $c$ is the ligand concentration; the ligand dissociates from a bound receptor at a rate $k_{off}$. In the bound state, the receptor may transition to the signaling state at a rate $k_f$, from which it can return to the unbound state again at rate $k_{off}$. While in the signaling state the receptor produces output molecules at a rate $k_{out}$. B) Ligand discrimination depends on the separability of output distributions dictated by the distance between their means and their widths. The blue distributions are easily separable, while the orange distributions that have the same distance between their means are inseparable due to greater variance in $n$. C) Ligand discrimination can be formulated as the estimation of the dissociation rate $k_{off}$ from the output $n$. The noise in the output $n$ described by the distribution $P(n|k_{off})$ translates into a distribution of errors in the estimate for $k_{off}$ which is centered on $k_{off}^* = k_{off}$. We plot the likelihood, $P(k_{off}|n)$, of $k_{off}$ for a particular $n$. The estimation function (black curve) determines how variance in $n$ translates into variance in $k_{off}^*$; we demonstrate two regimes (orange vs. blue) where the variance in $n$ has less or more effect on the variance in $k_{off}^*$ respectively.
The paper is structured as follows. In Section 2A we briefly review the KPR model. In Section 2B we show how the intrinsic molecular noise affects the output distribution in the KPR model and examine the signal-to-noise ratio of the model. In Section 2C we formulate the problem of specificity as the statistical estimation of ligand affinity and examine how receptor occupancy fluctuations affect the ability to identify ligands, an important precursor for considering the problem of ligand discrimination. In Section 2D we introduce a metric for ligand resolution derived from statistical estimation theory. In Section 2E we show that resolution enhancement by KPR is severely restricted due to the amplification of the relative noise by the KPR scheme. In Section 2F we investigate alternative KPR schemes and metrics to show that our results are inherent to the proofreading mechanism itself rather than a particular choice of the noise model of a discrimination performance metric. We conclude with a discussion of our results in the context of previous work, and their consequences for the understanding of the roles of KPR in cellular signaling.

2. Results

A. The KPR Model. We model a KPR receptor as a system with \( N + 2 \) states. These states represent the unbound receptor (state 0), the bound receptor in each of the \( N \) proofreading states, and the final signaling state of the receptor (numbered as \( N + 1 \), see Fig. 1A). Thus, \( N \) denotes the number of proofreading states and \( N = 0 \) corresponds to a non-proofreading receptor.

The ligand, present at concentration \( c \), can bind an unoccupied receptor with the overall rate \( k_{on}c \). Once bound by a ligand, the receptor can sequentially transition from the initial bound state to each subsequent proofreading state at a rate \( k_f \). The ligand can unbind from the receptor in any state at a rate \( k_{off} \). We assume that the ligand-receptor association rate \( k_{on} \) is diffusion limited and is independent of the ligand identity so that only the unbinding rate is related to the ligand-specific binding free energy (i.e. \( k_{off} \propto e^{-G_{bind}/k_BT} \)).

Note that the equilibrium dissociation constant, also known as the affinity, is \( K_d = k_{off}/k_{on} \) and in our case is uniquely determined by \( k_{off} \). In the terminal bound state the receptor produces downstream signaling molecules at a rate \( k_p \), accumulating \( n \) output molecules over time \( t \). In the following, we measure time in units of \( k_p^{-1} \) throughout this paper. This type of receptor output model is a common representation of a signaling mechanism by phosphorylation of intracellular molecules by a receptor-bound kinase (62–65). Other choices of receptor output models are investigated in Section 2F.

Ligand binding and unbinding are subject to intrinsic thermal and chemical fluctuations so that the transitions between the states in the model are probabilistic with the average rates described above. Together with the stochastic fluctuations in the production of the output molecules, this intrinsic noise leads to the distribution \( P(n|k_{off}) \) of possible outputs, \( n \), from the receptor in response to a ligand with the unbinding rate \( k_{off} \) (see Section 2F for models with minimal chemical noise).

At long times — defined as \( \min(k_{off},k_f) \gg 1 \) — or fast phosphorylation (\( k_p t \gg 1 \)) the output distribution for an \( N \)-step receptor is well approximated by the Normal distribution \( \langle n \rangle N \approx N\sigma^2 N \) with the mean output from the receptor \( \langle n \rangle N \), its variance \( \sigma^2 N \) (53). In some situations, the cell needs to make a ligand discrimination decision in a short time, before the output distribution has reached the Normal distribution. This regime is analysed in Section 2F and the main results of the paper are recapitulated in this regime as well.

