MAKING EVOLUTION RIGOROUS- THE ERROR THRESHOLD

NISHEETH K. VISHNOI

[Extended Abstract of this paper to appear in Innovations in Theoretical Computer Science (ITCS), 2013]

ABSTRACT. In a series of influential papers Eigen and his coworkers introduced the quasispecies model as a first-order approximation to (Darwinian) evolution and applied it to self-reproducing molecules such as RNA or DNA in an attempt to explain the origin of genetic information which gives rise to life. It was argued by Eigen that an important property in such a model of evolution is the existence of an error threshold: A rate of error during the reproduction phase below which genetic information is intact and above which it disappears. Besides the insights the quasispecies model has provided on the emergence of life, perhaps its most powerful impact has been on the study of viruses where error threshold phenomena has been leveraged to design drug strategies that attempt to mutate the virus to death. While the existence of error thresholds for specific settings has been verified by computer simulations, and has been the basis for the design of mutagenic drugs, a mathematical proof of this phenomena has remained elusive. The trouble is that one can construct pathological examples for which no non-trivial error threshold can exist. In this paper we present a proof of existence of a sharp error threshold in the quasispecies model for a large set of biologically relevant evolutionary parameters. Our analysis benefits from viewing the quasispecies model as an evolutionary process on the hypercube which permits the use of simple yet powerful ideas from linear algebra and Fourier analysis.

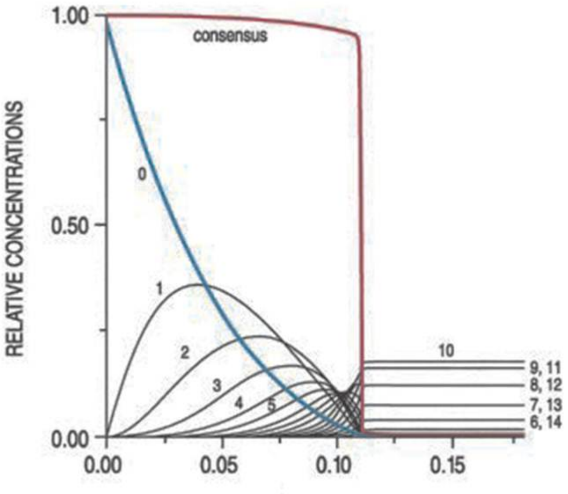

FIGURE 1. Simulation results for a single peak model on binary strings of length 20 where the curves shown correspond to the relative concentration of strings of the given Hamming weight at the steady state. The illustration implies a phase transition when the error rate is about 0.11. From [Eig02].

Nisheeth K. Vishnoi. Microsoft Research, Bangalore, India. Email: nisheeth.vishnoi@gmail.com.

1. INTRODUCTION

Eigen *et al.* [Eig71, ES77, EMS88, EMS89] introduced the quasispecies model, a remarkably clean mathematical model that captured the error-prone evolution of mol-ecules such as RNA and addressed the origins-of-life question. It was argued, see [Eig93], that an important property of such a model is the existence of an *error threshold*[1], *i.e.*, a replication error rate below which genetic information is intact and above which it is destroyed. Besides the insights the quasispecies model provides on the emergence of life, it has had a powerful impact on design of drug strategies for RNA-viruses such as Foot-and-Mouth-Disease [HD98], Poliovirus [CCA01] and even RNA *retroviruses* such as HIV [LEK$^+$99]. In all these cases, the error threshold phenomena is leveraged to attempt to mutate the virus to death, see [Vis12]. While the existence of error thresholds for specific settings has been verified by computer simulations [EMS89, SS82] (e.g., Figure 1) for various biological settings, a mathematical proof of this phenomena did not exist. Problematically, one can construct pathological examples for which no non-trivial error threshold exists. The main contribution of this paper is to rigorously study error thresholds in the quasispecies model of evolution, identify a class of biologically relevant parameters and prove the existence of error threshold in such cases.

1.1. **The Quasispecies Model.**

We describe the quasispecies model for discrete-time evolution of strings of length L comprising of two building blocks 0 and 1.[2] Thus, the model considers L-bit strings, $\{0,1\}^L$, and is parametrized by an *error* or *mutation rate* μ and a *fitness function* or *landscape* $a : \{0,1\}^L \mapsto \mathbb{Z}_{\geq 1}$. The mutation rate indicates the per-bit copying error during reproduction and for a string $\sigma \in \{0,1\}^L$, its fitness $a_\sigma \overset{\text{def}}{=} a(\sigma)$ captures its survivability with respect to an environment.[3] The process starts with a distribution over $\{0,1\}^L$ and it iterates deterministically applying the tenets of reproduction-selection-mutation. In the reproduction stage each σ produces a_σ copies of itself, in the selection stage we normalize, obtaining a new distribution over $\{0,1\}^L$, and finally, mutation occurs. The eventual outcome of the evolutionary process is not a single string (i.e., a species), rather it is an invariant distribution, justifying the term *quasispecies*. In this sense, it is a first-order approximation to Darwinian survival-of-the-fittest evolution.

Suppose $\mathbf{m}(0)$ is the starting distribution over the population at time 0 and, more generally, let $m_\sigma(t)$ denote the fraction of the population of σ at time t. If reproduction were error-free then $m_\sigma(t+1)$ would be proportional to $a_\sigma m_\sigma(t)$. Instead, with errors, for every string τ, $a_\tau m_\tau(t)\mu^{d_H(\sigma,\tau)}(1-\mu)^{L-d_H(\sigma,\tau)}$ (fractional) copies[4] of σ are produced (without normalization) where $d_H(\sigma,\tau)$ is the Hamming distance between σ and τ. This deterministic process is the expectation of the randomized copying process where each bit of τ flips independently with probability μ. This process can be described by the following coupled difference equations:

$$\forall \sigma \in \{0,1\}^L, \quad m_\sigma(t+1) \overset{\text{def}}{=} \frac{\sum_{\tau\in\{0,1\}^L} m_\tau(t) a_\tau Q^\mu_{\tau\sigma}}{\sum_{\tau\in\{0,1\}^L} m_\tau(t) a_\tau}. \tag{1}$$

Here, $Q^\mu_{\sigma\tau} \overset{\text{def}}{=} \mu^{d_H(\sigma,\tau)}(1-\mu)^{L-d_H(\sigma,\tau)} = Q^\mu_{\tau\sigma}$. If we let A be the diagonal matrix with $A_{\sigma\sigma} \overset{\text{def}}{=} a_\sigma$, then the evolutionary equation above can be succinctly expressed in the matrix form as

$$\mathbf{m}(t+1) \overset{\text{def}}{=} \frac{Q^\mu A\mathbf{m}(t)}{\|A\mathbf{m}(t)\|_1} = \frac{(Q^\mu A)^{t+1}\mathbf{m}(0)}{\|A\mathbf{m}(t)\|_1}. \tag{2}$$

If we assume $0 < \mu < 1$, then $Q^\mu A > 0$ and it follows from the Perron-Frobenius Theorem that starting from any initial distribution, the process converges to a unique limit $\mathbf{v}^\mu$, the principal right eigenvector of $Q^\mu A$. The population determined by $\mathbf{v}^\mu$ marks the culmination of the evolutionary process and is the quasispecies for the evolution described by μ, L and a.

