What risks lead to ruin

Narayana Santhanam Venkat Anantharam

***Abstract*— Insurance transfers losses associated with risks to the insurer for a price, the *premium*. Considering a natural probabilistic framework for the insurance problem, we derive a necessary and sufficient condition on loss models such that the insurer remains solvent despite the losses taken on. In particular, there need not be any upper bound on the loss—rather it is the structure of the model space that decides insurability.**

Insurance is a way of managing losses associated with risks—for example, floods, network outages, and earthquakes—primarily by transfering risk to another entity—the insurer, for a price, the *premium*. The insurer attempts to break even by balancing the possible loss that may be suffered by a few (risk) with the guaranteed payments of many (premium).

In 1903, Filip Lundberg [1] defined and formulated this scenario in its natural probabilistic setting as part of his thesis. In particular, Lundberg formulated a collective risk problem pooling together the risk of all the insured. There is an underlying risk model—a probability measure on loss sequences. Typically, the model itself is unknown, but can be imagined to belong to a known class of risk models. Suppose the insurance company sets some premium to be paid by the insured regularly—say, once at the beginning of every time interval. The losses incured by the insured will be of uncertain size in every time interval, governed according to the unknown underlying risk model. For a given class of risk models, how should the premiums be set so that the insurer compensates all losses in full, yet remains solvent?

Related to the insurance problem is the *pricing* problem that several researchers [2], [3] have considered for the Internet—these adopt, among other techniques, game theoretic principles to tackle the problem. A different approach, including that of Lundberg [1] involves studying the loss parametrically, using, for example, Poisson processes as the class of risk models. A more comprehensive theory of risk modeling has evolved [4] which incorporates several model classes for the loss other than Poisson processes, and which also includes some fat tailed distribution classes.

The later approach is very reminiscent of work in probability estimation, universal compression and prediction. Lately, there has been a lot of focus on choosing model classes for new applications such as language modeling, text compression, clustering and classification. Researchers have come up with new classes of models, *e.g.* [5], [6], as well as theoretical and practical approaches that balance the complexity of the model classes with their description power [7]. In particular, one would like to use a model class that is as general as possible, and is yet tractable.

This focus in compression literature is very pertinent to a new slew of scenarios for risk management. In settings like network outages, it is not clear what should constitute a reasonable risk model in the absence of usable information about what might cause the outages. If we are going to model these risks, how does one choose a class that is as general as possible, yet, one on which the insurer can set premiums to remain solvent?

A preliminary question is, then, what are necessary and sufficient conditions for a class of measures on infinite loss sequences to be *insurable*? In this paper, we provide a partial answer. If losses can be modelled as *i.i.d.* samples from a set $\mathcal{P}$ of distributions, in Theorems 4 and 6, we determine a necessary and sufficient condition on $\mathcal{P}$ for insurability.

We adopt the collective risk approach, namely, we abstract the problem without loss of generality to include just two players in the insurance game—the insured and the insurer. We denote the sequence of losses by $\{X_i\}_{i\geq 1}$, and we assume that $X_i \in \mathbb{N}$ for all $i \geq 1$, where $\mathbb{N}$ denotes the set of natural numbers, $\{0, 1, 2, , \ldots, \}$. $\mathcal{P}^\infty$ is a collection of measures on infinite length loss sequences. In this paper, we deal with only *i.i.d.* measures. Consequently, we denote by $\mathcal{P}$ the set of distributions on $\mathbb{N}$ obtained as single letter marginals of $\mathcal{P}^\infty$.

Let $\mathbb{N}^*$ be the collection of all finite length strings of natural numbers. The insurer's *scheme* Φ is a mapping from $\mathbb{N}^* \rightarrow \mathbb{R}^+$, and is interpreted as the premium demanded by the insurer from the insured after a loss sequence is observed. The insurer can observe the loss for a time prior to entering the insurance game. However, we require the insurer enters the game with probability 1 no matter what loss models are in force, and the insurer cannot quit once entered.

We adopt another abstraction without loss of generality: at any stage if the insurer is surprised by a loss bigger than the premium charged in *that* round, the insurer goes bankrupt. To see why this simplification does not involve any loss of generality, imagine the sequence of premiums set in the paper to represent the cummulative premium thus far.