The expressions for the mean and the variance can be derived from the Master equation of the KPR model (see Materials and Methods):

\[
\langle n \rangle N = k_p t \left( \frac{1}{1 + g} \right)^N \frac{x}{1 + x} \tag{1}
\]

\[
\sigma^2 N(x,t) = \langle n \rangle N \cdot \left( 1 + 2 \frac{k_p}{k_{off}} \right) + 2k_p^2 x^2 (2 + x + \frac{N}{2} + x + N x) \frac{(1 + g)^{2N+1}}{(1 + x)^{2N+2}} \tag{2}
\]

where \( N \) is the number of proofreading steps (\( N = 0 \) is a non-proofreading receptor), \( g = k_{off}/k_f \) is the proofreading strength, and \( x = c/K_d \) is the non-dimensionalized ligand concentration. Equation (1) expresses the fact that at long times the number of output molecules produced by time \( t \) is simply the product of \( k_p t \) and the probability to be in the signaling state \( P_{N+1} = x/(1 + x)(1 + g)^N \). Equation (2) reflects the fact that the variance in the number of output molecules is the sum of the intrinsic variance of the receptor occupancy distribution and the variance of the Poissonian process producing output molecules at rate \( k_p \) (53).

B. Noise in the receptor output space. As illustrated in Fig. 1B-C, the width of the distribution \( P(n|k_{off}) \) is an important factor controlling the ability to resolve two ligands. In this section we explore how proofreading affects the width of the receptor output distribution.

The signal-to-noise ratio (SNR), \( \langle n \rangle N / \sqrt{\sigma^2 N} \), is commonly used to quantify the effect of noise in any signaling system, and is higher when the signal distribution is more tightly centered about its mean. In Fig. 2A we show the SNR for a KPR receptor as a function of the number of proofreading steps, \( N \), and the proofreading strength, \( g = k_{off}/k_f \). Critically, the SNR decreases with increasing proofreading strength (larger \( k_{off} \) or smaller \( k_f \)). This trend can be understood from Eqs. 1-2, which show that in the high \( g \) limit the mean signal, \( \langle n \rangle N \), scales like \( O(1/g^N) \) while the noise, \( \sqrt{\sigma^2 N} \), scales like \( O(1/g^{N/2}) \). Thus, the SNR scales as \( O(g^{-N/2}) \), and strong proofreading yields a lower SNR. Physically, the decrease in the SNR with increasing proofreading strength stems from the fact that a higher proofreading strength increases the variance in both the receptor occupancy distribution and the Poissonian process producing output molecules (see Eq. 2), amplifying the relative noise strength. We show in Section 2F that the increase in the receptor occupancy variance is inherent to KPR and sufficient to decrease the receptor’s signal processing performance.

The SNR also reveals that the output distributions from different ligands increasingly overlap as the number of proofreading steps increases, despite the increased separation of their means, as shown in Fig. 2B. The difference in the mean outputs in response to each ligand (blue and orange lines) increases with \( N \), illustrating the classical proofreading effect. However, despite this increasing separation between the means, the output distributions become progressively more overlapped [colored envelopes represent \( \langle n \rangle N(k_{off},i) \pm \sqrt{\sigma^2 N(k_{off},i)} \)]. The
SNR therefore already provides an indication that the noise associated with proofreading makes ligand resolution harder.

In the next section we formulate receptor specificity in a precise mathematical manner and show that the intrinsic noise of the KPR scheme poses a challenge for KPR as a mechanism for enhanced specificity.

### C. Proofreading decreases specificity

To give a precise meaning to the problem of ligand resolution, and thereby quantify the effects of noise on ligand resolution, we frame it as a problem of determining the ligand identity based on the observed receptor output, $n$. This entails statistical estimation of $k_{\text{off}}$ using an estimator function $k_{\text{off}}^*(n)$. For Normally distributed outputs considered here, choosing $k_{\text{off}}^*$ to be the maximum likelihood estimator (MLE) — that is, the estimate for which $P(n|k_{\text{off}})$ is maximized — guarantees an unbiased estimate (i.e., $E[P(n|k_{\text{off}})|k_{\text{off}}^*] = k_{\text{off}}$) with a minimum variance over different estimators (35, 50, 60). At long times, the MLE is well approximated by the inverse function $\langle n \rangle^{-1}(k_{\text{off}})$ obtained by solving Eq. 1 for $k_{\text{off}}$ (53, 60).

As illustrated in Fig. 1C, the distribution of receptor outputs, $P(n|k_{\text{off}})$, gives rise to the distribution of the estimates of $k_{\text{off}}$, $P(k_{\text{off}}|k_{\text{off}}^*)$ because the (deterministic) estimate $k_{\text{off}}^*$ depends on $n$ sampled from the distribution $P(n|k_{\text{off}})$ (53, 60).

Since the estimator is unbiased, $P(k_{\text{off}}|k_{\text{off}}^*)$ is centered at the correct estimate, $k_{\text{off}}^* = k_{\text{off}}$. The width of the distribution of estimates depends on both the width of the output distribution, $P(n|k_{\text{off}})$, and the estimator function, $k_{\text{off}}^*(n)$, as illustrated in Fig. 1C. The Cramér-Rao lower bound provides an expression for the minimum expected variance of $P(k_{\text{off}}|k_{\text{off}}^*)$ as: $\mathbb{E}[\text{Var}(k_{\text{off}}^*)] \geq -I^{-1}$ where $I = -\mathbb{E}[\partial^2 \ln P(n|k_{\text{off}})/\partial k_{\text{off}}^2]$ is the Fisher Information and all expectations are taken over $P(n|k_{\text{off}})$ (60). Intuitively, the more information about $k_{\text{off}}$ that is contained in $P(n|k_{\text{off}})$, the lower is the variance in the estimate $k_{\text{off}}^*$. For a Normally distributed receptor output, $n$, the Fisher Information can be calculated as (53, 60):

$$I = \frac{\langle \partial(n|n) / \partial k_{\text{off}} \rangle^2}{\sigma_{k_{\text{off}}}^2} \simeq \frac{1}{\sigma_{k_{\text{off}}}^2}.$$  

