Martingale Families and Dimension in P

Philippe Moser *

Abstract

We introduce a new measure notion on small complexity classes (called F-measure), based on martingale families, that gets rid of some drawbacks of previous measure notions: it can be used to define dimension because martingale families can make money on all strings, and it yields random sequences with an equal frequency of 0's and 1's. As applications to F-measure, we answer a question raised in [1] by improving their result to: for almost every language A decidable in subexponential time, $\mathsf{P}^A = \mathsf{BPP}^A$. We show that almost all languages in PSPACE do not have small non-uniform complexity. We compare F-measure to previous notions and prove that martingale families are strictly stronger than Γ-measure [1], we also discuss the limitations of martingale families concerning finite unions. We observe that all classes closed under polynomial many-one reductions have measure zero in EXP iff they have measure zero in SUBEXP. We use martingale families to introduce a natural generalization of Lutz resource-bounded dimension [13] on P, which meets the intuition behind Lutz's notion. We show that P-dimension lies between finite-state dimension and dimension on E. We prove an analogue to the Theorem of Eggleston in P, i.e. the class of languages whose characteristic sequence contains 1's with frequency α, has dimension the Shannon entropy of α in P.

1 Introduction

Resource-bounded measure has been successfully used to understand the structure of the exponential time classes E and EXP, see [12] for a survey. Recently resource-bounded measure has been refined via effective dimension which is an effectivization of Hausdorff dimension, yielding applications in a variety of topics, including algorithmic information theory, computational complexity, prediction, and data compression [13, 17, 14, 4, 2, 6].

Unfortunately both Lutz's resource-bounded measure and dimension formulations [10, 13] only work on classes containing E (apart from finite-state dimension). One reason for this is that when a martingale is to bet on some string x depending on the history of the language for strings $y < x$, the history itself is exponentially larger than the string x. Thus even reading the history is far above the computational power of P.

One way to overcome this difficulty was proposed in [1], with a measure notion (called Γ-measure) defined via martingales betting only on a sparse subset of strings of the history, with the drawback that the class of sparse languages does not have measure zero. Nevertheless it seems that sparse languages and more generally languages whose characteristic sequences satisfy some frequency property should be small for an appropriate measure notion on P, because there exists simple (exponential-time computable) martingales always making the

*Email: mosersan@gmail.com. This research was supported in part by Swiss National Science Foundation Grant PB GE2–104820.

same fixed bet that succeed on such languages. Such martingales are relatively "simple": exponential computational power is only required to keep track of the current capital. This also shows how important it is for a martingale to be able to bet on all strings, in order to succeed. This "betting on all strings" property becomes crucial in Lutz's recent formulation of effective Hausdorff dimension [13].

A stronger measure notion called dense martingale measure (denoted Γ_d) was then proposed in [23], with the surprising result that the polynomial time version of Lutz's hypothesis "NP does not have measure zero in E" does *not* hold [3]. Γ_d-measure does not satisfy the finite union property though; it was then showed that a restricted version (denoted $\Gamma/(\mathsf{P})$) of it does, unfortunately $\Gamma/(\mathsf{P})$-measure has some unnatural properties: a language with infinitely many easy instances can still be random.

Another limitation of previous martingale-based measure notions on P from [1, 23] and on PSPACE [18] is the inability of the corresponding martingales to bet on all strings. Γ-martingales can only bet on a polynomial number out of the exponentially many strings of length n, whereas Γ_d and $\Gamma/(\mathsf{P})$ martingales can only double their capital a polynomial number of times while betting on (the exponentially many) strings of size n, with the direct consequence that neither can be used to define a dimension notion, because the ability to bet on every string is essential for this purpose (notice that simply keeping track of the capital won by a martingale doubling its capital on every string is impossible in polynomial time). Moreover the random sequences yielded by either of those two measure notions do not have an equal frequency of 0's and 1's in the limit, whereas this property is captured by Lutz's resource-bounded measure notion on E, corresponding to the intuitive idea of a random sequence.

In this paper we introduce a measure notion on P based on martingale families (called F-measure), where martingale families can double their capital on all strings, thus enabling us to define dimension in P. F-measure gets rid of the unnatural random sequences of $\Gamma/(\mathsf{P})$-measure [23], and yields random sequences with an equal frequency of 0's and 1's, similarly to Lutz resource-bounded measure [10]. Moreover F-measure is strictly stronger than Γ-measure. *United, we stand; divided, we fall* is the key idea behind F-measure, i.e. whereas a single polynomial time computable martingale is not able to make money on all exponentially many strings of size n, a family of martingales working together and sharing their capital *can*. The idea is to separate the exponentially many strings of size n into groups of polynomial size, where each member of the family bets on one of these groups of strings. The family shares a common bank account: When such a martingale bets on a string x, the capital at its disposal amounts to the capital currently gathered by its family on predecessors of x, although it has no information about how much this (possibly) exponentially large capital is.

Constructing a perfect measure on P has turned out to be much more difficult as previously thought; it is now widely admitted that the perfect measure on P might be very difficult to achieve, and that for any measure notion on P some desirable properties must be abandoned; and F-measure is no exception. Similarly to Γ_d-measure [23], martingale families do not satisfy the finite union property, but only satisfy the union property in some non-general sense: we can only guarantee the union property for families with the same bank account structure; however this is usually enough to prove theorems where the union property is needed.

We show in Section 3.1 that except for general unions, martingale families satisfy the basic measure properties, i.e. every single language has measure zero, and the whole class P does not have measure zero, we then introduce uniform P-unions and show that the union

property holds for those. We observe that it is easy to derive a F-measure notion on classes between P and E like $\mathsf{QUASIPOLY}$, SUBEXP and PSPACE; for BPP see [21].

Next we show that the concept of randomness yielded by F-measure is optimal regarding frequency: every language L such that there are infinitely many n with $|L[1 \cdots n]| \leq \epsilon n$ (with $\epsilon < 1/2$), has measure zero in P (Section 3.2).

As applications to F-measure, we answer a question raised in [1], improving their result to: almost all (all except a measure zero class) languages computable in subexponential time, are hard enough to derandomize BPP, i.e. a polynomial time algorithm can use almost every language $L \in \mathsf{SUBEXP}$ to derandomize every probabilistic polynomial time algorithm, even if the probabilistic algorithm has also oracle access to L.

We also investigate the nonuniform complexity of languages of PSPACE, and show that almost all languages in PSPACE do not have small nonuniform complexity, thus reducing the resource-bounds of a similar result in [11].

Next we compare F-measure to previous measure notions on P, and show that F-measure is strictly stronger than Γ-measure, i.e. every Γ-measure zero set has F-measure zero, and there are classes with Γ-measure non-zero that have F-measure zero. Due to their intrinsic differences, we cannot compare Γ_d-measure and $\Gamma/(\mathsf{P})$-measure [23] to F-measure. Nevertheless all sets proved to be small for $\Gamma/(\mathsf{P})$-measure in [23] are also small for F-measure. Regarding density arguments, F-measure performs better; indeed a (Lebesgue) random language has $(1/2 - o(1))2^n$ words of length n (with high probability), and this property is captured by F-measure, whereas for $\Gamma/(\mathsf{P})$-measure, the set of languages having $o(2^n)$ words of length n has $\Gamma/(\mathsf{P})$-measure zero. The advantage of $\Gamma/(\mathsf{P})$-measure over F-measure is that it satisfies the finite union property. Concerning Γ_d-measure and F-measure, both their respective strengths are different, whereas Γ_d-measure cannot be used to define dimension in P, F-measure fails to capture the Γ_d-measure zero sets in [3].

We also show that all classes closed under polynomial many-one reductions have measure zero in EXP iff they have F-measure zero in E_α, which reduces the time bounds of many results [8, 19, 8, 7] from measure on E to measure on SUBEXP.

The second part of the paper is devoted to dimension in P. Lutz resource-bounded dimension [13], has been introduced on a wide variety of complexity classes ranging from finite state automata, exponential time and space up to the class of recursively enumerable languages [17], with the exception of small classes like P.