---

[1] Sometimes called the error catastrophe.

[2] Note that RNA and DNA molecules consist of strings of four building blocks, the neucleotides A, G, C and U (RNA) or T (DNA), so the restriction to an alphabet of size two is not without loss of generality. However, as is standard in expositions related to the quasispecies model including Eigen's, results for the binary case extend to the quaternary case. Indeed, our results have analogues in the RNA/DNA setting.

[3] As is the norm, we assume the environment and fitness function are fixed throughout the evolutionary process.

[4] The population in this model is assumed to be infinite justifying fractions.

1.2. **The Error Threshold.** If μ is close to 0 then one would expect $\mathbf{v}^{\mu}$ to have its mass concentrated on sequences of high fitness but not entirely on the fittest or *master* sequence. On the other hand note that if $\mu = 1/2$, then $v^{\mu}_{\sigma} = 1/2^{L}$ for all σ. This conforms with the intuition that if each bit, during copying, can be 0 or 1 with probability $1/2$, then no genetic information can be retained by the quasispecies. Interestingly, Eigen and his coworkers observed that when the fitness landscape is single-peaked, *i.e.*, the master sequence has fitness above 1 while the rest have fitness 1, this phenomena occurs much below $\mu = 1/2$, in fact around $1/L$. This led them to hypothesize that there exists a particular mutation rate (below $1/2$) beyond which all the 2^L strings become nearly-equally abundant. They called this critical mutation rate the *error threshold* as, beyond this rate the master sequence's genetic information is lost, and called this transition an *error catastrophe*; the mutation induced loss of genetic information. Subsequently error threshold has been observed empirically for many other fitness functions, see [DSV12].

The notion of an error threshold is widely recognized, however no consensus exists on its definition; see Wilke [Wil05]. Since $\mathbf{v}^{\mu}$ will hardly ever be $\mathscr{U}$, the uniform distribution on $\{0,1\}^L$, the goal is to find the smallest μ such that $\mathbf{v}^{\mu}$ is *close* to $\mathscr{U}$. To define the error threshold one needs a function that measures closeness. Since both the uniform distribution and $\mathbf{v}^{\mu}$ are probability distributions, we can use the most prevalent way in probability to measure closeness: $\|\mathbf{v}^{\mu} - \mathscr{U}\|_1 \overset{\text{def}}{=} \sum_{\tau \in \{0,1\}^L} |v^{\mu}_{\tau} - 1/2^L|$.[5] In this paper, we define $\mu_{\star}(\varepsilon) \overset{\text{def}}{=} \min\{\mu \in (0,1) : \|\mathbf{v}^{\mu} - \mathscr{U}\|_1 \leq \varepsilon\}$. As noted before, when $\mu = 1/2$, the steady state vector $\mathbf{v}^{\mu}$ is *exactly* $\mathscr{U}$. Hence, $\mu_{\star}(\varepsilon) \leq 1/2$ for all $\varepsilon > 0$. Thus, we will focus on μ in $(0, 1/2]$.

Finally, the notion of error catastrophe is important in antiviral drug-design as several important viruses are RNA viruses and their evolution can be captured to the first order by the quasispecies model. From the virus' point of view, a high mutation rate implies greater diversity, which in turn could mean greater adaptability and greater ability to escape the hosts immune responses and pressure from drug therapy. At the same time, too high a mutation rate induces a loss of genetic information. Thus, if we could increase the mutation rate past the error threshold, we would severely compromise the virus' identity. Intriguingly, this strategy is already employed by the body which can produce antibodies that increase mutation rate. Artificially, this effect can also be accomplished by mutagenic drugs such as ribavirin, see [CCA01].

1.3. **What Makes the Quasispecies Model Hard?** Given L, a, and ε, we wish to analytically determine[6] the smallest value of μ for which the vector $\mathbf{v}^{\mu}$ comes ε-close to the uniform distribution in the ℓ_1 norm. The matrix Q^{μ} is well studied in TCS and is often referred to as the *noisy hypercube* matrix. *All* its eigenvalues and eigenvectors can be written explicitly in terms of the parameters μ and L. Additionally, A is a diagonal matrix whose eigenvectors and eigenvalues are explicit. Hence, it may be tempting to believe that the eigenvalues and eigenvectors for $Q^{\mu}A$ can be easily determined. The problem is that A is not explicit (it is given as input) and the fact that it is a diagonal matrix is superfluous (we could always make either matrix diagonal). Thus, since the eigenspaces interact in a complex manner, the largest eigenvector of $Q^{\mu}A$ does not seem to be derivable in closed form from the spectral data of Q^{μ} and A. In short, the Boolean function a, which gives the quasispecies model the ability to elegantly capture complex evolutionary interactions, turns out to be the reason this process is difficult to analyze.

1.4. **Our Contribution.** It seems difficult to analytically determine the error threshold when the fitness landscape, a, is arbitrary. Indeed, even the existence of a low error threshold in the simple-sounding, single-peak landscape was only argued heuristically by Eigen and his coworkers. We first present a class of fitness functions which generalize Eigen's single-peak model and is implicit in almost all biological settings where the quasispecies model is being deployed. First, consider how one obtains a fitness function. In applications L is at least a few thousand and, for HIV-1, about 10000. Hence, it is not possible to expect $2^L = 2^{10000}$ distinct numbers as input. Moreover, experimentally obtaining each of these numbers is a laborious and expensive process. Typically, the fitness is estimated for a number of strings which are considered to carry genetic information and the rest are assumed to be 1. Additionally, it can be argued that the fitness function has little to no dependance on L. Hence, a reasonable

---

[5] Other notions such as $\|\mathbf{v}^{\mu} - \mathscr{U}\|_2, \|\mathbf{v}^{\mu} - \mathscr{U}\|_{\infty}$ or the difference in Shannon entropies of $\mathbf{v}^{\mu}$ and $\mathscr{U}$, can also be studied and $\mu_{\star}$ may change accordingly.