For sharply peaked output distributions the distribution of estimates, $P(k_{\text{off}}^*|k_{\text{off}}^*)$, is well approximated by $N(k_{\text{off}}, \sigma_{k_{\text{off}}}^2)$, where $\sigma_{k_{\text{off}}}^2 = -I^{-1}$ is the estimate variance. The estimate distribution for general $N$ has variance:

$$\langle \sigma_{k_{\text{off}}}^2 \rangle_N = \frac{(1+g)^N + 2 \frac{k_{\text{off}}}{k_{\text{off}}^*} (1+x)^3}{\frac{k_{\text{off}}}{k_{\text{off}}^*} x (1+g (1+N+Nx)^2)^3} - 2(1+g)^{N+2} x (1+x) \left(2+x+g(2+2N+n+Nx)\right)^2,$$

For strong proofreading the variance of the estimate scales as $O(g^N)$. We and others have recently quantified receptor specificity by the ratio $k_{\text{off}}^*/\sigma_{k_{\text{off}}}$, which is similar to the SNR but measures the relative width of the estimation distribution, $P(k_{\text{off}}^*|k_{\text{off}}^*)$ (48, 53). The specificity ratio is inspired by the measure of accuracy of ligand concentration sensing first calculated by Berg and Purcell (34). Specificity is a critical prerequisite for good ligand resolution because greater specificity allows the ligand affinities to be closer together without their estimates being statistically indistinguishable.
We demonstrate in Fig. 2C that, except for low N in the weak proofreading regime, g \geq 1, the specificity (calculated from Eq. 4) decreases with increasing number of proofreading steps, N. Notably, this detrimental effect of proofreading persists throughout the regime where weakly binding ligands are strongly proofread but strongly binding ligands are not (corresponding to 1 < g < 10 for the weakly binding ligand in Fig. 2C), which is the regime where KPR is commonly invoked - such as in TCR signaling. The decrease in specificity at large N is a consequence of the increased relative noise in the receptor output as proofreading steps are added (see Fig. 2B). Note that we interpolate the results for non-integer N, which can be thought of as the effective number of proofreading steps for imperfect proofreading such as when the proofreading steps are not perfectly irreversible (14). Specificity is maximized for low but nonzero values of N in the weak proofreading regime because of a balance between the detrimental effect of molecular noise and the beneficial effect of increased sensitivity to changes in k_{off} arising from the term |\partial(n(N))/\partial k_{off}| in Eq. 3. At very small g the receptor spends effectively no time in the proofreading states and the specificity is decreased again. The balance of these effects results in the maximization of the specificity for low but nonzero values of N in the weak proofreading regime; the location of this maximum in (g, N)-space also depends on the ligand concentration (see Fig. S1 C). However, the enhancement of specificity compared to a non-proofread receptor is very moderate (less than 10-fold for any N with \( \zeta = 0.1 \), Fig. S1 G) and is substantially lower than the classical and commonly invoked exponentially strong enhancement in the strong proofreading regime. Thus, the fact that specificity is only enhanced in this regime casts a new light on the performance and the purpose of KPR in such applications.

Another approach to quantify the effects of noise on ligand discrimination is through the mutual information, I(n, k_{off}), between the input variable (ligand affinity), k_{off}, and the receptor output, n (18, 22, 59). Although not the main thrust of this paper, it has been shown that mutual information is closely related to the Fisher Information which we considered in this section (66, 67). As shown in the Supporting Information (S.I.) Section A, proofreading also decreases the discrimination when quantified through the mutual information.

We have shown that both strong proofreading and a large number of proofreading steps generally lead to very noisy estimates for a single ligand affinity. The greater the spread in the estimate distribution for a single ligand (i.e., the specificity), the less well resolved two different ligands are expected to be. In the next section we give this notion a precise mathematical meaning by constructing a resolution metric for two ligands.

D. A metric for two-ligand resolution. While the SNR and the specificity are useful for quantifying noise in KPR performance both metrics describe the signaling statistics for only a single ligand. We now extend the idea of specificity to the problem of discriminating between two ligands that are are distinguished by their respective ligand-receptor dissociation rates k_{off,1} and k_{off,2}. The differences in the affinity between the ligands is characterized by the ratio \( \zeta = k_{off,2}/k_{off,1} \) (we define \( \zeta \in (0, 1) \) so that \( k_{off,1} > k_{off,2} \)). An appropriate resolution metric should encapsulate how well the difference in affinity between the two ligands, \( \zeta \), can be resolved in the presence of noise in the output distribution of each ligand.