Hausdorff dimension is a refinement of Lebesgue measure, where every measure zero class of languages is assigned a real number between 0 and 1, called its Hausdorff dimension. The key idea of Lutz is to receive a tax after each round (even if the martingale did not bet during that round): the largest tax rate which can be received without preventing the martingale from succeeding on a given class represents the dimension of the class.

Trying to bridge the gap between finite state automata and exponential time requires a measure notion which is able to bet and double the capital at every round. Whereas all previous measure notions on P [1, 23] are unable to do so, it is not a problem for martingale families. This leads to a natural generalization of Lutz resource-bounded dimension [13] on P, which meets the idea behind Lutz's notion.

We give some evidence that P-dimension is a natural extension to P of previously existing dimension notions, by showing that it lies exactly between finite-state dimension and dimension on E, i.e. we show that for any sequence S, $\dim_{\mathrm{FS}}(S) \geq \dim_{\mathsf{P}}(S) \geq \dim_{\mathsf{E}}(S)$.

Finally we prove an analogue to the Theorem of Eggleston [5] in P, i.e. the class of languages whose characteristic sequences contain 1's with frequency α, has strong dimension

the Shannon entropy of α in P.

Due to space constraints, all proofs are postponed to the Appendix.

2 Preliminaries

Let us fix some notations for strings and languages. A *string* is an element of $\{0,1\}^n$ for some integer n. For a string x, its length is denoted by $|x|$. $s_0, s_1, s_2 \ldots$ denotes the standard enumeration of the strings in $\{0,1\}^*$ in lexicographical order, where $s_0 = \lambda$ denotes the empty string. Note that $|w| = 2^{O(|s_{|w|}|)}$. If x, y are strings, we write $x \leq y$ if $|x| < |y|$ or $|x| = |y|$ and x precedes y in alphabetical order. A *sequence* is an element of $\{0,1\}^\infty$. If F is a string or a sequence and $1 \leq i \leq |w|$ then $w[i]$ and $w[s_i]$ denotes the ith bit of F. Similarly $w[i \ldots j]$ and $w[s_i \ldots s_j]$ denote the ith through jth bits.

For two string x, y, the concatenation of x and y is denoted xy. If x is a string and y is a string or a sequence extending x i.e. $y = xu$, where u is a string or a sequence, we write $x \sqsubseteq y$. We write $x \sqsubset y$ if $x \sqsubseteq y$ and $x \neq y$.

A *language* is a set of strings. A *class* is a set of languages. The cardinal of a language L is denoted $|L|$. Let n be any integer. The set of strings of size n of language L is denoted $L^{=n}$. Similarly $L^{\leq n}$ denotes the set of strings in L of size at most n.

A language L is said to be polynomially dense if there exists a polynomial p, such that $|L^{=n}| \geq 2^n/p(n)$. We identify language L with its characteristic function χ_L, where χ_L is the sequence such that $\chi_L[i] = 1$ iff $s_i \in L$. Thus a language can be seen as a sequence in $\{0,1\}^\infty$. $L \upharpoonright s_n$ denotes the initial segment of L up to s_n given by $L[s_0 \cdots s_n]$.

We use standard notation for traditional complexity classes; see for instance [22]. For $\epsilon > 0$, denote by E_ϵ the class $\mathsf{E}_\epsilon = \bigcup_{\delta<\epsilon} \mathsf{DTIME}(2^{n^\delta})$. SUBEXP is the class $\cap_{\epsilon>0}\mathsf{E}_\epsilon$, and quasi polynomial time refers to the class $\mathsf{QUASIPOLY} = \cup_{k\geq 1}\mathsf{DTIME}(n^{\log^k n})$.

2.1 Martingales

Lutz measure on E [11] is obtained by imposing appropriate resource-bounds on a game theoretical characterization of classical Lebesgue measure, via martingales. A martingale is a function $d : \{0,1\}^* \to \mathbb{R}_+$ such that, for every $w \in \{0,1\}^*$, $2d(w) = d(w0) + d(w1)$ This definition can be motivated by the following betting game in which a gambler puts bets on the successive membership bits of a hidden language A. The game proceeds in infinitely many rounds where at the end of round n, it is revealed to the gambler whether $s_n \in A$ or not. The game starts with capital 1. Then, in round n, depending on the first $n-1$ outcomes $w = \chi_A[0 \ldots n-1]$, the gambler bets a certain fraction $\epsilon_w d(w)$ of his current capital $d(w)$, that the nth word $s_n \in A$, and bets the remaining capital $(1-\epsilon_w)d(w)$ on the complementary event $s_n \notin A$. The game is fair, i.e. the amount put on the correct event is doubled, the one put on the wrong guess is lost. The value of $d(w)$, where $w = \chi_A[0 \ldots n]$ equals the capital of the gambler after round n on language A. The player wins on a language A if he manages to make his capital arbitrarily large during the game, i.e. $\limsup_{n\to\infty} d(\chi_A[0 \ldots n]) = \infty$.

3 A New Measure on P via Martingale Families

The following equivalent alternative to martingales will be useful.

Definition 3.1 *A rate-martingale is a function* $D : \{0,1\}^* \to [0,2]$ *such that for every* $w \in \{0,1\}^*$ $D(w0) + D(w1) = 2$.

A rate-martingale outputs the factor by which the capital is increased after the bet, whereas a martingale outputs the current capital.

The key idea to define our measure on small complexity classes is that instead of considering a single martingale as usual, we consider families of rate-martingales which share their wins. These rate-martingales are computed by Turing machines with random access to their input, i.e. machines that have oracle access to their input and can query any bit of it. To enable such machines to compute the length of their input F without reading it, we also provide them with $s_{|w|}$; this convention is denoted by $M^w(s_{|w|})$. Since these Turing machines need to approximate real numbers, we assume their output to be two binary numbers (a, b) corresponding to the rational number $\frac{a}{b}$. With this convention, rational numbers such as $1/3$ can be said to be computed exactly. Here is a definition of such a family of rate-martingales.

Definition 3.2 *(martingale families) A* P*-family of rate-martingales* $(\{D_i\}_i, \{Q_i\}_i, \mathrm{ind})$*, is a family of rate-martingales* $\{D_i\}_i$*, where* $Q_i : \mathbb{N} \to \mathcal{P}(\{0,1\}^*)$ *are disjoint polynomial-printable query sets (i.e. there is a Turing machine that on input* $(i, 1^n)$ *outputs all strings in* $Q_i(n)$ *in time polynomial in* n*), i.e.* $Q_i(n) \cap Q_j(n) = \emptyset$ *and* $Q_i(m) \subseteq Q_i(n)$ *for* $m < n$*,* $\mathrm{ind} : \{0,1\}^* \to \mathbb{N}$ *is a polynomial time computable function, such that* $D_i(L \upharpoonright x)$ *is computable by a random access Turing machine* M *in time polynomial in* $|x|$ *i.e.* $M^{L\upharpoonright x}(x, i) = D_i(L \upharpoonright x)$ *where* M *queries its oracle only on strings in* $Q_i(|x|)$*, and* $\mathrm{ind}(x)$ *is an index* i *such that* $x \notin Q_j(|x|)$ *for every* $j \neq i$*.*

For simplicity we omit the indexes and denote the family of rate-martingales by (D, Q, ind), unless needed. Each rate-martingale D_i of the family only bets on strings inside its query set Q_i. The function ind on input a string x, outputs which rate-martingale is to (possibly) bet on x. The idea is that the rate-martingales share their wins, and have the ability to divide the bets along all members of the family. We are interested in the total capital such a family wins.

Definition 3.3 *Let* (D, Q, ind) *be a* P*-family of rate-martingales such that* $D_i(\lambda) \leq 1$ *for every* i*. The wins of a* P*- family of rate-martingales is the function* $W_D : \{0,1\}^* \to \mathbb{Q}$*, where* $W_D(L \upharpoonright x) = \prod_{i \leq 2^{|x|}} \prod_{y \leq x} D_i(L \upharpoonright y)$.