To eliminate trivial schemes that do not enter the game at all, we require that for all $p \in \mathcal{P}$, the insurer enters the game with probability 1.

Definition 1. A class $\mathcal{P}^\infty$ of measures is insurable if $\forall\ \eta > 0$, there exists a premium scheme Φ such that $\forall$ $p \in \mathcal{P}^\infty$, $p(\Phi \text{ goes bankrupt }) < \eta$ and if, in addition, for all $p \in \mathcal{P}^\infty$, $\lim_{n\to\infty} p(\{X^n : \Phi(X^n) < \infty\}) = 1$. □

In Section II, we consider an example each of insurable and non-insurable classes.

I. RESULTS

We model the loss at each time by numbers in $\mathbb{N} = \{0, 1, \ldots\}$. A loss distribution is a distribution over $\mathbb{N}$, and let $\mathcal{P}$ be a set of loss distributions. $\mathcal{P}^\infty$ is the collection of *i.i.d.* measures over infinite sequences from $\mathbb{N}$ such that the set of marginals over $\mathbb{N}$ they induce is $\mathcal{P}$. We call $\mathcal{P}$ the set of *single letter marginals* of $\mathcal{P}^\infty$. Each $p \in \mathcal{P}$ is assumed to have finite support, and the *span* of $p \in \mathcal{P}$ is the highest number which has probability > 0 under p.

An insurer's *scheme* Φ is a mapping from $\mathbb{N}^* \to \mathbb{R}^+$, and is interpreted as the premium demanded by the insurer from the insured after a loss sequence is observed. For convenience, we assume $\Phi(x^n) = \infty$ on every sequence x^n of losses on which Φ has not entered.

Note however that the supremum over all distributions $p \in \mathcal{P}$ of the span of p need not bounded. Thus, we do not assume an upper bound on the possible loss.

The crux of insurability is this: we would like close distributions to be similar in their span. We first define what distributions are close, followed by what distributions have "similar" span. We will then specify the necessary and sufficient conditions for insurability.

A. Close distributions

Insurability of $\mathcal{P}^\infty$ depends on the neighborhoods of the probability distributions among its single letter marginals $\mathcal{P}$. The relevant "distance" between distributions in $\mathcal{P}$ that decides the neighborhood is

$$\mathcal{J}(p, q) = D\left(p \Big|\Big| \frac{p+q}{2}\right) + D\left(q \Big|\Big| \frac{p+q}{2}\right).$$

B. Cummulative distribution functions

In this paper, we phrase the notion of similarity in span in terms of the cummulative distribution function. Note that we are dealing with distributions over a discrete (countable) support, so a few non-standard definitions related to the cummulative distribution functions need to be clarified.

For our purposes cummulative distribution function of any distribution p is a function from $\mathbb{R} \to [0, 1]$, and will be denoted by F_p. We obtain F_p by first defining F_p on points in the support of p and the point at infinity. We define F_p for all other points by linearly interpolating between the values in the support of p.

Let $F_p^{-1}(1)$ be the smallest number y such that $F_p(y) = 1$, and let $F_p^{-1}(x) = 0$ for all $0 \le x < F_p(0)$. Note that for $0 \le x \le 1$, $F_p^{-1}(x)$ is now uniquely defined.

Two technical observations are in order since we are dealing discrete distributions. Consider a distribution p with support $\mathcal{A} \subset \mathbb{N}$. For $\delta > 0$, let (T for tail)

$$T_\delta = \{y \in \mathcal{A} : y \ge F^{-1}(1-\delta)\},$$

and let (H for head)

$$H_\delta = \{y \in \mathcal{A} : y \le 2F^{-1}(1-\delta/2)\}.$$

It is easy to see that

$$p(T_\delta) > \delta \text{ and } p(H_\delta) > 1 - \delta.$$

Suppose, for some δ, $F_p^{-1}(1-\delta) > 0$ and the premium is set to $F^{-1}(1-\delta)$, the probability under p of the loss exceeding the premium is $\ge \delta$. If the premium is set to $2F_p^{-1}(1-\delta/2)$, the probability that the loss exceeds the premium is $\le \delta$. We will use these observations in the proofs to follow.