The standard metric for ligand resolution in the KPR model, first used by Hopfield and Ninio, uses the ratio of the mean outputs for the two ligands (1-3):

\[
\eta_{HN} = \frac{(n)(k_{off,2}, c, t)}{(n)(k_{off,1}, c, t)}
\]

where \((n)(k_{off,i}, c, t)\) is the mean signal produced by time \( t \) by the receptor in response to a ligand with dissociation rate \( k_{off,i} \), at fixed concentration \( c \). In our convention \( \zeta < 1 \) and \( \eta_{HN} \geq 1 \) since the weaker binding ligand forms less long-lived ligand-receptor complexes and therefore produces less output. The larger \( \eta_{HN} \), the better these ligands can be discriminated. For concreteness, we choose \( \zeta = 0.1 \) here and throughout this paper as informed by the typical value for self-peptides versus weak antigens to be separated by the TCR (3, 12, 14, 18, 68).

As discussed above, the metric \( \eta_{HN} \) relies only on the average receptor outputs and does not account for the noise in the output distributions which we have shown to be so important for ligand discrimination.

To take the noise into account, a resolution metric must quantify the distance between the two estimate distributions of \( P(k_{off,1}^\ast | n) \) and \( P(k_{off,2}^\ast | n) \). Intuitively, two distributions are well separated when their means are separated by more than the sum of their standard deviations (e.g., see Fig. 1B).

This intuition is encapsulated in the Fisher’s linear discriminant (69, 70), based on which we define our ligand resolution metric:

\[
\eta_{FLD} = \frac{(\langle k_{off,1}^\ast \rangle - \langle k_{off,2}^\ast \rangle)^2}{\sigma_{k_{off,1}}^2 + \sigma_{k_{off,2}}^2}
\]

where \( \langle k_{off,i}^\ast \rangle \) is the mean estimate for \( k_{off,i} \) generated by ligand with \( k_{off,i} \), and \( \sigma_{k_{off,i}}^2 \) is the variance about this mean. The averages in Eq. 6 are taken over \( P(n|k_{off,i}) \), and we assume an unbiased estimator so that \( \langle k_{off,i}^\ast \rangle = k_{off,i} \) (60). Note that in the limit of similar ligands we have \( k_{off,2} \approx k_{off,1} + \delta k_{off} \) and \( \sigma_{k_{off,2}}^2 \approx \sigma_{k_{off,1}}^2 + \delta k_{off}^2 \), so that:

\[
\eta_{FLD} \approx \frac{(k_{off,1} - k_{off,2})^2}{2\sigma_{k_{off}}^2}
= \frac{(\langle n \rangle N / \partial k_{off})^2 (k_{off,1} - k_{off,2})^2}{\sigma_N^2}
\lesssim \frac{(\langle n \rangle (k_{off,1}) - (n)(k_{off,2}))^2}{2\sigma_n^2}
\]

where we have used the Fisher Information of (Eq. 3) and in the last line we have interpreted \( \delta k_{off} = k_{off,1} - k_{off,2} \). Thus, the Fisher linear discriminant formulated in \( k_{off}-\)space is very closely related to the equivalent formulation in \( n \)-space that conveys the intuitive meaning about the ability to separate the two output distributions. For both \( \eta_{HN} \) and \( \eta_{FLD} \), a higher value indicates higher resolution. However, we show in the following section that KPR has severely restricted benefits for ligand discrimination, relative to a non-proofreading receptor, when noise is accounted for in quantifying the resolution.

E. Increasing the number of proofreading steps does not enhance discriminatory power. For a KPR receptor with \( N \) proofreading steps, the Hopfield-Ninio resolution (Eq. 5) is:

\[
\eta_{HN}^N = \left( \frac{1 + \frac{1}{\zeta}}{\frac{1}{\zeta} + 1} \right)^N \frac{1 + x}{\zeta + x}
\]
using \( x = k_{\text{on},c}/k_{\text{off},1} \), \( g = k_{\text{off},1}/k_f \) is the proofreading strength for the weaker binding ligand, and \( \zeta = k_{\text{off},2}/k_{\text{off},1} \). In the limit of strong proofreading (\( 1/g \to 0 \)) and low concentration (\( x \ll 1 \)), the resolution described by Eq. (7) scales as \( \eta_N^{\text{HN}} \sim O(1/C^{N+1}) \). At high concentrations (\( x \gg 1 \)) and strong proofreading the resolution scales as \( \eta_N^{\text{HN}} \sim O(1/C^N) \) due to the saturation of the receptor by the ligand (Section 2C). Thus, the enhancement of the resolution by \( N \) proofreading steps is given by:

\[
\frac{\eta_N^{\text{HN}}}{\eta_0^{\text{HN}}} = \left( \frac{1 + g}{1 + \zeta g} \right)^N. \tag{8}
\]

In the strong proofreading regime this simplifies to \( \frac{\eta_N^{\text{HN}}}{\eta_0^{\text{HN}}} \sim O(\zeta^{-N}) \), while in the weak proofreading regime this simplifies to \( O(1) \). The behavior of \( \eta_N^{\text{HN}} \) expresses the well known result that in the strong proofreading regime, additional proofreading steps exponentially enhance mean resolution relative to a non-proofreading receptor. However, when noise is included in the resolution metric, as in \( \eta_{\text{FLD}} \), increasing the number of steps does not necessarily increase the discrimination strength because the separation of the means \((n)(k_{\text{off},1})\) grows slower with \( N \) compared to the variances \( \sigma_n^2 \), obscuring the distinctions between ligands.