For simplicity we simply write i for the index of the first product, unless needed. Remember that $D_i(L \upharpoonright x)$ is the factor by which the capital is multiplied after the bet on x. Thus the product in Definition 3.3 is exactly the total capital the whole family of rate-martingales would win, would they be able to share their wins after each bet. Note that the function W_D is not polynomial, but only exponential time computable. This is a major difference to previous measure notions on P: computing the global wins of the family of rate-martingales is above the computational power of P.

A class has measure zero if there is a family of rate-martingales whose wins on the languages of the class are unbounded. Here is a definition.

Definition 3.4 *A class* C *of languages has* P*-measure zero, denoted* $\mu_{\mathsf{P}}(C) = 0$*, if there is a* P*-family of rate-martingales* (D, Q, ind) *such that for every* $L \in C$*,* $\limsup_{n\to\infty} W_D(L \upharpoonright s_n) = \infty$.

Whenever D's capital grows unbounded on L, we say that the family of rate-martingales succeeds on L, and write $L \in S^\infty[D]$. We call our measure notion F-measure.

It is easy to see that at higher complexity levels such as EXP, F-measure is equivalent to Lutz's measure notion [10], by taking a family containing a unique rate-martingale.

To prove a non-general union property we need rate-martingales that succeed independently, i.e. where every member in the family succeeds starting from any capital.

Definition 3.5 *The independent success set of a* P*-family of rate-martingales* (D, Q, ind) *denoted* $S_I^\infty[D]$ *is the set of languages* L *such that for every* $\alpha > 0$, $\limsup_{n\to\infty} \prod_i \alpha \prod_{y \le s_n} D_i(L \upharpoonright y) = \infty$.

It is sometimes more convenient to output the current capital of a rate-martingale, rather than the factor of increase. It is easy to check that Definition 3.2 can be reformulated by taking families of martingales instead of rate-martingales. We call such a family a P-family of martingales. Both definitions are equivalent, i.e. if (D, Q, ind) is a P-family of rate-martingales then (d, Q, ind) with $d_i(L \upharpoonright x) = \prod_{\{y | y \le x \text{ and } y \in Q_i(|x|)\}} D_i(L \upharpoonright y)$ is a P-family of martingales with the same win function. For the other direction take $D_i(L \upharpoonright x) = \frac{d_i(L\upharpoonright x)}{d_i(L\upharpoonright x-1)}$. Since both definitions are equivalent we shall switch from one to the other depending on which is the most appropriate in a given context.

Sometimes we need approximable martingales instead of exactly computable ones. Here is a definition.

Definition 3.6 *A* P*-approximable family of martingales* $(\{d_i\}_i, \{Q_i\}_i, \text{ind})$*, is a family of martingales* $\{d_i\}_i$*, where* Q_i *and* ind *are as in Definition 3.2 and such that* $d_i(L \upharpoonright x)$ *is* k*-approximable by a random access Turing machine* M *in time polynomial in* $|x| + k$*, i.e.* $|M^{L\upharpoonright x}(x, i, k) - d_i(L \upharpoonright x)| \le 2^{-k}$ *where* M *queries its oracle only on strings in* $Q_i(|x|)$.

3.1 The Basic Measure Properties

Let us show the union property for the following non-general case, where the query sets Q_i are the same for each family of rate-martingales to be considered for the union.

Definition 3.7 *A* P*-union of measure zero sets is a family of classes* $\{C_j\}_j$ *such that there exists a* P*-family of rate-martingales* $(\{D_{i,j}\}_{i,j}, \{Q_i\}_i, \text{ind})$ *such that for every* $j \ge 1$, $C_j \subseteq S_I^\infty[\{D_{i,j}\}_i]$.

As the following result shows, the basic measure properties hold for F-measure, as long as we restrict ourselves to P-unions.

Theorem 3.1

1. *Let* L *be any language in* P*, then* $\{L\}$ *has* P*-measure zero.*
2. P *does not have* P*-measure zero.*
3. *Let* $\{C_j\}_j$ *be a* P*-union of measure zero sets, and let* $C = \bigcup_j C_j$*, then* C *has* P*-measure zero.*

It is easy to check that F-measure on P can be extended to a measure notion on $\mathsf{QUASIPOLY}$, E_ϵ, and PSPACE, by taking the corresponding time and space bounds. For a measure on BPP we refer the reader to [21].

3.2 Applications: Some Classes of Measure Zero

3.3 Smallness of Languages with Low Density

As mentioned earlier martingale families can bet on every string, thus yielding a randomness notion which is optimal in terms of density of random languages.

Theorem 3.2 *Let $0 \leq \epsilon < 1/2$. The set D_ϵ of languages L such that for infinitely many n $|L[s_1, s_2, \cdots, s_n]| \leq \epsilon n$, has P-measure zero.*

An immediate Corollary of Theorem 3.2 is that the class SPARSE of languages containing few information is small in P, as opposed to Γ-measure [1].

Corollary 3.1 SPARSE *has* P*-measure zero.*

3.4 Almost Every Language in SUBEXP Can Derandomize BPP

We improve a former result of [1] by showing that almost every language A in E_ϵ can derandomize BPP^A.

Theorem 3.3 *For every $\epsilon > 0$, the set of languages A such that $\mathsf{P}^A \neq \mathsf{BPP}^A$ has E_ϵ -measure zero.*

3.5 Almost Every Language in PSPACE does not have Small Circuit Complexity

The following result shows that almost every language in PSPACE does not have small nonuniform complexity.

Theorem 3.4 *Let $c > 0$, $\mathsf{SIZE}(n^c)$ has PSPACE-measure zero.*

3.6 Comparison with Previous Measure Notions

The following result shows that F-measure is strictly stronger than Γ-measure [1].

Theorem 3.5 *μ_{P} is stronger than μ_Γ, i.e. for every class C, $\mu_\Gamma(C) = 0$ implies $\mu_{\mathsf{P}}(C) = 0$ and there are classes C such that $\mu_\Gamma(C) \neq 0$ and $\mu_{\mathsf{P}}(C) = 0$.*

We cannot compare F-measure to $\Gamma/(\mathsf{P})$-measure [23] directly, due to their intrinsic differences: a language L is said to have $\Gamma/(\mathsf{P})$-measure zero if there exists a "game strategy" which succeeds on *any* subsequences of L. This leads to the unnatural situation where for any random language L, $L \cup \{0\}^*$ does not have $\Gamma/(\mathsf{P})$-measure zero, although there are infinitely many easy instances. It is easy to check that such a set has P-measure zero. Nevertheless all sets proved to be small for $\Gamma/(\mathsf{P})$-measure in [23] are also small for F-measure. Regarding density arguments, F-measure performs better; indeed a (Lebesgue) random language has with high probability $(1/2 - o(1))2^n$ words of length n, and this property is captured by F-measure in Theorem 3.2, whereas for $\Gamma/(\mathsf{P})$-measure, the set of languages having $o(2^n)$ words of length n has $\Gamma/(\mathsf{P})$-measure zero. The advantage of $\Gamma/(\mathsf{P})$-measure over F-measure is that it satisfies the finite union property. Since $\Gamma/(\mathsf{P})$-measure is derived from Γ_d-measure [23], we cannot compare Γ_d-measure to F-measure, and both their respective strength are different: whereas Γ_d-measure cannot be used to define dimension in P, F-measure fails to capture the Γ_d-measure zero sets in [3].

3.7 Equivalence Between Measure on EXP and SUBEXP

Many results have been obtained from the plausible hypothesis $\mu_{\mathsf{E}}(\mathsf{NP}) \neq 0$ see for instance [16, 8], and the E-measure of all classes $\mathsf{ZPP}, \mathsf{RP}, \mathsf{BPP}, \mathsf{SPP}$ is now well understood, [19, 8, 7]. The following theorem shows that all these results follow from the a priori weaker assumption in terms of measure in E_ϵ.

Theorem 3.6 *Let C be a class downward closed under $\leq_m^p$-reducibilities, and let $\alpha > 0$. We have $\mu_{\mathsf{E}_\alpha}(C) \neq 0$ iff $\mu_{\mathsf{EXP}}(C) \neq 0$.*

4 Dimension on P

To define a dimension notion from F-measure, we need some minor modification for technical reasons. From now on we only consider P-families where the query sets of Definition 3.2 cover all strings of some size, and where the number of martingales allowed to bet on strings of size n is bounded, i.e. we require $\cup_{i \leq 2^n/n} Q_i(n) = \{0,1\}^{\leq n}$.