C. Necessary and sufficient conditions for insurability

Existence of close distributions with very different spans is what kills insurability. A scheme could be "deceived" by some process $p \in \mathcal{P}^\infty$ into setting low premiums, while a close enough distribution lurks with a high loss. The conditions for insurability of $\mathcal{P}^\infty$ are phrased in terms of its single letter marginals $\mathcal{P}$.

Formally, a distribution p in $\mathcal{P}$ is *deceptive* if $\forall$ neighborhoods $\epsilon > 0$, $\exists\ \delta > 0$ so that no matter what function $f : \mathbb{R} \to \mathbb{R}$ is chosen, $\exists$ a bad distribution $q \in \mathcal{P}$ such that

$$\mathcal{J}(p, q) \le \epsilon$$

and

$$F_q^{-1}(1-\delta) > f(F_p^{-1}(1-\delta)),$$

In Sections III and IV, we show that a collection $\mathcal{P}^\infty$ of finite support *i.i.d.* process are insurable iff there are no deceptive distributions among its single letter marginals $\mathcal{P}$.

II. EXAMPLES

The set $\mathcal{N}^\infty$ is the class of *i.i.d.* processes whose single letter marginals have finite moment. Namely, $\forall p \in \mathcal{N}^\infty$, $E_p X_1 < \infty$.

Theorem 1. $\mathcal{N}^\infty$ is not insurable.

Proof Note that the loss measure that puts probability 1 on the all-0 zero sequences exists in $\mathcal{N}^\infty$. Since we consider only schemes that enter with probability 1 no matter what $p \in \mathcal{N}^\infty$ is in force, every insurer must therefore enter after seeing a finite number of zeros.

Fix any scheme. Denote the premiums charged at time i by $\Phi(X^i)$. Suppose the scheme enters the game after seeing N losses of size 0. To show that $\mathcal{N}^\infty$ is not insurable, we show that $\exists \eta > 0$ such that for all schemes Φ, $\exists p \in \mathcal{N}^\infty$ such that

$$p(\text{ } \Phi \text{ goes bankrupt }) \geq \eta.$$

Fix some $\delta = 1 - \eta$. Let ϵ be small enough that

$$(1-\epsilon)^N > 1 - \delta/2,$$

and let M be a number large enough that

$$(1-\epsilon)^M < \delta/2.$$

Note that since $1 - \delta/2 \geq \delta/2$, $N < M$.

Let L be greater than any of premiums charged by Φ for the sequences $0^N, 0^{N+1}, \ldots 0^M$. Let $p \in \mathcal{N}^\infty$ satisfy, for all i,

$$p(X_i) = \begin{cases} 1-\epsilon & \text{if } X_i = 0 \\ \epsilon & \text{if } X_i = L. \end{cases}$$

For the process p, the insurer is bankrupted on all sequences that contain loss L in between the N'th and M'th step. The sequences in question have probabilities (under p)

$$(1-\epsilon)^N \epsilon, (1-\epsilon)^{N+1}\epsilon, \ldots, (1-\epsilon)^{N+M-1}$$

and they also form a prefix free set. Therefore, summing up the geometric series and using the assumptions on ϵ above,

$$p(\text{ } \Phi \text{ is bankrupted }) \geq 1 - \delta/2 - \delta/2 = \eta. \qquad \Box$$

One can verify that every distribution in $\mathcal{N}^\infty$ is deceptive.

A monotone distribution on numbers satisfies for all i, probability of $i \geq$ probability of $i+1$. Let $\mathcal{M}^\infty$ be the set of all monotone *i.i.d.* loss processes with finite support. It will follow from Section III that

Theorem 2. $\mathcal{M}^\infty$ is not insurable. □

The above results mean that while insurability seems related to weak compressibility [8], it is not identical.

Consider $\mathcal{U}$, the collection of all uniform distributions over a finite support of form $\{m, \ldots, M\}$, with m and M being arbitrary. Let the losses be sampled *i.i.d.* from one of the distributions in $\mathcal{U}$—call these processes $\mathcal{U}^\infty$.

Theorem 3. $\mathcal{U}^\infty$ is insurable.