[6] Often computing $\mu_{\star}(\varepsilon)$ suffices, see e.g. [DSV12, TBVD12]. However, analytical answers are useful as they provide the ability to ask if-then questions and understand for what parameters the error threshold occurs much below $1/2$.

model is one where at most k strings have fitness $1 < a_\sigma \le c$ and the rest have fitness 1 and where k, c are relatively small compared to 2^L. This leads us to the following definition.

Definition 1.1 ((c,k)-finite and bounded landscape)**.** *Let c and k be positive integers. A (c,k)-finite and bounded landscape is a fitness function $a : \{0,1\}^L \mapsto \mathbb{Z}_{\ge 1}$ such that there exist $\sigma_1, \dots, \sigma_k$ with $1 < a_{\sigma_i} \le c$ and for all $\sigma \notin \{\sigma_1, \dots, \sigma_k\}$, $a_\sigma = 1$.*

Note that when $k = 1$ this is equivalent to Eigen's single-peak landscape model. For (c,k)-finite and bounded landscapes we prove the following theorem.

Theorem 1.2 (Main Theorem)**.** *For a given L, (c,k)-finite and bounded fitness landscape and small enough $\varepsilon > 0$,*

$$\Omega\left(\frac{1}{L}\ln\frac{c}{\varepsilon}\right) \le \mu_\star(\varepsilon) \le O\left(\frac{1}{L}\ln\frac{ckL}{\varepsilon}\right).$$

Thus, consistent with *in vitro* and *in silico* observations, a sharp error threshold exists for (c,k)-finite and bounded landscapes much below $1/2$. As one might expect, the sharpness of the error threshold deteriorates as c and k increase. One way to interpret ℓ_1 closeness is the following. If the probability vector $\mathbf{v}^\mu$ is ε close to the uniform distribution in the ℓ_1 norm, then the number of indices σ, for which $|v^\mu_\sigma - 1/2^L| \ge \varepsilon/C$ is at most C. A quantitatively similar threshold can also be shown for closeness in ℓ_2 and average fitness, and is implicit in our work. The average fitness, defined to be $\sum_\sigma v^\mu_\sigma a_\sigma$, is always at least 1, and an error catastrophe is said to occur if it drops below $1 + \varepsilon$. Note that both ℓ_2 and average are weaker notions of convergence: It can happen that $\mathbf{v}^\mu$ normalized to have ℓ_2 norm 1 is ε close to the corresponding uniform distribution but its distance in ℓ_1 norm is $\varepsilon \cdot 2^{L/2}$. Similarly, the largest eigenvalue being at most $1 + \varepsilon$ (it has to be at least 1) gives a bound on ℓ_1 closeness which is not useful as such.

1.5. **Open Problems.** While we present the first formal proof of existence of sharp error threshold for a class fitness functions motivated by practice, several questions remain open. Natural questions are to further tighten our theorem and find more general and useful fitness functions for which such results can be proved. Another question is if $\|\mathbf{v}^\mu - \mathscr{U}\|_1$ converges monotonically to 0 as μ goes to $1/2$. This question arises when one tries to compute the error threshold by incrementing μ discretely. Our theorem implies that the distance goes to 0 at a certain rate for (c,k)-finite and bounded landscapes, but does not establish monotonicity.

The next question concerns the spectral ratio of the $Q^\mu A$ matrix. The spectral ratio determines the speed of evolution the quasispecies model, and hence is an important parameter from a biological perspective.[7] First notice that for positive fitness function a, and $0 < \mu < 1/2$, all eigenvalues of $Q^\mu A$ are positive real numbers. Ordering these eigenvalues as $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_{2^L}$, we call the ratio λ_2/λ_1 the spectral ratio of QA. It can be shown by standard spectral techniques that the spectral ratio of $Q^\mu A$ captures the convergence properties of the quasispecies model; at time t, the distance $\|\mathbf{m}(t) - \mathbf{v}^\mu\|_1$ roughly goes down geometrically with respect to the spectral ratio. One tempting conjecture is that for $\mu \in (0, 1/2)$, the spectral ratio is at most $1 - 2\mu$, that of the noisy hypercube. This is false in general but can be shown easily when the fitness function is a tensor. Our techniques to prove Theorem 1.2, along with the rank-one update formulas for eigenvalues can be used to establish near-tight bound on the spectral ratio for the single-peak model. We omit the details. Can we establish upper bounds on the spectral ratio for the more general (c,k)-finite and bounded landscapes?

1.6. **Technical Overview.** The proof of Theorem 1.2 combines simple tools from linear algebra and Fourier analysis of Boolean functions. Note that the fitness function a can be thought of as a Boolean function since its domain is $\{0,1\}^L$. It is useful to go back and forth between strings $\sigma \in \{0,1\}^L$ and subsets S of $[L]$. The matrix Q^μ is the noisy hypercube matrix whose eigenvectors are the characters of the hypercube. A character $\chi_S : \{0,1\}^L \mapsto \{-1,1\}$ corresponding to $S \subseteq [L]$ is $\chi_S(\sigma) \stackrel{\text{def}}{=} (-1)^{\sum_{i \in S} \sigma_i}$. These characters are orthogonal and can be normalized to have ℓ_2 norm 1. Let U be the matrix whose columns are these orthonormal characters. Thus, Q^μ can be written as $UDU^\top$ where D is a diagonal matrix with $D_{S,S} \stackrel{\text{def}}{=} (1-2\mu)^{|S|}$. Hence, one can rewrite $Q^\mu A = UDU^\top A$ which is similar to $DU^\top AU$ since $U^\top U = I$. Note that $U^\top AU$ is a rewriting of A in the Fourier basis and turns out to have a nice form: Its diagonal entries are $1 + \alpha$ for some positive α and the off-diagonal entries are bounded in absolute value

---

[7] When modeling the effect of a mutagenic drug under the quasispecies model, the convergence rate determines the minimum required duration of treatment.