Lutz's key idea to define resource-bounded dimension is to tax the martingales' wins. The following definition formalizes this tax rate notion.

Definition 4.1 *Let $s \in [0,1]$ and (D, Q, ind) be a P-family of rate-martingales, and let L be a language. We say D s-succeeds on L, if $\limsup_{n\to\infty} 2^{(s-1)n} W_D(L \upharpoonright n) = \infty$.*

Similarly D s-succeeds on class C, if D s-succeeds on every language in C.

The dimension of a complexity class is the highest tax rate that can be received on the martingales' wins without preventing them from succeeding on the class.

Definition 4.2 *Let C be a class of languages. The P-dimension of C is defined as* $\dim_{\mathsf{P}}(C) = \inf\{s \in [0,1] :$ *there is a P-family of rate-martingales D that s-succeeds on C $\}$.*

We say C has dimension s in P denoted $\dim(C|\mathsf{P})$ if $\dim_{\mathsf{P}}(C \cap \mathsf{P}) = s$. If lim sup is replaced with lim inf in Definition 4.1, we say strongly s-succeed, and denote by $\mathrm{Dim}_{\mathsf{P}}$ the associated dimension notion. This is similar to the packing dimension notion from [2].

P-dimension satisfies a non-general union property, as shown in the following result.

Theorem 4.1 *Let $\{C_j\}_j$ be a family of classes, and let $\{s_j\}_j$ with $s_j \in [0,1]$ such that for every $\epsilon > 0$ there exists a P-family of martingales $\{d_{i,j}\}_{i,j}$ such that $\{d_{i,j}\}_i$ $(s_j + \epsilon)$-succeeds on C_j. Let $C = \bigcup_j C_j$, then $\dim_{\mathsf{P}}(C) \leq \sup_j\{s_j\}$.*

It is easy to check that P-dimension can be extended to classes above P like QUASIPOLY, subexponential time and PSPACE; for BPP see [21].

4.1 Finite-State Dimension versus P-dimension

The following result gives some evidence that P-dimension is a natural extension of previous dimension notions to the class P.

Theorem 4.2 *Let S be a language. Then $\dim_{\mathrm{FS}}(S) \geq \dim_{\mathsf{P}}(S) \geq \dim_{\mathsf{E}}(S)$.*

4.2 Application: Connecting Frequency and Shannon Entropy

In this section we show a polynomial time version of the Theorem of Eggleston [5], i.e. we prove that the class of languages with asymptotic frequency α have strong dimension the Shannon entropy of α in P. Analogue version of the theorem of Eggleston have been proved for various resource bounds [4, 13].

Let us introduce the following notations. First the Shannon entropy refers to the following continuous function $H : [0,1] \to [0,1]$, $H(\alpha) = \alpha \log \frac{1}{\alpha} + (1-\alpha) \log \frac{1}{1-\alpha}$.

For a language A and $n \in \mathbb{N}$, let $\text{freq}_A(n) = \frac{\#(1, A[0 \dots n-1])}{n}$, where $\#(1, A[0 \dots n-1])$ is the number of 1's in $A[0 \dots n-1]$. For $\alpha \in [0,1]$, let $\mathsf{FREQ}(\alpha) = \{A \in \{0,1\}^\infty | \lim_{n\to\infty} \text{freq}_A(n) = \alpha\}$.

The following is a polynomial time version of the Theorem of Eggleston [5].

Theorem 4.3 *For all* E*-computable* $\alpha \in [0,1]$*, we have* $\text{Dim}(\mathsf{FREQ}(\alpha)|\mathsf{P}) = H(\alpha)$.

5 Conclusion

From the quest for the perfect measure on P a widespread consensus has emerged that for measure on small complexity classes some properties need to be renounced. The main contribution of our measure notion is that unlike previous measure notions on P, it leads to a reasonable way to define dimension in P. The price to pay is that martingale families only satisfy a non-general union property. We expect our measure and dimension notions to be useful for further measure-based investigations in small complexity classes.

Acknowledgements

I thank Jack Lutz for providing me his notes on normal sequences [15], enabling me to prove Theorem 4.2.

References

[1] E. Allender and M. Strauss. Measure on small complexity classes, with application for BPP. *Proc. of the 35th Ann. IEEE Symp. on Found. of Comp. Sci.*, pages 807–818, 1994.

[2] K. B. Athreya, J. M. Hitchcock, J. H. Lutz, and E. Mayordomo. Effective strong dimension in algorithmic information and computational complexity. *Proceedings of the Twenty-First Symposium on Theoretical Aspects of Computer Science*, pages 632–643, 2004.

[3] J.-Y. Cai, D. Sivakumar, and M. Strauss. Constant-depth circuits and the lutz hypothesis. *Proc. 38'th Foundations of Computer Science Conference*, pages 595–604, 1997.

[4] J. Dai, J. Lathrop, J. Lutz, and E. Mayordomo. Finite-state dimension. *Theoretical Computer Science*, 310:1–33, 2004.

[5] H. Eggleston. The fractional dimension of a set defined by decimal properties. *Quarterly Journal of Mathematics*, 20:31–36, 1949.

[6] L. Fortnow and J. Lutz. Prediction and dimension. *Journal of Computer and System Sciences*, 70:570–589, 2005.

[7] J. M. Hitchcock. The size of SPP. *Theoretical Computer Science*, 320:495–503, 2004.

[8] R. Impagliazzo and P. Moser. A zero-one law for RP. *Proceedings of the 18th Conference on Computational Complexity*, pages 48–52, 2003.

[9] A. Klivans and D. van Melkebeek. Graph nonisomorphism has subexponential size proofs unless the polynomial hierarchy collapses. *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 659–667, 1999.

[10] J. Lutz. Category and measure in complexity classes. *SIAM Journal on Computing*, 19:1100–1131, 1990.

[11] J. Lutz. Almost everywhere high nonuniform complexity. *Journal of Computer and System Science*, 44:220–258, 1992.

[12] J. Lutz. The quantitative structure of exponential time. In L. Hemaspaandra and A. Selman, editors, *Complexity Theory Retrospective II*, pages 225–260. Springer, 1997.

[13] J. Lutz. Dimension in complexity classes. *Proceedings of the 15th Annual IEEE Conference on Computational Complexity*, pages 158–169, 2000.

[14] J. Lutz. The dimensions of individual strings and sequences. *Information and Computation*, 187:49–79, 2003.

[15] J. Lutz. Introduction to normal sequences. *Lecture notes*, 2004.

[16] J. Lutz and E. Mayordomo. Cook versus Karp-Levin: Separating completeness notions if NP is not small. *SIAM Journal on Computing*, 164:141–163, 1996.

[17] J. H. Lutz. Effective fractal dimensions. *to appear in Mathematical Logic Quarterly*, 2003.

[18] E. Mayordomo. Measuring in PSPACE. *Proceedings of the 7th International Meeting of Young Computer Scientists (IMYCS'92). Gordon-Breach Topics in Computer Science 6*, 136:93–100, 1994.

[19] D. Melkebeek. The zero-one law holds for BPP. *Theoretical Computer Science*, 244(1-2):283–288, 2000.

[20] P. Moser. Baire's categories on small complexity classes. *14th Int. Symp. Fundamentals of Computation Theory*, pages 333–342, 2003.

[21] P. Moser. Resource-bounded measure on probabilistic classes. *submitted*, 2004.

[22] C. Papadimitriou. *Computational complexity*. Addisson-Wesley, 1994.

[23] M. Strauss. Measure on P- strength of the notion. *Inform. and Comp.*, 136:1:1–23, 1997.

[24] J. Ziv and A. Lempel. Compression of individual sequences via variable-rate coding. *IEEE Transaction on Information Theory*, pages 530–536, 1978.