Proof If the threshold probability of ruin is η, set the premiums Φ as follows. For all sequences $\overline{x}$ with length $\leq \log \frac{1}{\eta} + 1$, $\Phi(\overline{x}) = \infty$. For all sequences longer than $\log \frac{1}{\eta} + 1$, the premium is twice the largest loss observed thus far. It is easy to see this scheme is bankrupted with probability $\leq \eta$. □

III. NECESSARY CONDITION FOR INSURABILITY

Note that according to the conventions adopted with defining cummulative distribution functions in Section I-B, if for a sequence x, $F_q^{-1}(1-\delta) > \Phi(x)$, the scheme Φ will be bankrupted with probability $\geq \delta$ in the next step.

$\mathcal{P}^\infty$ is a set of *i.i.d.* measures over infinite sequences from $\mathbb{N}$, and let $\mathcal{P}$ denote the collection of their single letter marginals.

Theorem 4. If $\mathcal{P}^\infty$ is insurable, then no $p \in \mathcal{P}$ is deceptive.

Proof We prove the contrapositive of the theorem: if some $p \in \mathcal{P}$ is deceptive, then $\mathcal{P}^\infty$ is not insurable. Lemma 7 necessary for the following proof can be found in the Appendix.

In what follows, we will denote by p (or q) both a measure $\in \mathcal{P}^\infty$ as well as the single letter marginal distribution $\in \mathcal{P}$. The context will clarify which of the two is meant if need be.

To show the contraposition, let $p \in \mathcal{P}$ denote a deceptive distribution. Namely, for all $\epsilon > 0$, $\exists \delta > 0$ such that $\forall f_p : \mathbb{R} \to \mathbb{R}$ ϵ, $\exists$ q satisfying $\mathcal{J}(p,q) < \epsilon$

$$F_q^{-1}(1-\delta) > f_p(F_p^{-1}(1-\delta)).$$

We prove that $\mathcal{P}^\infty$ is not insurable by fixing some $\eta > 0$ and finding for each scheme Φ, a process $q \in \mathcal{P}^\infty$ such that

$$q(\text{ } \Phi \text{ goes bankrupt }) \geq \eta.$$

Pick $0 \leq \alpha < h^{-1}(\frac{1}{2})$ and let $0 < \eta < (1 - 2h(\alpha))\big(1 - \frac{1}{e}\big)$, the bounds chosen in order to satisfy the technical requirements of the proof of Lemma 7.

Consider any Φ that enters on p with probability 1. For all n, let

$$R_n = \{x^n : \Phi(x^n) < \infty\}$$

be the set of sequences of length n on which Φ has entered and let N be a number such that

$$p(R_N) > 1 - \alpha.$$

Furthermore, let

$$R_{p,n} = \{x^n \in R_n : p(x^n) > 0\}.$$

We will prove that $\exists q \in \mathcal{P}^\infty$ such that

$$q(\text{ } \Phi \text{ goes bankrupt }) \geq (1 - \frac{2}{N} - 2h(\alpha))\bigg(1 - \frac{1}{e}\bigg). \tag{1}$$

By picking N large enough, we can make this probability of error arbitrarily close to $(1-2h(\alpha))\big(1-\frac{1}{e}\big)$.

For all $0<\delta'<1$,

$$f_p(F_p^{-1}(1-\delta')) \overset{\text{def}}{=} \max_{\substack{x^i: N\le i\le N+\lceil \frac{1}{\delta'}\rceil \\ x^i\in R_{p,i}}} \Phi(x^i).$$

and let

$$f_p(0) = \sup_{\substack{0<\delta'\le 1\\ F_p^{-1}(1-\delta')=0}} \max_{\substack{x^i: N\le i\le N+\lceil \frac{1}{\delta'}\rceil \\ x^i\in R_{p,i}}} \Phi(x^i).$$