by α. If A corresponds to a (c,k)-finite and bounded landscape, then we show that α is roughly $\frac{ck}{2^L}$. Very roughly, the spectrum of $Q^\mu A$ behaves like $D(I+\alpha J)$ which captures the effect of A. (See Lemma 2.4. Also note that if all strings have the same fitness, $k=0$ and $\alpha=0$ gives us back D.) More precisely, since $Q^\mu A$ and, hence, $DU^\top AU$ is not symmetric, we work with its symmetrization $W \overset{\text{def}}{=} D^{1/2}U^\top AUD^{1/2}$. What is the relation between eigenvalues and eigenvectors of $Q^\mu A$ and W? It is easy to see that if $\mathbf{w}$ is an eigenvector of W with eigenvalue λ then $UD^{1/2}\mathbf{w}$ is an eigenvector of $Q^\mu A$ with the same eigenvalue.[8] This connection allows us to show that the largest eigenvalue of $Q^\mu A$ is at most $1+\alpha(1-\mu)^L$, α as before and the smallest at least about $c(1-\mu)^L$ (see Corollary 2.7, Lemma 2.8 and Theorem 2.5). The fact that the eigenvalues of the noisy hypercube drop geometrically is crucial in the proofs. Thus, when μ is above $\ln ck/L$, the largest eigenvalue comes close to 1. Note that the largest eigenvalue is always at least 1. Since the largest eigenvalue captures the average fitness (see Lemma 2.3) of the quasispecies, it being close to 1 means, in one way, that the quasispecies has lost its genetic information.

Even if the eigenvalue is close to 1 that does not necessarily imply that the eigenvector $\mathbf{v}^\mu$ is close to the uniform vector in ℓ_1 norm as required by Theorem 1.2. Going from the largest eigenvector of W to that of $Q^\mu A$ we must worry about the norms due to the matrix D. One key observation here is that because $\mathbf{v}^\mu$ is non-negative, if $\mathbf{w}$ were the largest eigenvector of W with ℓ_2 norm 1, then the ℓ_1 norm of the corresponding eigenvector, $\mathbf{v} \overset{\text{def}}{=} UD^{1/2}\mathbf{w}$ of $Q^\mu A$ is *exactly* $2^{L/2}w_\emptyset$. (See Lemma 2.9.) This allows us to show that the closeness of $\mathbf{v}^\mu = \frac{\mathbf{v}}{\|\mathbf{v}\|_1}$ to the uniform vector is governed by $w_\emptyset$ and the relation between the ℓ_1 and ℓ_2 norm of $\mathbf{v}$. In particular it can be shown that if $w_\emptyset \geq 1-\gamma$, for some small enough γ, then the ℓ_1 and ℓ_2 norm of this are close to $2^{L/2}$ and 1 respectively, see Corollary 2.12. Thus, for such a vector Cauchy-Schwarz is nearly tight and the vector has to be almost parallel to the uniform vector. This, along with a lower bound on $w_\emptyset$ (see Lemma 2.16) turns out to be sufficient to establish an upper bound on μ for $\mathbf{v}^\mu$ to be ε-close to the uniform vector in ℓ_1. The lower bound on the error threshold follows more easily via eigenvalue considerations. The formal proof appears in the next section.

2. PROOFS

We start with some preliminaries with basic notations in Section 2.1.1, basics about diagonalizing Q^μ in the Fourier basis in Section 2.1.2 and some basic facts about the matrix $Q^\mu A$ in Section 2.1.3. Section 2.2 contains the proof of Theorem 1.2.

2.1. Preliminaries.

2.1.1. *Notation.* As above, vectors are denoted in bold and are indexed by the elements of Boolean hypercube $\{0,1\}^L$. In the sections below we also index by subsets $S \subseteq [L] \overset{\text{def}}{=} \{1,\ldots,L\}$ for convenience. The ℓ_1 norm of a vector $\mathbf{v}$ is denoted $\|\mathbf{v}\|_1 \overset{\text{def}}{=} \sum_\sigma |v_\sigma|$ and the ℓ_2 norm $\|\mathbf{v}\|_2 \overset{\text{def}}{=} \sqrt{\sum_\sigma v_\sigma^2}$. Let $\mathscr{U}_1$ denote the uniform probability vector over $\{0,1\}^L$, i.e., the vector with each entry set to $1/2^L$, and let $\mathscr{U}_2$ be the vector that has every entry set to $1/2^{L/2}$.

2.1.2. *The Matrix Q^μ and its Spectrum.* As discussed in the introduction, it suffices to consider $0<\mu\leq 1/2$. Recall that the matrix Q^μ is defined such that $Q^\mu_{\sigma\tau} = \mu^{d_H(\sigma,\tau)}(1-\mu)^{L-d_H(\sigma,\tau)}$ where $d_H(\cdot,\cdot)$ measures the Hamming distance between σ and τ. Note that Q^μ is symmetric and, for μ as above, is positive semi-definite. For $S\subseteq[L]$, define $\chi_S(\sigma) \overset{\text{def}}{=} (-1)^{\sum_{i\in S}\sigma_i}$. Let $\hat{\chi}_S \overset{\text{def}}{=} \frac{1}{2^{L/2}}\chi_S$. Then it follows that $\langle\hat{\chi}_S,\hat{\chi}_T\rangle = 1$ if $S=T$ and 0 otherwise. Let U be the matrix whose columns are $\hat{\chi}_S$; i.e., $U_{\sigma,S} \overset{\text{def}}{=} \hat{\chi}_S(\sigma)$. Then, from the orthonormality of the set $\{\hat{\chi}_S\}_{S\subseteq[L]}$ it follows that U is a unitary matrix; i.e., $UU^\top = I$. Moreover, it is well known that the eigenvectors of Q^μ are columns of U. This gives the following proposition.

Proposition 2.1. *$Q^\mu = UDU^\top$, where $U_{\sigma,S} = \hat{\chi}_S(\sigma)$ and D is a diagonal matrix with $D_{S,S} = (1-2\mu)^{|S|}$.*

---

[8] Note that the norm of this latter vector need not be 1 in ℓ_1 or ℓ_2 even if the ℓ_2 norm of $\mathbf{w}$ was one.

2.1.3. *The Matrix $Q^\mu A$.* While both Q^μ and A are symmetric their product is, in general, not. However, for $0 < \mu \le 1/2$, the matrix Q^μ is positive semi-definite (PSD) and since A is diagonal and positive $Q^\mu A$ is strictly positive. Hence, the Perron Frobenius Theorem implies that its right eigenvector that corresponds to its largest eigenvalue is strictly positive and unique. We record this observation here.