A Appendix

A.1 Proof of Theorem 3.1

Notation: $L|s_{n+1}$ denotes the initial segment of L up to s_n given by $L[s_0 \cdots s_n]$. Let $L \in \mathsf{P}$ and M be a polynomial time Turing machine deciding L. Divide $\{0,1\}^n$ into $2^n/n$ zones of n consecutive strings denoted B_i^n, with $i = 1, 2, \cdots 2^n/n$. Consider the following P-family of rate-martingales (D, Q, ind) where $Q_i(n) = \bigcup_{j=1}^{n} B_i^j$ and $\text{ind}(x)$ is the index i such that $x \in Q_i(|x|)$. Let A be any language. Strategy D_i bets all its capital on strings in Q_i according to M, i.e. let $x \in B_i^n$, then $D_i(A \upharpoonright x) = 2$ whenever $A(x) = M(x)$, otherwise $D_i(A \upharpoonright x) = 0$. It is easy to check that (D, Q, ind) is a P-family of rate-martingales. $L \in S^\infty[D]$ because the family of rate-martingales doubles its capital after every bet, i.e. $\limsup_{n\to\infty} W_D(L \upharpoonright s_n) = \limsup_{n\to\infty} \prod_i \prod_{y \leq s_n,} D_i(L \upharpoonright y) = \limsup_{n\to\infty} 2^n = \infty$, which ends the proof of the first property.

For the second property, let (D, Q, ind) be a P-family of rate-martingales. Consider the following language $L \in \mathsf{P}$. Let $x \in \{0,1\}^*$, define $L(x) = 0$ iff $D_i((L|x)0) \leq 1$ where $i = \text{ind}(x)$. L is computable in polynomial time because the machine computing $D_i((L|x)0)$ only queries $L|x$ on strings contained in $Q_i(|x|)$, therefore requiring only a polynomial number of recursive steps. Because the Q_i's are disjoint, only computations of D_i have to be performed. Thus $L \in \mathsf{P}$. The strategy family does not succeed on L, since $\limsup_{n\to\infty} W_D(L \upharpoonright s_n) = \limsup_{n\to\infty} \prod_i \prod_{y \leq s_n,} D_i(L \upharpoonright y) \leq 1$ i.e. $L \notin S^\infty[D]$, which ends the proof.

For the third property, we need the following Lemma.

Lemma A.1 *Let (d, Q, ind) be a P-approximable family of martingales, then there exists a P-computable family of martingales (d', Q, ind) with the same query set and* ind *function, such that for any $w \in \{0,1\}^*$ and every i $d'_i(w) \geq d_i(w)$.*

Let (d, Q, ind) be as above and let $i \geq 1$. Denote by $\{d_{i,k}\}$ the approximation of d_i where $|d_{i,|w|}(w) - d_i(w)| \leq \frac{1}{|w|^2}$. Consider the following martingale d'_i, with initial capital $d'_i(\lambda) = 2$ where for wb with $w \in \{0,1\}^*$, and $b \in \{0,1\}$ is the membership bit of some string x, with $x \in Q_i(|x|)$, is defined as follows. Let $d'_i(wb) = d'_i(w) + \frac{d_{i,|wb|}(wb) - d_{i,|wb|}(w\bar{b})}{2}$. If $x \notin Q_i(|x|)$, then $d'_i(wb) = d'_i(w)$. Since $Q_i(|x|)$ is poly-printable, computing $d'_i(wb)$ only requires a polynomial number of recursive steps. It is easy to check that d'_i is a martingale, thus (d', Q, ind) is P-family of martingales. Let us check that $d'_i(w) \geq d_i(w) + \frac{1}{|w|}$ by induction. The inequality holds for $w = \lambda$. Let $w \in \{0,1\}^*$ and $b \in \{0,1\}$, we have $d'_i(wb) = d'_i(w) + \frac{d_{i,|wb|}(wb) - d_{i,|wb|}(w\bar{b})}{2} \geq d_i(w) + \frac{1}{|w|} + \frac{d_{i,|wb|}(wb) - d_{i,|wb|}(w\bar{b})}{2}$ by induction hypothesis. Since $d_{i,|wb|}(wb) \geq d_i(wb) - \frac{1}{|wb|^2}$ we have $d'_i(wb) \geq d_i(w) + \frac{1}{|w|} + \frac{d_i(wb) - d_i(w\bar{b})}{2} - \frac{1}{|wb|^2}$. Because d_i is a martingale, we have $d_i(w) - \frac{1}{2} d_i(w\bar{b}) = \frac{1}{2} d_i(wb)$ thus $d'_i(wb) \geq d_i(wb) + \frac{1}{|w|} - \frac{1}{|wb|^2} \geq d_i(wb) + \frac{1}{|wb|}$ which ends the proof of the lemma.

Let us prove the theorem. Let $\{C_j\}_j$ be a P-union of measure zero sets, and let (d, Q, ind) be a family of rate-martingales witnessing this fact. Consider the following family of martingales given by $d'_i(w) = \sum_{j \geq 1} \frac{1}{2^j} d_{i,j}(w)$. Let us show that d'_i is P-approximable. Consider the following approximation $d'^k_i(L \upharpoonright x) = \sum_{j=1}^{q(k+|x|)} \frac{1}{2^j} d_{i,j}(L \upharpoonright x)$ where q is a polynomial to be determined later. Because all $d_{i,j}$'s have polynomial size query set, so does d'^k_i and therefore it is polynomial time computable in $|x| + i + k$. We have $|d'_i(L \upharpoonright x) - d'^k_i(L \upharpoonright x)| \leq$

$\sum_{j>q(k+|x|)} \frac{1}{2^j} d_{i,j}(L \upharpoonright x)$. Since $d_{i,j}(L \upharpoonright x) \leq 2^{|x|^c}$ for some $c > 0$, we have $|d'_i(L \upharpoonright x) - d'^k_i(L \upharpoonright x)| \leq \frac{2^{|x|^c}}{2^{q(k+|x|)}} \leq 2^{-k}$ by choosing $q(y) = y^{c+1}$.

By Lemma A.1 there exists a P-computable family of martingale $\bar{d}_i$ such that $\bar{d}_i(L \upharpoonright x) \geq d'_i(L \upharpoonright x)$ for all strings x, and $\frac{1}{2}\bar{d}_i(\lambda) \leq 1$. Thus $\frac{1}{2}\bar{d}_i(L \upharpoonright x) \geq \frac{1}{2 \cdot 2^j} d_{i,j}(L \upharpoonright x)$ for all i, j, x. Let $j > 0$ and let $L \in S^\infty_I[\{d_{i,j}\}_i]$. We have $\limsup_{n\to\infty} \prod_i \frac{1}{2}\bar{d}_i(L \upharpoonright s_n) \geq \limsup_{n\to\infty} \prod_i \frac{1}{2^{j+1}} d_{i,j}(L \upharpoonright s_n) = \infty$ i.e. $C_j \subseteq S^\infty[\bar{d}]$. □

A.2 Proof of Theorem 3.2

Let $0 \leq \epsilon < 1/2$ and let $\alpha = 1/2 - \epsilon$. Divide the strings of size n into $2^n/n$ blocks of size n denoted $B_1, \ldots, B_{2^n/n}$. Consider the following family of rate-martingales $\{D_i\}_i$, where D_i bets a fraction α of its current capital that the strings in B_i have membership bits zero. It is easy to check that $\{D_i\}_i$ is a P-family of rate-martingales; thus whenever D_i' s bet is correct (resp. false), the capital is multiplied by a factor $1 + \alpha$ (resp. $1 - \alpha$). Let $L \in D_\epsilon$, we have for infinitely many n $W_D(L \upharpoonright s_n) = \prod_i \prod_{y \leq s_n} D_i(L \upharpoonright y) \geq [(1+\alpha)^{(\frac{1}{2}+\alpha)}(1-\alpha)^{(\frac{1}{2}-\alpha)}]^n$. Since $(1+\alpha)^{(\frac{1}{2}+\alpha)}(1-\alpha)^{(\frac{1}{2}-\alpha)} > 1$ we have $L \in S^\infty[D]$ □

A.3 Proof of Theorem 3.3

We use the standard model of oracle Boolean circuits see [22] for more details. For a bound function t we denote by $\mathsf{SIZE}(t(n))$ the set of languages decided by a family of circuits of size $t(n)$, where n is the size of the input. The circuit complexity of a Boolean function $f : \{0,1\}^n \to \{0,1\}$, is the size of the smallest circuit computing f.