A closed form expression for the resolution of the KPR receptor using the metric \( \eta_{\text{FLD}} \) is cumbersome (provided for completeness in S.I.), but can easily be evaluated numerically. In Fig. 3A we show how \( \eta_{\text{FLD}} \) depends on the number of proofreading steps, \( N \), and proofreading strength, \( g \) (fixing \( k_f = 1 \) so that \( g = k_{\text{off}} \)). Unlike \( \eta_N^{\text{HN}} \), here an intermediate proofreading strength for the weaker binding ligand \((g \sim O(1))\) and a few proofreading steps yield the greatest resolution, similar to the specificity in Fig. 2C, reflecting the mathematical similarity of the two metrics (discussed in Section D).

Although the location of the maximum of the resolution, \( \eta_{\text{FLD}} \), depends on the ligand concentration \( c \) (see Fig. S1 D-E), the maximum is attained at weak proofreading strengths \((g < 1 \text{ for both ligands})\) for all concentrations (see inset of Fig. S1 D). The commonly considered context for KPR is where weakly binding non-agonists are suppressed by proofreading \((1 < g_{\text{weak}} < 10)\) while strongly binding agonists are hardly suppressed \((0.1 < g_{\text{strong}} < 1)\) (12, 14). In this biologically important regime, denoted with a dashed line in Fig. 3A, the resolution monotonically decreases as proofreading steps are added.

We visualise the resolution enhancement relative to the non-proofreading receptor by the ratio \( \eta_{\text{FLD}}^N/\eta_{\text{FLD}}^0 \) in Fig. 3B, which shows that the resolution does not improve with additional proofreading steps. Instead, the ratio scales inversely with \( N \) as \( N \) becomes large (green lines in Fig. 3B). While there can be moderate resolution enhancement for a few proofreading steps in the regime of weak proofreading, this enhancement is always ablated in the large \( N \) limit since the ratio scales as

\[
\frac{\eta_{\text{FLD}}^N}{\eta_{\text{FLD}}^0} \propto O\left( \frac{(N^2g)\zeta^3}{(1 + g)^{N+2}} \right) \tag{9}
\]

which tends to zero for \( N \to \infty \) — opposite to the scaling of \( \eta_N^{\text{HN}}/\eta_0^{\text{HN}} \) shown in Fig. 3B (black line). The resolution enhancement for a broad range of proofreading strengths is provided in the S.I. (Figure S1 G), which also shows that the fold-change in resolution given by \( \eta_{\text{FLD}}^N/\eta_{\text{FLD}}^0 \) is significantly less than that indicated by \( \eta_N^{\text{HN}}/\eta_0^{\text{HN}} \).

The resolution at large \( N \) decreases because the relative noise of the receptor output is always increased by increasing the number of proofreading steps (see Fig. 2A). In principle, the molecular noise of the receptor could be reduced in different variants of the KPR mechanism. However, in the next section we demonstrate that even in models with minimal noise, the inherent stochasticity of the KPR mechanism still results in diminished resolution capacity compared to a non-proofread receptor (see also S.I. Section E).

F. Alternative metrics and models of noise reveal intrinsic disadvantages of proofreading.

F.1. Low noise models. The KPR model considered so far has several sources of noise. These include the intrinsic noise inherent to the probabilistic nature of the sequential state transitions during proofreading, the variability of the residence times in different states, and the noise in the production of the receptor output \( n \). Furthermore, for mathematical tractability all rate processes have been assumed to be Poisson processes

![Fig. 3. Kinetic proofreading does not enhance resolution.](https://doi.org/10.1101/2023.01.13.523988)
with exponentially distributed reaction times.

In this section, we investigate alternative KPR schemes with diminished noise to show that our results are not an artefact of a particular choice of the (high noise) model. In the first scheme (see Section D of the S.I.), to understand the role of the noise in the production of the output variable, we consider a KPR receptor which produces bursts of $\beta$ output molecules immediately upon arriving in the active state $(N + 1^{\text{th}})$ state. This model can serve as a representation of a fast signaling output burst by a G-protein coupled receptor (53, 71, 72). Unlike a Poisson rate process with rate $k_{\text{on}}$, this burst process does not add extra noise associated with the output generation.

The mean number of the output molecules, $\langle m \rangle$, produced by this burst model is

$$\langle m \rangle_N = \frac{\beta k_{\text{on}} t x}{(1 + g)^N (1 + x)}.$$  \hspace{1cm} \text{[10]}

The variance in the number of output molecules as a function of $N$ is derived in the Section D of the S.I.