Proposition 2.2. *For $0 < \mu < 1/2$ and diagonal A with strictly positive entries, the dominant right eigenvector, $\mathbf{v}^\mu$ of $Q^\mu A$ is strictly positive and unique. Moreover, for any probability vector $\mathbf{m}$, $\lim_{t\to\infty}(Q^\mu A)^t \mathbf{m} = \mathbf{v}^\mu$.*

Henceforth, for the sake of readability, we will denote Q^μ by Q. The following lemma gives a characterization of the largest eigenvalue of QA.

Lemma 2.3. *If λ_1 is the largest right eigenvalue of QA with eigenvector $\mathbf{v}$ such that $\|\mathbf{v}\|_1 = 1$, then $\lambda_1 = \sum_\sigma a_\sigma v_\sigma$.*

Hence, the largest eigenvalue measures the average fitness of the corresponding eigenvector. The proof is straightforward.

Proof. By definition, $QA\mathbf{v} = \lambda_1 \mathbf{v}$. Hence,

$$\lambda_1 = \lambda_1 \sum_\sigma v_\sigma = \sum_\sigma \sum_\tau Q_{\tau\sigma} a_\tau v_\tau = \sum_\tau a_\tau v_\tau \sum_\sigma Q_{\tau\sigma} = \sum_\tau a_\tau v_\tau.$$

□

2.2. **Error Threshold in Finite and Bounded Fitness Landscapes.** This section contains the proof of Theorem 1.2. It is a sequence of simple results and we break it down into three parts. The reader can refer to Section 1.6 to recall the flow of the proof. Section 2.2.1 is key to the following section and shows how to write the matrix A in the Fourier basis. Section 2.2.2 then uses the results from the previous sections and Section 2.2.1 to establish a near-tight bound on the largest eigenvalue of $Q^\mu A$. Section 2.2.3 argues about the dominant eigenvector of $Q^\mu A$ and establishes constraints on μ such that $\mathbf{v}^\mu$ comes close to the uniform distribution and, thus, proves the main theorem. This section relies on the Fourier connection and requires estimating the ℓ_1 and ℓ_2 norm of vector the vector $UD^{1/2}\mathbf{w}$, which is parallel to $\mathbf{v}^\mu$ and where $\mathbf{w}$ is the largest eigenvector in the Fourier domain.

2.2.1. *A Structural Lemma.* Let a be a (c,k)-finite and bounded fitness landscape where $\sigma_1,\ldots,\sigma_k$ are such that $1 < a_{\sigma_i} \le c$ and for all $\sigma \notin \{\sigma_1,\ldots,\sigma_k\}$, $a_\sigma = 1$. Proposition 2.1 implies that the spectrum of QA is the same as that of $DU^\top AU$, where U is the matrix whose columns are $\hat{\chi}_S$; i.e., $U_{\sigma,S} = \hat{\chi}_S(\sigma)$ and $\hat{\chi}_S = \frac{1}{2^{L/2}}\chi_S$. Note that

$$(U^\top AU)_{S,T} = \sum_\sigma \hat{\chi}_S(\sigma) a_\sigma \hat{\chi}_T(\sigma) = \frac{1}{2^L}\left(\sum_{i=1}^k \chi_{S\Delta T}(\sigma_i)a_{\sigma_i} + \sum_{\sigma\notin\{\sigma_1,\ldots,\sigma_k\}} \chi_{S\Delta T}(\sigma)\right).$$

If $S = T$, then

$$(U^\top AU)_{S,S} = \frac{1}{2^L}\sum_\sigma a_\sigma = \frac{2^L - k}{2^L} + \frac{\sum_{i=1}^k a_{\sigma_i}}{2^L} = 1 + \frac{\sum_{i=1}^k (a_{\sigma_i} - 1)}{2^L} \le 1 + \frac{(c-1)k}{2^L}.$$

Note that the above also implies that $(U^\top AU)_{S,S} \ge 1$. On the other hand, when $S \ne T$, $(U^\top AU)_{S,T}$ is

$$\frac{1}{2^L}\left(\sum_{i=1}^k \chi_{S\Delta T}(\sigma_i)a_{\sigma_i} + \sum_{\sigma\notin\{\sigma_1,\ldots,\sigma_T\}} \chi_{S\Delta T}(\sigma)\right) = \frac{1}{2^L}\left(\sum_{i=1}^k \chi_{S\Delta T}(\sigma_i)a_{\sigma_i} - \sum_{i=1}^k \chi_{S\Delta T}(\sigma_i)\right) = \frac{\sum_{i=1}^k \chi_{S\Delta T}(\sigma_i)(a_{\sigma_i} - 1)}{2^L}.$$

Here we have used the fact that for any $V \ne \emptyset$, $\sum_\sigma \chi_V(\sigma) = 0$. Hence, when $S \ne T$,

$$\left|(U^\top AU)_{S,T}\right| \le \frac{(c-1)k}{2^L}.$$

We record this in the following lemma.

Lemma 2.4. *For the matrix A corresponding to a (c,k)-finite and bounded landscape,*

$$\forall S \subseteq [L], \quad (U^\top AU)_{S,S} = 1 + \frac{\sum_{i=1}^k (a_{\sigma_i} - 1)}{2^L} \le 1 + \frac{(c-1)k}{2^L}$$

and

$$\forall S \neq T \subseteq [L], \quad (U^\top AU)_{S,T} = \frac{\sum_{i=1}^{k} \chi_{S\Delta T}(\sigma_i)(a_{\sigma_i}-1)}{2^L} \text{ and } \left|(U^\top AU)_{S,T}\right| \leq \frac{(c-1)k}{2^L}.$$

2.2.2. *Largest Eigenvalue.* In this section we prove the following theorem.

Theorem 2.5. *Let A correspond to a (c,k)-finite and bounded fitness landscape and $0 < \mu \leq 1/2$, if λ_1 is the largest (right) eigenvalue of QA, then*

$$\max\left\{1, \frac{(1-2\mu+2\mu^2)^L}{(1-\mu)^L} + (c-1)(1-\mu)^L\right\} \leq \lambda_1 \leq 1 + (c-1)k(1-\mu)^L.$$

We will prove this theorem in two parts. First we establish an upper bound. In the process we upper bound the distance of the largest eigenvector to the uniform vector; this will be useful for the error threshold calculation. We consider the symmetric matrix $D^{-1/2}DU^\top AUD^{1/2}$ which has the same eigenvalues as $DU^\top AU$ which, by Proposition 2.1, has the same spectrum as that of QA.