Definition A.1 *Let A be any language. The hardness $H^A(G_{m,n})$ of a random generator $G_{m,n} : \{0,1\}^m \longrightarrow \{0,1\}^n$, is defined as the minimal s such that there exists an n-input circuit C with oracle gates to A, of size at most s, such that $|\Pr_{x\in\{0,1\}^m}[C(G_m(x)) = 1] - \Pr_{y\in\{0,1\}^n}[C(y) = 1]| \geq \frac{1}{s}$.*

We need the following pseudorandom generator from [9].

Theorem A.1 (Klivans-Melkebeek) *Let A be any language. There is a polynomial-time computable function F such that for every $\epsilon > 0$, there exists $a, b \in \mathbb{N}$ such that $F : \{0,1\}^{n^a} \times \{0,1\}^{b \log n} \to \{0,1\}^n$, and if r is the truth table of a $(a \log n)$-variables Boolean function of A-oracle circuit complexity at least $n^{\epsilon a}$, then the function $G_r(s) = F(r, s)$ is a generator with hardness $H^{\mathrm{A}}(G_r) > n$.*

Lt us prove Theorem 3.3. Let $\epsilon > 0$, let $0 < \delta < \max(\epsilon, 1/2)$, and $b > 0$ be some constant to be determined later. Consider the following martingale d betting only on strings of size $m = n + \frac{1}{b}\log n$ for some integer n. Let Z_m be the set of strings of the form $0^{2^{b|u|}}u$ where $u \in \{0,1\}^{\frac{1}{b}\log n}$, clearly $Z_m \subset \{0,1\}^m$. Denote by $C^w(l,t)$ with $l \leq t$ the set of F-oracle l-inputs circuits of size less than t, and denote by $C^w(l,t,u)$ the set of circuits C in $C^w(l,t)$ such that for every $z = 0^{2^{b|v|}}v \in Z_m$ whose membership bit is is in the u zone of wu, where wu is viewed as the prefix of the characteristic sequence of some language, we have $C(v) = wu[z]$. It is well known [22] that $|C^w(l,t)| \leq 2^{t \log t}$. Let $B(w,u,m)$ denote the number of $z \in Z_m$ whose membership bits are in the u zone of wu. Let F be the prefix of the characteristic sequence of some language L, coding for words up to size $\leq m-1$, and let $u \in \{0,1\}^*$, with $0 < |u| \leq 2^m$.

Let $d(wu) = \frac{|C^w(\frac{1}{b}\log n, n^{\delta/b}, u)|}{|C^w(\frac{1}{b}\log n, n^{\delta/b})|} 2^{B(w,u,m)} d(w)$. It is easy to check that d is a martingale. d is computable in time 2^{m^ϵ}, because there are $2^{n^{2\delta/b}}$ circuits to simulate which takes time less than 2^{m^ϵ} for an appropriate choice of b. For the dependency set, since the circuits to be simulated have size less than $n^{\delta/b}$, they can only query F on the membership bits of strings of size at most $n^{\delta/b}$, moreover d only bets on strings in Z_m, thus $G(m) = \bigcup_{j=1}^{m} Z_j \cup \{0,1\}^{\leq n^{\delta/b}}$, which has size less than $2^{n^{\delta/b}} + mn^{1/b}$ which is less than 2^{m^ϵ}.

Let A be any language and consider $F(A) := \{u | 0^{2^{b|u|}} u \in A\}$. It is clear that $F(A) \in \mathsf{E}^A$. Consider H_δ^A the set of languages L such that every n-input circuits with oracle gates for A of size less than $2^{\delta n}$ fails to compute L. We have, $F(A) \in H_\delta^A$ implies $\mathsf{P}^A = \mathsf{BPP}^A$ by Theorem A.1.

We show that d succeeds on every language A such that $F(A) \notin H_\delta^A$. Let A be any such language, let F be the prefix of A coding for strings up to size $m-1$ as above, and let $u \in \{0,1\}^{2^m}$, thus for n large, $d(wu) = \frac{|C^w(\frac{1}{b}\log n, n^{\delta/b}, u)|}{|C^w(\frac{1}{b}\log n, n^{\delta/b})|} 2^{B(w,u,m)} d(w) \geq \frac{1}{|C^w(\frac{1}{b}\log n, n^{\delta/b})|} 2^{n^{1/b}} d(w)$ $\geq \frac{2^{n^{1/b}}}{2^{n^{2\delta/b}}} d(w) \geq 2^{n^{1/2b}} d(w)$, i.e. $A \subseteq S^\infty[\{d_i\}_i]$. □

A.4 Proof of Theorem 3.4

Let $c > 0$. For $n \leq t$ denote by $C(n,t)$ the number of n-inputs Boolean circuits of size t. Divide the strings of size n into consecutive blocks of size n^{c+1} denoted $R_1^n, \cdots, R_{2^n/n^{c+1}}^n$. Consider the following family of martingales $\{d_i\}_i$, where d_i bets on strings in R_i. Let F be the initial segment of a language L coding for strings up to R_{i-1}^n, and let $0 < |u| \leq n^{c+1}$. Consider $d_i(wu) = \frac{C(n,n^c,w,u)}{C(n,n^c)} 2^{|u|} d_i(w)$ where $C(n,t,w,u)$ is the number of n-inputs Boolean circuits of size t deciding some language $A \in \{0,1\}^n$ such that $u \sqsubseteq A[R_i^n]$. It is easy to check that d_i is a martingale. $\{d_i\}_i$ is a $\mathsf{DSPACE}(n^{c+2})$-family of martingales because $C(n,n^c,w,u)$ and $C(n,n^c)$ are computable in $\mathsf{DSPACE}(n^{c+2})$ by constructing all corresponding circuits and reading the input on u, thus $Q_i(n) = \bigcup_{j \leq n} R_i^j$.

Let L be a language in $\mathsf{SIZE}(n^c)$, then for $w, |u| = n^{c+1}$ as above we have $d_i(wu) = \frac{C(n,n^c,w,u)}{C(n,n^c)} 2^{n^{c+1}} d_i(w) \geq \frac{1}{C(n,n^c)} 2^{n^{c+1}} d_i(w) \geq 2^{n^{c+1} - n^c c \log n} d_i(w) \geq 2^{\frac{n^{c+1}}{2}} d_i(w)$. Thus $L \in S^\infty[d]$. □

A.5 Proof of Theorem 3.5

The Γ-measure introduced in [1] is defined through single P-computable martingales with poly-printable query sets. Let (d, Q_d) be such a martingale, running in time n^c. Divide the strings of size n into blocks of size n denoted $R_1^n, \cdots, R_{2^n/n}^n$. Consider the following family of rate-martingales $\{d_i\}_i$, where $d_0 = d$ $d_i \equiv 1$ for $i \geq 1$ $Q_i(m) = \bigcup_{j=1}^{m} R_i^j - Q_d(m)$ and $Q_0(m) = Q_d(m)$. Let $\mathrm{ind}(x) = 0$ for all x. It is easy to check that $\{d_i\}_i$ is a P-family of martingales, whose win function is equal to the single martingale d. Finally, it is shown in [1] that the class SPARSE does not have Γ-measure zero, thus Theorem 3.2 ends the proof. □

A.6 Proof of Theorem 3.6

Let $\alpha > 0$. Let C be a class downward closed under $\leq_m^p$-reducibilities, and such that $\mu_{\mathsf{EXP}}(C) = 0$; Let d denote the martingale witnessing this fact, and suppose d runs in time