In the second alternative scheme, we investigate the role of variance in the residence time in the active state of the receptor. To this end, in addition to the burst output the residence time in the active state, rather than obey an exponential distribution, is a constant $\tau = 1/k_{\text{off}}$ for each ligand. Furthermore, in the S.I. we present another version of this model where all transition times between states are $\delta$-rather than exponentially distributed. This hybrid model has been investigated using stochastic simulations (see Methods).

The results of these minimal noise models are shown in Fig. 4A. The burst model shows a significant decrease in the resolution for the addition of a single proofreading step and a steady decline as $N$ increases. The model with a fixed residence time is less noisy and consequently shows better resolution than either burst or Poisson types of output with exponentially distributed times. Yet, both minimal noise models exhibit decreasing discriminating power with increasing number of proofreading steps, similar to the original KPR model (Eq. 9).

As above, the resolution of these models is worse in the strong proofreading regime ($g = 10$, solid line) compared to the weak proofreading regime ($g = 3$, dashed line).

F.2. Alternative metrics of ligand resolution. In the previous sections we have assumed that the measurement time is sufficiently long so that the distribution of the output, $n$, is well approximated by a normal distribution. However, in some situation – notably the problem of antigen detection by TCR – it is possible that a cell must determine a ligand identity very quickly, ruling out the possibility for many binding-unbinding events to occur during the measurement time. We investigated this regime using stochastic simulation (see S.I. Section G), and the results are shown in Fig. 4B, which shows that even at very short times $\eta_{\text{LD}}$ decreases with $N$, as seen for long measurement times.

Furthermore, in some cases – including the problem of TCR discrimination – it can be argued that the primary purpose of KPR for a T cell is the maximal suppression of false positive errors from the weakly binding ligand (even at the expense of the somewhat increased false negative rate for the strongly binding ligand). Spurious activation of the immune system by weak binders - which carries the potential of substantial damage to the host - may be more dangerous to the host compared to false negative errors of occasional non-detection of the strongly binding ligand because the latter can be compensated for by other clonal cells that identify the non-self antigen (73, 74).

We quantify these notions through a metric defined as the ratio of the probability of false activation to the probability of all activation events. We assume that activation occurs when the output variable $n$ crosses an activation threshold $n^*$, with probability $P(n(k_{\text{off}, \text{weak}}, c) > n^*)$. The threshold is set to a value such that the probability of activation by the strongly binding ligand is $10\%$: $P(n(k_{\text{off}, \text{strong}}, c) > n^*) = 0.1$ (the results are robust with respect to the specific choice of this value). In Fig. 4C we show the probability ratio of false activation to the total activation as a function of the number of proofreading steps This plot reveals that the ratio is minimized for a small number of proofreading steps and increases as $N$ is increased further. The location of this minimum shifts to larger $N$ as the proofreading strength decreases, and the minimum shifts to smaller $N$ as the proofreading strength increases. We show the ratio for $g_{\text{weak}} = 10$ and $g_{\text{weak}} = 3$, which represent a typical range of proofreading strengths for the TCR according to recent experimental measurements (12–14).

Overall, our results demonstrate that KPR has severely limited benefits for ligand discrimination, only providing moderate enhancement in ligand resolution in the non-classical regime of fast proofreading and a few proofreading steps.

3. Discussion

In this paper we have shown that intrinsic molecular noise greatly affects the KPR-based ability of a receptor to discriminate between ligands, and that resolution metrics must account for this noise to avoid misleading interpretations of the effects of KPR. To capture the role of the noise, we used the metric $\eta_{\text{LD}}$ (Eq. 6) which, unlike the classical metric $\eta_{\text{HN}}$ (Eq. 5), incorporates information not only about the means but also about the variances of the distributions of the estimated ligand identity, $P(k_{\text{off}, i}/k_{\text{off}})$. Taking the noise into account reveals that the benefits of proofreading for ligand discrimination are severely restricted to a moderate effect at low number of proofreading steps and low proofreading strength.

The central result of the classical analysis of KPR (based on $\eta_{\text{HN}}$) posits that the discrimination between two ligands improves with the number of proofreading steps as $O((N + 1)^{-1})$ in the strong proofreading regime, where $\zeta = k_{\text{off}, 2}/k_{\text{off}, 1}$ is the ratio of the affinities of the two ligands to be distinguished. By contrast, in Sections 2 B-E we demonstrated that proofreading amplifies the effect of intrinsic molecular noise in both the signaling state occupancy and the downstream output, thereby making ligand discrimination more difficult. In all regimes, for large $N$, the resolution enhancement, $\eta_{\text{LD}}^N/\eta_{\text{LD}}^0$, scales inversely with $N$ and tends to zero as $N \to \infty$.