Lemma 2.6. *Let $\mathbf{w} = (w_S)_{S\subseteq[L]}$ be a vector such that $\sum_S w_S^2 = 1$. Then,*

$$\mathbf{w}^\top D^{1/2}U^\top AUD^{1/2}\mathbf{w} \leq 1 + (c-1)k(1-\mu)^L.$$

Proof. It follows from Lemma 2.4 that if $\mathbf{w} = (w_S)_{S\subseteq[L]}$, then

$$\begin{aligned}
\mathbf{w}^\top D^{1/2}U^\top AUD^{1/2}\mathbf{w} &= \sum_S w_S^2(1-2\mu)^{|S|} + \frac{1}{2^L}\sum_{i=1}^{k}(a_{\sigma_i}-1)\sum_{S,T} w_S w_T (1-2\mu)^{|S|/2}(1-2\mu)^{|T|/2}\chi_S(\sigma_i)\chi_T(\sigma_i) \\
&= \sum_S w_S^2 + \frac{1}{2^L}\sum_{i=1}^{k}(a_{\sigma_i}-1)\left(\sum_S w_S\chi_S(\sigma_i)(1-2\mu)^{|S|/2}\right)^2 \\
&\leq 1 + \frac{1}{2^L}\sum_{i=1}^{k}(a_{\sigma_i}-1)\big(\sum_S w_S^2\chi_S^2(\sigma_i)\big)\sum_S(1-2\mu)^{|S|} \\
&= 1 + \frac{1}{2^L}\sum_{i=1}^{k}(a_{\sigma_i}-1)\sum_{j=0}^{L}\binom{L}{j}(1-2\mu)^j = 1 + \frac{1}{2^L}\sum_{i=1}^{k}(a_{\sigma_i}-1)(1+1-2\mu)^L \\
&= 1 + \sum_{i=1}^{k}(a_{\sigma_i}-1)(1-\mu)^L \leq 1 + (c-1)k(1-\mu)^L.
\end{aligned}$$

□

Thus, we have the following corollary.

Corollary 2.7 (Upper Bound)**.** *For a (c,k)-finite and bounded fitness landscape, the largest eigenvalue of QA is at most $1+(c-1)k(1-\mu)^L$.*

The following lemma gives a lower bound which is comparable to the upper bound except for the factor of k.

Lemma 2.8 (Lower Bound)**.** *For a (c,k)-finite and bounded fitness landscape, the largest eigenvalue of QA is at least $\max\{1, \frac{(1-2\mu+2\mu^2)^L}{(1-\mu)^L} + (c-1)(1-\mu)^L\}$.*

Proof. First let $\mathbf{w}$ be a vector such that $w_\emptyset = 1$ and $w_S = 0$ for all $S \neq \emptyset$. The Rayleigh quotient of this vector is 1. This proves that the largest eigenvalue, which maximizes the Rayleigh quotient, has to be at least 1.

To show the other term in the lower bound, without loss of generality, let σ_1 be the string such that $a_{\sigma_1} = c$. Consider the vector $\mathbf{w}$ such that $w_S \overset{\text{def}}{=} \chi_S(\sigma_1)(1-2\mu)^{|S|/2}$. The norm of this vector is $\sum_S(1-2\mu)^{|S|} = 2^L(1-\mu)^L$. The corresponding quadratic form, again using Lemma 2.4, is

$$\mathbf{w}^\top D^{1/2}U^\top AUD^{1/2}\mathbf{w} = \sum_S(1-2\mu)^{2|S|} + \frac{1}{2^L}\sum_{i=1}^{k}(a_{\sigma_i}-1)\left(\sum_S \chi_S(\sigma_1)(1-2\mu)^{|S|}\chi_S(\sigma_i)\right)^2$$

$$\geq (1+(1-2\mu)^2)^L + \frac{1}{2^L}(c-1)(2-2\mu)^{2L}.$$

Dividing by the squared-norm of $\mathbf{w}$, which is $(2-2\mu)^L$, we get the lemma by noticing that the largest eigenvector is the maximizer of the Rayleigh quotient. $\square$

Theorem 2.5 follows from Corollary 2.7 and 2.8.

2.2.3. *The Error Threshold- Proof of Theorem 1.2.* We now head towards establishing bounds on μ for the error threshold and prove Theorem 1.2. First let us recall the formal relation between eigenvectors of QA and eigenvectors of $D^{1/2}U^\top AUD^{1/2}$: If $\mathbf{w}$ is an eigenvector of the latter with eigenvalue λ, then $UD^{1/2}\mathbf{w}$ is an eigenvector of QA with the same eigenvalue. This is because if $\mathbf{w}$ is such that $D^{1/2}U^\top AUD^{1/2}\mathbf{w} = \lambda\mathbf{w}$, then $UDU^\top AUD^{1/2}\mathbf{w} = \lambda UD^{1/2}\mathbf{w}$ which is the same as $QA(UD^{1/2}\mathbf{w}) = \lambda(UD^{1/2}\mathbf{w})$. If $\mathbf{w}_0 = (1,0,\ldots,0)$, then $UD^{1/2}\mathbf{w}_0 = \mathscr{U}_2$, which is the uniform vector in ℓ_2 norm. If $\mathbf{w}_1$ is the largest eigenvector of $D^{1/2}U^\top AUD^{1/2}$, then, $\mathbf{v} \stackrel{\text{def}}{=} UD^{1/2}\mathbf{w}_1$ is the largest right eigenvector of QA. By Proposition 2.2, $\mathbf{v}$ is positive. We now calculate its ℓ_1 and ℓ_2 norm. Assume that $\|\mathbf{w}_1\|_2 = 1$.