2^{n^k}. For a language L denote by L' the following padded version of L where $L' = \{0^{|x|^{k/\alpha}}1x : x \in L\}$. Clearly $L' \leq_m^p L$, thus $L' \in C$. For a prefix X of some characteristic sequence, let X' be given by $X'(y) = X(0^{|y|^{k/\alpha}}1y)$. Consider the following E_α-computable martingale d' that bets only on strings of the form $0^{|x|^{k/\alpha}}1x$, and defined by $d'(X \upharpoonright 0^{|x|^{k/\alpha}}1x) = d(X' \upharpoonright x)$. It is easy to check that d' is computable in time 2^{n^α}, and has a query set of size 2^{n^α}. Let $L \in C$, thus $L' \in C$, and $d'(L \upharpoonright 0^{|x|^{k/\alpha}}1x) = d(L' \upharpoonright x)$. Since $L' \in S^\infty[d]$ this ends the proof. □

A.7 Proof of Theorem 4.1

The proof is similar to Theorem 3.1. Let $\epsilon > 0$, $s = \sup_j\{s_j\}$ and let $\{d_{i,j}\}_{i,j}$ be a P-family of martingales such that $\{d_{i,j}\}_j$ $(s_j + \epsilon/2)$-succeeds on C_j. Denote by d'_i the sum of the family of martingales as in Theorem 3.1. Let us check that d'_i $(s+\epsilon)$-succeeds on C. Let $L \in C_j$ for some j, we have $d'(w) \geq \frac{1}{2^j} d_{i,j}(w)$ for every i, and $\frac{1}{2}d'(\lambda) \leq 1$. Let d' denote $\frac{1}{2}d'$, we have $\limsup_{n\to\infty} 2^{(s+\epsilon-1)n} W_{d'}(L \upharpoonright s_n) = \limsup_{n\to\infty} 2^{(s+\frac{\epsilon}{2}-1)n} 2^{(\frac{\epsilon}{2})n} \prod_i d'_i(L \upharpoonright s_n) \geq \limsup_{n\to\infty} 2^{(s+\frac{\epsilon}{2}-1)n} 2^{(\frac{\epsilon}{2}-\frac{j+1}{\log n})n} \prod_i d_{i,j}(L \upharpoonright s_n) \geq \limsup_{n\to\infty} 2^{\frac{\epsilon n}{4}} 2^{(s_j+\frac{\epsilon}{2}-1)n} W_{\{d_{i,j}\}_i}(L \upharpoonright s_n) = \infty$. □

A.8 Proof of Theorem 4.2

Finite-state dimension is defined via martingale computable by finite-state machines (called FSG: finite-state gamblers), see [4] for more details.

Let $p : \{0,1\}^l \to [0,1]$ be a probability measure, i.e. $\sum_{x\in\{0,1\}^l} p(x) = 1$. The Shannon entropy of p is given by

$$H(p) = \mathrm{E}_p \log \frac{1}{p(x)} = \sum_{x\in\{0,1\}^l} p(x) \log \frac{1}{p(x)}.$$

The Kullback-Leibler divergence between two probability measures p, q is

$$D(p,q) = \sum_{x\in\{0,1\}^l} p(x) \log \frac{p(x)}{q(x)}.$$

Any binary sequence S, yields a probability measure called its empirical block probability measure $\pi_n^l(S)(w) : \{0,1\}^l \to [0,1]$ given by

$$\pi_n^l(S)(w) = \mathrm{freq}(w, S[1\ldots ln]) = \frac{\#(w, S[1\ldots ln])}{ln}$$

where $\#(w,x) = |\{u \in \{0,1\}^* : uw \sqsubseteq x\}|$ is the number of occurrences of w in x.

The following theorem is implicit in [24].

Theorem A.2 *For a binary sequence S, and $s \in [0,1]$, the following are equivalent.*

1. *There is a FSG that s-succeeds on S.*
2. *There exist $l \in \mathbb{Z}^+$ such that $\liminf_{n\to\infty} H(\pi_n^l(S)) < sl$.*

Let us prove Theorem 4.2 (the proof is similar to [15]). Let S be a binary sequence, and let $s > \dim_{\mathrm{FS}}(S)$. By definition of finite-state dimension, there is a FSG that s-succeeds on S. By Theorem A.2, there exists $l \in \mathbb{Z}^+$ such that

$$\liminf_{n\to\infty} H(\pi_n^l(S)) < sl.$$

Since the set of probability measures on $\{0,1\}^l$ is compact, there is a probability measure $\psi : \{0,1\}^l \to [0,1]$ and a sequence of natural numbers $n_1 < n_2 < \ldots$ such that, $H(\psi) < sl$ and for any $w \in \{0,1\}^l$, $\lim_{k\to\infty} \pi_{n_k}^l(S)(w) = \psi(w)$. Let $\delta = \frac{sl - H(\psi)}{4}$. Because D is continuous, there exist a positive rational probability measure $\psi' : \{0,1\}^l \to \mathbb{Q} \cap [0,1]$ such that $D(\psi|\psi') < \delta$. Since H, D are continuous, there is an integer m such that for $k \geq m$ we have

$$H(\pi_{n_k}^l(S)) < H(\psi) + \delta$$

$$D(\pi_{n_k}^l(S)|\psi') < D(\psi|\psi') + \delta.$$

Let us divide $\{0,1\}^{=n}$ into $2^n/nl$ groups of nl consecutive strings, $I_1, I_2, \ldots, I_{2^n/nl}$, where martingale d_i will bet on I_i. Extend ψ' to a function $\psi' : \{0,1\}^{\leq l} \to [0,1]$ by $\psi'(u) = \sum_{v\in\{0,1\}^{l-|u|}} \psi'(uv)$. For a string xw, where w corresponds to the membership bits of the strings in I_i, let

$$d_i(xw[0\ldots k]) = 2^{k+1} d_i(x)\psi'(w[0\ldots k]).$$

It is easy to check that d_i is a martingale. Moreover d_i is computable in polynomial time, because ψ' is. Therefore the d_i's are a P-computable family of martingales, whose wins function is given by (for $x \in \{0,1\}^{ln}$ and $u \in \{0,1\}$)

$$W(ux) = W(u)2^{nl} \prod_{w\in\{0,1\}^l} \psi'(w)^{\#(w,x)}.$$

Denote $x_k = S[0\ldots n_k l - 1]$, hence taking tax rate $1 - s$ we have

$$\begin{aligned}
\log \frac{W(x_k)}{2^{(1-s)|x_k|}} &= sn_k l + \sum_{w\in\{0,1\}^l} \#(w, x_k) \log \psi'(w) \\
&= n_k[sl + \sum_{w\in\{0,1\}^l} \pi_{n_k}^l(S)(w) \log \psi'(w)] \\
&= n_k[sl - \mathrm{E}_{\pi_{n_k}^l(S)} \log(\frac{1}{\pi_{n_k}^l(S)(w)} \cdot \frac{\pi_{n_k}^l(S)(w)}{\psi'(w)})] \\
&= n_k[sl - H(\pi_{n_k}^l(S)) - D(\pi_{n_k}^l(S)|\psi')].
\end{aligned}$$

Therefore

$$\log \frac{W(x_k)}{2^{(1-s)|x_k|}} \geq n_k[sl - H(\psi) - 3\delta] = \delta n_k.$$

Since $\delta > 0$, W s-succeeds on S, i.e. $\dim_{\mathsf{P}}(S) \leq s$, thus $\dim_{\mathsf{P}}(S) \leq \dim_{\mathrm{FS}}(S)$. □

A.9 Proof of Theorem 4.3

The following result gives an upper bound on the strong P-dimension of $\mathsf{FREQ}(\alpha)$.

Theorem A.3 *For all $\alpha \in [0,1]$, we have* $\mathrm{Dim}_{\mathsf{P}}(\mathsf{FREQ}(\alpha)) \leq H(\alpha)$.