One of the most common biological signaling systems where the KPR model has been invoked is T cell discrimination between foreign antigens and self peptides (43). These applications of the KPR model commonly focus on the scenario where the number of proofreading steps is large (18, 21, 27, 61) and the weakly binding self peptides are strongly proofread and attenuated, because this is where the proofreading is most effective according to the classical, noiseless, HN metric. However, the large $g$, large $N$ regime is also the regime where we showed the receptor output distribution to be most noisy leading to the degradation in discrimination performance (Section 8).
While experimental measurements of KPR parameters in biological systems have remained difficult, recent work has started to provide experimental estimates of the effective number of proofreading steps for the TCR, estimated to be $N \approx 3$ (13, 14), and estimates of the proofreading rate have been found to be $k_f \approx 0.3 \text{s}^{-1}$ (14). Thus, for a typical agonist binding time of 1 - 10 seconds, a typical non-agonist binding time of 0.3 - 3 seconds, and taking $k_{\text{off}} = \tau^{-1}$ for the binding time $\tau$ (12), the TCR operates with moderate proofreading strength: $0.3 < g < 3$ for the strongly binding ligand and $1 < g < 10$ for the weakly binding ligand. Weakly binding ligands may be proofread even more strongly according to recent re-measurement of ligand affinities for the TCR (14). These experimental measurements of $N$ and $g$ for the TCR are thus consistent with the conclusions of our theoretical framework, although full comparison of the theoretical predictions with the detailed the experimental data is outside the scope of the present work and will be investigated in the future.

The limited utility of KPR in ligand discrimination and identifiability stems from its two inherent properties: attenuation of the occupancy of the signaling state, and the accompanying amplification of the effects of noise with each additional proofreading step. The problems posed by signal attenuation were already noted by McKeithan (3), who proposed a variant KPR scheme in which the receptor is stabilized in the final signaling state. Such stabilization can be introduced into our model by reducing the unbinding rate from the final proofreading state by a factor $\alpha < 1$ (see Methods: Eq. 11). The mean signal for a stabilized KPR receptor is

$$\langle n \rangle_{N,\text{stabilized}} = k_f \frac{x}{\alpha(1+g)(1+x) + (1-\alpha)x}.$$

While stabilizing the signaling state does increase the mean signal, the standard deviation still scales like $O(g^{-N/2})$ so that the increased noise in the signal distribution still outpaces the increased difference in signals as proofreading steps are added (see also Fig. S2C).

The results of this paper are not an artefact of a particular choice of the noise model. The variability of the signaling output arises from several different components. The first two are the Poisson noise of the output production in the signaling state and the variability of the residence and arrival times to the signaling state. The third is the probabilistic nature of the binding-unbinding events and state transition steps, which are intrinsic to the proofreading mechanism. As shown in Section 2F, the results of this paper - the decrease of the resolution capacity with stronger proofreading and increasing number of proofreading steps - hold true even in the minimal noise models where the signaling output is produced in discrete bursts and the receptor residence times are not exponentially but rather $\delta$-distributed. Our results also hold true for fast decision-making processes which occur at short times when multiple binding-unbinding events are unlikely.

Our results rest on the appropriate choice of resolution metric $\eta_{\text{FLD}}$ that is widely used in statistical inference as a measure of the estimate quality and discriminability. However, in different biological situations, a different metric can be more appropriate - such as the fraction of the false positive activations. As shown in Section 2F, the results of this paper are not an artefact of a particular metric choice, and hold also for different measures of the discrimination quality; the
fraction of false positives generally decreases with an increase in the number of proofreading steps, except for a moderate minimum at low $N$ at low proofreading strength.

It is important to distinguish these results from other works that have sought to balance ligand discrimination performance with the extrinsically imposed costs of proofreading such as input energy and decision speed, finding that different choices of the extrinsic costs optimize the proofreading at different numbers of proofreading steps (22, 39, 44–46, 75). These approaches typically posit that larger $N$ will enhance ligand discrimination so long as the extrinsic costs can be met. By contrast our result indicate that the discrimination capacity itself decreases with the proofreading strength due to the intrinsically stochastic nature of KPR.

Our results naturally lead to a question: if KPR commonly does not enhance the resolution of a sensory system, what other purpose might it serve in receptor systems like the TCR? One possibility is that KPR might enable multiplexing of signals in receptor mixtures by producing multiple statistically independent outputs from different proofreading states (53, 76, 77). These outputs can be used to infer the presence of multiple ligands in a mixture (53), convey information about ligand concentration or other environmental conditions in addition to ligand identity (77), or perform other complex signaling processing tasks (78). Mounting experimental evidence indicates that the TCR (as a case study for KPR) may indeed produce multiple intracellular second messengers which may enable T cells to detect antigen quantity alongside antigen quality (78–82). Our results lay the ground for further theoretical work aimed at understanding the effects of noise in multi-output receptor signaling pathways (78, 83–85).

### Materials and Methods

The KPR receptor states are enumerated by index $i$, with $i = 0$ representing the unbound state and $i = N$ the signaling state. The system is fully characterized by the pair of state variables $(i, n)$ representing the current receptor state and number of copies of the downstream signal molecule respectively. The probability of the system to be in the state $i$ at time $t$ is $P^n_i(c, k_{\text{off}}, t)$ which depends on the ligand concentration $c$ and the unbinding rate $k_{\text{off}}$. The dynamics of this probability distribution is described by the Master equation

$$ \frac{\partial P^n_i}{\partial t} = -k_{\text{on}}cP^n_0 + k_{\text{off}} \sum_{i} P^n_i + k_fP^n_0 - (k_f + k_{\text{off}})P^n_i \ldots $$

where $k_{\text{on}}$ association rate of the free ligands with an unoccupied receptor, $k_f$ is the rate at which the receptor transitions from state $i$ to state $i + 1$, $k_{\text{off}}$ is the rate of signal production from the terminal receptor state; we assume that the ligands are in a large extracellular volume so that $c$ remains constant. A stabilized KPR receptor mentioned in the Discussion has $k_{\text{off}}$ reduced by a factor $\alpha$ in the last line of Eq. 11. For the burst model of Section 2F the terms involving $k_f$ are not present and the term $k_fP^n_i$ becomes $k_fP^n_{i-1} - \beta$ (see also S.I.). The mean and variance for the burst model are solved in a similar manner to that of the phosphorylation model, described below.