Lemma 2.9. *For any* $\mathbf{w}$ *such that* $UD^{1/2}\mathbf{w}$ *is a non-negative vector,*

$$\|UD^{1/2}\mathbf{w}\|_1 = 2^{L/2}w_\emptyset.$$

Proof. Since $UD^{1/2}\mathbf{w}$ is a non-negative vector,

$$\|UD^{1/2}\mathbf{w}\|_1 = \sum_\sigma (UD^{1/2}\mathbf{w})(\sigma) = \sum_\sigma\sum_S \frac{1}{2^{L/2}} w_S(1-2\mu)^{|S|/2}\chi_S(\sigma) = \sum_S \frac{1}{2^{L/2}} w_S(1-2\mu)^{|S|/2}\sum_\sigma \chi_S(\sigma) = 2^{L/2}w_\emptyset.$$

Here we have used the fact that $\sum_\sigma \chi_S(\sigma) = 0$ if $S \neq \emptyset$ and is 2^L if $S = \emptyset$. $\square$

Corollary 2.10. *Fix* $\mathbf{w}$ *such that* $\|\mathbf{w}\|_2 = 1$ *and* $w_\emptyset \geq 1-\gamma$*. Then for* $\mathbf{v} \stackrel{\text{def}}{=} UD^{1/2}\mathbf{w}$*, we have* $\|\mathbf{v}\|_1 \geq (1-\gamma)2^{L/2}$ *and* $\|\mathbf{v}\|_2 \geq (1-\gamma)$*.*

Proof. The first claim follows from Lemma 2.9. To lower bound $\|\mathbf{v}\|_2$, use Cauchy-Schwarz: $2^{L/2}\|\mathbf{v}\|_2 \geq \|\mathbf{v}\|_1$. $\square$

Lemma 2.11. *Fix* $\mathbf{w}$ *such that* $\|\mathbf{w}\|_2 = 1$*. Then for* $\mathbf{v} \stackrel{\text{def}}{=} UD^{1/2}\mathbf{w}$, $\|\mathbf{v}\|_2 \leq 1$

Proof. First note that $\|\mathbf{v}\|_2 = \|UD^{1/2}\mathbf{w}\|_2 = \sqrt{\mathbf{w}^\top D\mathbf{w}}$ since U is unitary. To complete the proof note that $\mathbf{w}^\top D\mathbf{w} \leq w_\emptyset^2 + (1-2\mu)(1-w_\emptyset^2) \leq 1$. $\square$

Thus, combining Corollary 2.10 and Lemma 2.11, we obtain the following corollary.

Corollary 2.12. *Fix* $\mathbf{w}$ *such that* $\|\mathbf{w}\|_2 = 1$ *and* $w_\emptyset \geq 1-\gamma$*. Let* $\mathbf{v} \stackrel{\text{def}}{=} UD^{1/2}\mathbf{w}$*. Then*

(1) $1-\gamma \leq \|\mathbf{v}\|_2 \leq 1$

(2) $(1-\gamma)2^{L/2} \leq \|\mathbf{v}\|_1 \leq 2^{L/2}$.

As a consequence, the vector $\mathbf{v}$ is almost parallel to the all 1 vector.

Lemma 2.13. *Let* $\mathbf{v} \geq 0$ *be such that* $\|\mathbf{v}\|_2 \leq 1$ *and* $(1-\gamma)2^{L/2} \leq \|\mathbf{v}\|_1$ *for a small enough constant* $\gamma > 0$*. Then*

(1) $\left\|\frac{\mathbf{v}}{\|\mathbf{v}\|_1} - \mathscr{U}_1\right\|_2 \leq O\left(\frac{\sqrt{\gamma}}{2^{L/2}}\right)$ *and*

(2) $\left\|\frac{\mathbf{v}}{\|\mathbf{v}\|_1} - \mathscr{U}_1\right\|_1 \leq O(\sqrt{\gamma})$.

Proof. $\left\|\frac{\mathbf{v}}{\|\mathbf{v}\|_1} - \mathscr{U}_1\right\|_2^2 = \sum_\sigma\left(\frac{v_\sigma}{\|\mathbf{v}\|_1} - \frac{1}{2^L}\right)^2 = \sum_\sigma \frac{v_\sigma^2}{\|\mathbf{v}\|_1^2} - \frac{2}{2^L}\sum_\sigma \frac{v_\sigma}{\|\mathbf{v}\|_1} + \frac{1}{2^L}$. Since $\mathbf{v} \geq 0$, $\|\mathbf{v}\|_1 = \sum_\sigma v_\sigma$. Hence,

$$\left\|\frac{\mathbf{v}}{\|\mathbf{v}\|_1} - \mathscr{U}_1\right\|_2^2 = \frac{\|\mathbf{v}\|_2^2}{\|\mathbf{v}\|_1^2} - \frac{1}{2^L} \leq \frac{1}{(1-\gamma)^2 2^L} - \frac{1}{2^L} = O\left(\frac{\gamma}{2^L}\right)$$

for small enough constant γ. It follows from Cauchy-Schwarz that

$$\left\|\frac{\mathbf{v}}{\|\mathbf{v}\|_1} - \mathscr{U}_1\right\|_1 \leq 2^{L/2}\left\|\frac{\mathbf{v}}{\|\mathbf{v}\|_1} - \mathscr{U}_1\right\|_2 \leq O(\sqrt{\gamma}).$$

$\square$

Lemma 2.14. *For $\delta > 0$, suppose c,k,μ and L are such that $(c-1)k(1-\mu)^L \leq \delta$. Let $\mathbf{w}_1$ be the largest eigenvector of $D^{1/2}U^\top AUD^{1/2}$ with $\|\mathbf{w}_1\|_2 = 1$ and let $w_\emptyset$ be the mass of $\mathbf{w}_1$ on the coordinate corresponding to $\emptyset$. Then, $w_\emptyset^2 \geq 1 - \frac{\delta}{2\mu}$.*

Proof. We noted in Theorem 2.5 that $\lambda_1 \geq 1$. Hence,

$$1 \leq \lambda_1 = \mathbf{w}_1^\top D^{1/2}U^\top AUD^{1/2}\mathbf{w}_1 \leq w_\emptyset^2 + (1-2\mu)(1-w_\emptyset^2) + k(c-1)(1-\mu)^L \leq 1 - 2\mu + 2\mu w_\emptyset^2 + \delta.$$

This implies that $w_\emptyset^2 \geq 1 - \frac{\delta}{2\mu}$.

☐

Hence, we can show the following which completes one half of Theorem 1.2.