Proof. Wlog let $\alpha \in (0, \frac{1}{2}]$. Let $s > H(\alpha)$, $\delta = s - H(\alpha) > 0$, and let $\epsilon > 0$ such that $(\frac{\alpha}{1-\alpha})^\epsilon \geq 2^{-\frac{\delta}{2}}$. Divide $\{0,1\}^n$ into consecutive blocks of size n, denoted $R_1^n, R_2^n, \cdots, R_{2^n/n}^n$. Consider the following P-family of martingales d_i, where d_i bets a fraction $1-2\alpha$ of its current capital that the membership bit of strings in R_i is 0. Whenever this bet is correct (resp. false), the capital is multiplied by a factor $2(1-\alpha)$ (resp. 2α). Let $A \in \mathsf{FREQ}(\alpha)$, and let $N \in \mathbb{N}$ be such that $\forall n \geq N, \mathrm{freq}_A(n) \leq \alpha + \epsilon$. Thus for $n \geq N$ we have, $\frac{W_d(A|n)}{2^{(1-s)n}} = \frac{(2\alpha)^{\#(1,A|n)}(2(1-\alpha))^{\#(0,A|n)}}{2^{(1-s)n}} = [\frac{(2\alpha)^{\mathrm{freq}_A(n)}(2(1-\alpha))^{1-\mathrm{freq}_A(n)}}{2^{1-s}}]^n = [2^s\alpha^{\mathrm{freq}_A(n)}(1-\alpha)^{1-\mathrm{freq}_A(n)}]^n \geq [2^s\alpha^{\alpha+\epsilon}(1-\alpha)^{1-(\alpha+\epsilon)}]^n = [2^s\alpha^\alpha(1-\alpha)^{1-\alpha}(\frac{\alpha}{1-\alpha})^\epsilon]^n \geq [2^{s-H(\alpha)-\frac{\delta}{2}}]^n = 2^{\frac{\delta}{2}n}$. Because $\delta > 0$, F is unbounded, i.e. d strongly s-succeeds on A. □

For the other direction, we need the following notation. Let d be a P-computable family of martingales, let $i \geq 1$, $w, v \in \{0,1\}^*$. Suppose that the ordered dependency set of d_i is of the form $Q_i = \{\cdots, s_{|w0|}, s_{|w1|}, s_{|w2|}, \cdots, s_{|w_{|v|}|}, \cdots\}$ where $s_{|w0|} \leq s_{|w|}$ and $s_{|w|} < s_{|w_i|}$ $\forall i = 1, 2, \cdots |v|$. Define $(wv)^* = \{wz : wz[s_{|w_i|}] = v_i \text{ for } i = 1, 2, \cdots, |v| \text{ and } s_{|wz|} = s_{|w_{|v|}|}\}$. and let $d_i((wv)^*) = d_i(wz)$ where $wz \in (wv)^*$. $d_i((wv)^*)$ is well defined because d_i only bets on strings whose membership bits correspond to v.

We need the following generalization of Kraft inequality, which says that there are only a few strings on which taxed martingales win money.

Lemma A.2 *Let $s \in [0,1]$, let d be a P-family of martingales. For all $w \in \{0,1\}^*$, $i, l \in \mathbb{N}$ there are less than 2^{sn} strings $u \in \{0,1\}^l$ such that $\frac{d_i((wv)^*)}{2^{(1-s)|v|}} > d_i(w)$.*

Proof. Let s, d, w, i, l be as above. Consider the following random variable X over $\{0,1\}^k$, $X(u) = d_i((wu)^*)$. Thus $E(X) = \sum_{u\in\{0,1\}^k} 1/2^k X(u) = 1/2^k \sum_{u\in\{0,1\}^k} d_i((wu)^*) = 1/2^{k-1}(\sum_{u\in\{0,1\}^{k-1}} d_i((w(u))^*)) = \ldots = d(w)$. Using $\Pr_{u\in\{0,1\}^k}[X(u) > \alpha E(X)] < 1/\alpha$ with $\alpha = 2^{(1-s)k}$ ends the proof. □

The following result gives a lower bound on the P-dimension of $\mathsf{FREQ}(\alpha)$.

Theorem A.4 *Let $\alpha \in [0,1]$ be E-computable, we have* $\mathrm{Dim}_{\mathsf{P}}(\mathsf{FREQ}(\alpha) \cap \mathsf{P}) \geq H(\alpha)$.

Proof. Let α be as above. Wlog $\alpha \in (0,1)$. Let d be a P-family of martingales. Let $0 < s < H(\alpha)$. Let α' denote the E-approximation of α, i.e. $|\alpha'(n) - \alpha| \leq \frac{1}{n}$, where $\alpha'(n)$ is computable in time polynomial in n. Consider $m(n) = \lfloor \log(2n) \rfloor$ and $k(n) = \lfloor \alpha'(m(n))m(n) \rfloor$. We have $\alpha'(m(n)) - \frac{1}{m(n)} \leq \frac{k(n)}{m(n)} \leq \alpha'(m(n))$, thus $|\frac{k(n)}{m(n)} - \alpha| \leq \frac{2}{m(n)}$. Therefore, $\lim_{n\to\infty} \frac{k(n)}{m(n)} = \alpha$. Because H is continuous we have $\lim_{n\to\infty} H(\frac{k(n)}{m(n)}) = \alpha$.

Let $D_n = \{u \in \{0,1\}^{m(n)} : \#(1,u) = k(n)\}$. Using, $e(\frac{n}{e})^n < n! < en(\frac{n}{e})^n$ for $n \geq 1$, yields $|D_n| = \binom{m(n)}{k(n)} > \frac{2^{m(n)H(\frac{k(n)}{m(n)})}}{ek(n)(m(n)-k(n))} \geq 4\frac{2^{m(n)H(\frac{k(n)}{m(n)})}}{em^2(n)} > 2^{m(n)H(\frac{k(n)}{m(n)})-2\log m(n)}$. By continuity of H there exists $s' > s$ such that for sufficiently large n, $H(\frac{k(n)}{m(n)}) \geq s'$. Thus for sufficiently large n, $|D_n| > 2^{sm(n)+(s'-s)m(n)-2\log m(n)} \geq 2^{sm(n)}$.

Consider the following language L. Let $x \in \{0,1\}^*$, with $|x| = n$. Compute $i = \mathrm{ind}(x)$, and $Q_i^{=n}(n)$. We have $|Q_i^{=n}(n)| = q(n)m(n) + r(n)$ where q is a polynomial and $0 \leq r(n) < m(n)$.

Order the strings in $Q_i^{=n}(n)$ lexicographically and divide them into consecutive blocks of size $m(n)$ denoted $B_1^n, B_2^n, \cdots, B_{q(n)}^n, B_{q(n)+1}^n$ except for the last one which has size $r(n)$. Let $w = L \upharpoonright B_k^n$ with $1 \leq k \leq q(n)$. Find the first string $u \in D_n$ such that $\frac{d_i((wu)^*)}{2^{(1-s)|u|}} \leq d_i(w)$ Such a string u exists by Lemma A.2. Define L to be u on strings in B_{k+1}^n, i.e. if x is the jth string of B_{k+1}^n, then $L(x) = u_j$. For strings in $B_{q(n)+1}^n$ repeat the construction by trying all u's of size $r(n)$.

L is polynomial time computable because since $Q_i(n)$ is poly-printable, only a polynomial number of recursive steps needs to be performed. There are less than $2n$ strings u to try by definition of D_n. Thus $L \in \mathsf{P}$.

Let us show that $L \in \mathsf{FREQ}(\alpha)$. Because d is a P-family, we have $Q_i(n) = \emptyset$ for $i > \frac{2^n}{n}$. Whenever $|Q_i^{=n}(n)| \equiv 0 \mod m(n)$ the part of L defined on strings in $Q_i^{=n}(n)$ has optimal frequency $\frac{k(n)}{m(n)}$. So suppose (worst case) $|Q_i^{=n}(n)| \equiv m(n) - 1 \mod m(n)$. We have $\mathrm{freq}(L^{=n}) = \frac{\#(1,L^{=n})}{2^n} \leq \frac{\frac{2^n}{n}(m(n)-1)+k(n)\frac{2^n - \frac{2^n}{n}(m(n)-1)}{m(n)}}{2^n}$, thus $\lim_{n\to\infty} \mathrm{freq}(L^{=n}) \leq \alpha$. Similarly $\lim_{n\to\infty} \mathrm{freq}(L^{=n}) \geq \alpha$, i.e. $L \in \mathsf{FREQ}(\alpha)$. Since d does not strongly s-succeed on L, this ends the proof. $\square$