The probability distribution of the output signal is $P(n|k_{\text{off}} = \sum_{i} P^n_i(c, k_{\text{off}})$. Accordingly, the mean downstream output signal

$$ \langle n \rangle = \sum_{n=0}^{\infty} \sum_{i=0}^{N} nP^n_i $$

Taking the derivative with respect to time and using the master equation above, the summation over $P_i$ reduces this expression to

$$ \frac{\partial \langle n \rangle}{\partial t} = k_{\text{off}} \sum_{n=0}^{\infty} n(P^n_{N+1} - P^n_0) $$

Re-indexing the term $P^n_{N+1}$ leads to

$$ \frac{\partial \langle n \rangle}{\partial t} = k_f \sum_{n=0}^{\infty} (n^2P^n_0 \ldots $$

where $P_N$ is the steady-state occupancy probability for the $N$th receptor state. The steady-state receptor occupancy probabilities $P_i(c, k_{\text{off}})$ can be obtained by algebraically solving the system of equations resulting from setting all time derivatives and $k_f$ to zero in Eq. 11 and summing over the state variable $n$ in $P^n_i$. Finally, the mean signal downstream of an $N$-state KPR receptor is

$$ \langle n \rangle = k_fP_N(c, k_{\text{off}}) $$

Using the expression for $P_N$ for a general $N$-step proofreading receptor that can be obtained from the Master equation Eq. 11 yields Eq. 1 for the mean signal (3): The case of $N = 0$ is a non-proofreading receptor. The same procedure can be used to compute the average output from the G-protein type KPR receptor, with the main difference being that the Master equation yields the relation

$$ \langle n(\infty) \rangle = k_{\text{on}}P_{N-1} \quad \text{and} \quad \langle n(\infty) \rangle / \langle m(\infty) \rangle = k_{\text{on}}cP_0 \quad \text{for} \ N = 0. $$

The variance in the downstream signal, $\sigma^2_N$, can be obtained from the second moment of the signal along the identity $\sigma^2_N = \langle n^2 \rangle - \langle n \rangle^2$. The second moment is defined as

$$ \langle n^2 \rangle = \sum_{n} \sum_{i=0}^{N} n^2P^n_i $$

Once again, taking the derivative with respect to time and substituting the Master equation yields a closed form expression:

$$ \frac{\partial \langle n^2 \rangle}{\partial t} = k_f \sum_{n} (2n+1)^2P^n_0 $$

Re-indexing the term $P^n_{N+1}$ leads to:

$$ \frac{\partial \langle n^2 \rangle}{\partial t} = k_f \sum_{n} (2n+1)P^n_0 $$

A closed form expression for the term $2k_f\sum_{n} nP^n_0$ can be obtained by solving the associated set of inhomogenous first order differential equations:

$$ \frac{\partial \langle n \rangle}{\partial t} = \sum_{n} \frac{nP^n_0}{\partial t} $$

for $i \in [0, N]$. Substituting the resulting expression for $\langle n \rangle_i$ into the differential equation for $\partial(n^2)/\partial t$, and integrating over time yields an expression for the second moment, which can then be used to find the variance $\sigma^2_N$. Using closed form expressions for $P_N$, the variance of the downstream signal is, in the long time limit $\min(k_{\text{on}}c, k_f) \to \infty$, Eq. 2. For the G-protein type KPR receptor the Master equation yields $\langle n(\infty) \rangle / \langle m(\infty) \rangle = k_f \sum_{n} (2n+1)P^n_{N-1}$, but the procedure is otherwise the same.

We view the T cell sensing problem as fundamentally a classification task based on estimating the ligand dissociation rate $k_{\text{off}}$. A maximum likelihood estimator (MLE) for $k_{\text{off}}$ can be constructed from the distribution $P(n|k_{\text{off}}, c) = \mathcal{N}(n, \sigma^2_N)$. At long times the limit of large numbers applies, making the distribution $P(n|k_{\text{off}}$ narrowly peaked about its mean so that this MLE is well approximated by (35, 53):

$$ k_{\text{off}} = \langle n \rangle^{-1}(k_{\text{off}})_{n=\langle n \rangle} \quad [12] $$
The authors are indebted to Golan Bel, \( \ell \), where \( \delta n \)
10. MA Daniels, et al., Thymic selection threshold defined by compartmentalization of Ras/MAPK
 Kirby et al.