Corollary 2.15 (Upper Threshold)**.** *For a small enough constant $\varepsilon > 0$ and a (c,k)-finite and bounded landscape, $\mu_\star(\varepsilon) \leq O\left(\frac{1}{L}\ln\frac{ckL}{\varepsilon^2}\right)$.*

Proof. Given ε, we pick $\gamma \overset{\text{def}}{=} O(\varepsilon^2)$, and $\delta \overset{\text{def}}{=} O(\gamma\mu)$ such that

$$(c-1)k(1-\mu)^L \leq \delta. \tag{3}$$

Then, it follows from Lemma 2.14 that if $\mathbf{w}_1$ is the largest eigenvector of $D^{1/2}U^\top AUD^{1/2}$, then, $w_\emptyset \geq w_\emptyset^2 \geq 1-\gamma$. Hence, from Corollary 2.12 it follows that for $\mathbf{v}^\mu$, which is by definition $UD^{1/2}\mathbf{w}_1$, $\|\mathbf{v}^\mu\|_1 \geq (1-\gamma)2^{L/2}$. Thus, by part (2) of Lemma 2.13, it follows that $\|\mathbf{v}^\mu - \mathscr{U}_1\|_1 \leq O(\sqrt{\gamma}) = O(\varepsilon)$. Finally, note that picking $\mu = O\left(\frac{1}{L}\ln\frac{ckL}{\varepsilon^2}\right)$ suffices to satisfy Equation (3). ☐

Lower threshold. Now we proceed to prove the other half of Theorem 1.2.

Lemma 2.16. *If the largest eigenvalue λ_1 of QA is at least $1+\delta$ for some $\delta > 0$ and $\mathbf{v} \overset{\text{def}}{=} \mathbf{v}^\mu$ is the corresponding vector with $\|\mathbf{v}\|_1 = 1$, then $\sum_{i=1}^k v_{\sigma_i} \geq \frac{\delta}{(c-1)}$. As a consequence,*

$$\|\mathbf{v} - \mathscr{U}_1\|_1 \geq \frac{\delta}{(c-1)} - \frac{k}{2^L}.$$

Proof. The largest eigenvalue $\lambda_1 = \sum_\sigma a_\sigma v_\sigma$. Hence,

$$1+\delta \leq \lambda_1 \overset{\text{Lemma 2.3}}{=} \sum_{i=1}^k a_{\sigma_i}v_{\sigma_i} + 1 - \sum_{i=1}^k v_{\sigma_i} = 1 + \sum_{i=1}^k (a_{\sigma_i}-1)v_{\sigma_i}.$$

Hence, $\sum_{i=1}^k a_{\sigma_i}v_{\sigma_i} \geq \delta$. Thus, $\sum_{i=1}^k v_{\sigma_i} \geq \frac{\delta}{(c-1)}$. The claim about the ℓ_1 distance from $\mathscr{U}_1$ follows by noticing that $\mathbf{v}$ is non-negative. ☐

Hence, we have the following corollary.

Corollary 2.17 (Lower Threshold)**.** *For a given ε and a (c,k)-finite and bounded landscape such that $ck < 2^L\varepsilon$,*

$$\mu_\star(\varepsilon) \geq \Omega\left(\frac{1}{L}\ln(c-1)\right).$$

Proof. Given $\varepsilon > 0$, we, rather slackly, we let $\delta \overset{\text{def}}{=} (c-1)\varepsilon + \frac{(c-1)k}{2^L}$. Thus, if we pick μ such that $\delta = (c-1)(1-\mu)^L - 1$, it follows from Lemma 2.16 and Lemma 2.8 that $\|\mathbf{v}^\mu - \mathscr{U}_1\|_1 \geq \varepsilon$. It is easy to check that the choice of μ claimed in the corollary satisfies the constraint on δ. ☐

To complete the proof of Theorem 1.2, we mention that by a tighter analysis, we can improve the lower bound to that claimed in the main theorem. We skip the tedious details in this version of the paper.

REFERENCES

[CCA01] Shane Crotty, Craig E. Cameron, and Raul Andino. RNA virus error catastrophe: Direct molecular test by using ribavirin. *Proceedings of the National Academy of Sciences*, 98(12):6895–6900, 2001.

[DSV12] Narendra Dixit, Piyush Srivastava, and Nisheeth K. Vishnoi. A finite population model of molecular evolution: Theory and computation. *Journal of Computational Biology*, 19(10):1176–1202, 2012.

[Eig71] M. Eigen. Selforganization of matter and the evolution of biological macromolecules. *Die Naturwissenschaften*, 58:456–523, 1971.

[Eig93] M. Eigen. The origin of genetic information: Viruses as models. *Gene*, 135:37–47, 1993.

[Eig02] M. Eigen. Error catastrophe and antiviral strategy. *Proceedings of the National Academy of Sciences of the United States of America*, 99(21):13374–13376, October 2002.

[EMS88] M. Eigen, J. McCaskill, and P. Schuster. Molecular quasi-species. *J. Phys. Chem.*, 92:6881–6891, 1988.

[EMS89] M. Eigen, J. McCaskill, and P. Schuster. The molecular quasi-species. *Adv. Chem. Phys.*, 75:149–263, 1989.

[ES77] M. Eigen and P. Schuster. The hypercycle, a principle of natural self-organization. part a: Emergence of the hypercycle. *Die Naturwissenschaften*, 64:541–565, 1977.

[HD98] John Holland and Esteban Domingo. Origin and evolution of viruses. *Virus Genes*, 16:13–21, 1998.

[LEK$^+$99] L A Loeb, J M Essigmann, F Kazazi, J Zhang, K D Rose, and J I Mullins. Lethal mutagenesis of hiv with mutagenic nucleoside analogs. *Proc Natl Acad Sci U S A*, 96(4):1492–1497, February 1999.

[SS82] J. Swetina and P. Schuster. Self-replication with errors: a model for polynucleotide replication. *Biophys. Chem.*, 16:329–345, 1982.

[TBVD12] Kushal Tripathi, Rajesh Balagam, Nisheeth K. Vishnoi, and Narendra M. Dixit. Stochastic simulations suggest that HIV-1 survives close to its error threshold. *PLoS Computational Biology*, 8(9):e1002684, 2012.

[Vis12] Nisheeth K. Vishnoi. Evolution without sex, drugs and Boolean functions. Expository article available at `http://research.microsoft.com/en-us/um/people/nvishno/Site/Publications.html`, 2012.

[Wil05] Claus Wilke. Quasispecies theory in the context of population genetics. *BMC Evolutionary Biology*, 5(1):44, 2005.