Set $\epsilon=\frac{1}{N^2}$. Since p is deceptive, for some $\delta>0$, there exists some distribution q satisfying

$$\mathcal{J}(p,q)<\epsilon=\frac{1}{N^2},$$

while

$$F_q^{-1}(1-\delta) > f_p(F_p^{-1}(1-\delta)).$$

For the distributions $p,q\in\mathcal{P}$ above, $\mathcal{J}(p,q)<\frac{1}{N^2}$. Applying Lemma 7 to distributions over length-n sequences induced by the measures $p,q\in\mathcal{P}^\infty$ corresponding to the distributions above,

$$q(R_{p,N}) \ge 1-\frac{2}{N}-2h(\alpha),$$

namely, Φ has entered with probability (under q) at least $1-\frac{2}{N}-2h(\alpha)$ for length N sequences. Since the insurer cannot quit once it has entered, the scheme has entered with probability (under q) at least $1-\frac{2}{N}-2h(\alpha)$ for all n length sequences where $n>N$. Namely for all $n\ge N$,

$$q(R_n)\ge 1-\frac{2}{N}-2h(\alpha),$$

We have set things up so that Φ is bankrupted with the δ tail of q. To see that the δ tail of q bankrupts Φ, note that

$$\begin{aligned} F_q^{-1}(1-\delta) &\ge f_p\big(F_p^{-1}(1-\delta)\big)\\ &= \max_{\substack{X^i\in R_{p,i}\\ N\le i\le N+\lceil\frac{1}{\delta}\rceil}} \Phi(X^i). \end{aligned} \tag{2}$$

Namely any loss in the δ tail of q is greater than the highest premium Φ charges for any sequence in $R_{p,i}$ with i between N and $N+\lceil\frac{1}{\delta}\rceil$.

For convenience, let $M=\lceil\frac{1}{\delta}\rceil$. Let the distribution q be in force. Conditioned on any sequence in $R_{p,i}$ with i between N and $N+M-1$, the scheme Φ fails with probability (under q) at least δ in the following step.

A sequence on which Φ has entered, but such that Φ has not been bankrupted on any the sequence's prefixes is a *surviving* sequence.

Consider a surviving sequence $\overline{x}\in R_{p,N}$ in the support of p at level N. Given $\overline{x}$, let the conditional probability that Φ is bankrupted in the following step be δ_N. From (2), and as mentioned in the note before the Theorem, conventions adopted in the definitions of cummulative distribution functions in Section I-B and Equation (2) imply that we have $\delta_N\ge\delta$.

Now, given $\overline{x}$, the conditional probability that Φ is bankrupted in at most two further steps is,

$$\delta_N+(1-\delta_N)\delta_{N+1}\ge\delta+(1-\delta)\delta,$$

where δ_{N+1} is interpreted as the weighted average (over surviving length-$(N+1)$ suffixes of $\overline{x}$) of the probability that Φ goes bankrupt in step $N+2$. In particular, note that the inequality above holds because $\delta_{N+1}\ge\delta$ thanks to Equation (2).

Similarly, given a sequence $\overline{x}$, the probability that Φ is bankrupted on suffix sequences of $\overline{x}$ with length between N and $N+M$ is

$$\delta_N+(1-\delta_N)\delta_{N+1}+\ldots+(1-\delta_N)\delta_{N+M}\prod_{i=N+1}^{N+M}(1-\delta_i)$$

for some $\delta_N,\delta_{N+1},\ldots,\delta_{N+M}$, all of which are $\ge\delta$.

Let q_1 be the probability of all survivors in $R_{p,N}$, and q_2 be the probability of all sequences in $R_{p,N}$ where Φ has already been bankrupted. Therefore $q_1+q_2=q(R_{p,N})$.

Therefore Φ is bankrupted with probability

$$\begin{aligned}
&\ge q_2+q_1\left(\delta_N+\ldots+\delta_{N+M}\prod_{i=N}^{N+M}(1-\delta_i)\right)\\
&= q_2+q_1\times\\
&\big(\delta_{_N}+\overline{\delta_{_N}}\big(\delta_{_{N+1}}+\overline{\delta_{_{N+1}}}\big(\ldots\big(\delta_{_{N+M-1}}+\overline{\delta_{_{N+M-1}}}\delta_{_{N+M}}\big)\big)\big)\big)\\
&\ge q_2+q_1\big(\delta+(1-\delta)\delta+\ldots+(1-\delta)^M\delta\big)\\
&= q_2+q_1\Big(1-(1-\delta)^{\lceil 1/\delta\rceil}\Big)\\
&\ge q(\varphi_N)\Big(1-(1-\delta)^{\lceil 1/\delta\rceil}\Big)\\
&\ge \left(1-\frac{1}{N}-h(\alpha)\right)\Big(1-(1-\delta)^{\lceil 1/\delta\rceil}\Big),
\end{aligned}$$

where $\overline{\delta}$ stands for $1-\delta$. The theorem follows. □

Corollary 5. The set of monotone distributions is not insurable. □

IV. SUFFICIENT CONDITION FOR INSURABILITY

The necessary condition in Section III is also sufficient for insurability. As per conventions adopted in Section I-B, recall that a distribution q is bankrupted with probability $<\delta$ if the premium set is at least $2F^{-1}(1-\delta/2)$.

$\mathcal{P}^\infty$ is a set of *i.i.d.* measures over infinite sequences from $\mathbb{N}$, and let $\mathcal{P}$ denote the collection of their single letter marginals.

Theorem 6. If no $p\in\mathcal{P}$ is deceptive, then $\mathcal{P}^\infty$ is insurable.

Proof The proof is constructive. For any $0 < \eta < 1$, we obtain a scheme Φ such that for all $p \in \mathcal{P}^\infty$, $p(\Phi$ goes bankrupt $) < \eta$. For convenience, we will represent types by probability distributions and the sequence length. Namely, a string 123111 has a length-6 type $(4/6, 1/6, 1/6)$.

Since no $p \in \mathcal{P}$ is deceptive, it follows that for all $p \in \mathcal{P}$, $\exists \epsilon_p > 0$ such that for all $\delta > 0$, $\exists f_{p,\delta} : \mathbb{R} \to \mathbb{R}$ so that all q with $\mathcal{J}(p,q) < \epsilon_p$ satisfy

$$F_q^{-1}(1-\delta) < f_{p,\delta}(F_p^{-1}(1-\delta)).$$

We say that ϵ_p is the *reach* of p.

For $p \in \mathcal{P}$, define

$$B_p = \{q : \mathcal{J}(p,q) \le \epsilon_p\},$$

which will play the role of the set of distributions which will not be bankrupted by setting premiums assuming p is in force. Furthermore, for $p \in \mathcal{P}$, let

$$I_p = \left\{ q : |p-q|_1 \le \frac{{\epsilon_p}^2 (\ln 2)^2}{16} \right\}.$$

For large n, I_p will play the role of the set of length-n types on which the proposed scheme Φ will have entered, in order to ensure that Φ enters with probability 1 on strings generated by p. Note that if ϵ_p is small enough, $I_p \subset B_p$.

In an abuse of notation, let $2^{\mathbb{N}}$ be the set of all finite subsets of $\mathbb{N}$. Note that $2^{\mathbb{N}}$ is countable. Hence, associate with each finite subset $A \subset \mathbb{N}$, its index $i(A)$.

We now describe the scheme Φ.

a) Preliminaries: Consider a length-n sequence x on which Φ has not entered thus far. Let the type of the sequence be q, and let A denote the set of symbols in the sequence. Let

$$\mathcal{P}'_q = \{p' \in \mathcal{P} : q \in I_{p'}\}$$

be the set of distributions in $\mathcal{P}$ which potentially capture q. If $\mathcal{P}'_q \ne \emptyset$, we will refine set of distributions that could capture q is further refined to $\mathcal{P}_q \subset \mathcal{P}'_q$, as by requiring (3) to hold as follows. For $p' \in \mathcal{P}'_q$, let the reach of p' be $\epsilon_{_{p'}}$, and define

$$D_{_{p'}} \stackrel{\text{def}}{=} \frac{\epsilon_{_{p'}}^4 (\ln 2)^3}{512}.$$

If n satisfies

$$2^{-nD_{p'}} \le \frac{1}{(n+1)^{|A|}} \frac{\eta}{2i(A)^2 n^2} \frac{36}{\pi^4}, \tag{3}$$

then we place $p' \in \mathcal{P}_q$.

b) Description of Φ: If $\mathcal{P}_q = \emptyset$, the scheme does not enter yet. If $\mathcal{P}_q \ne \emptyset$, let p_q denote any distribution in $\mathcal{P}_q$.

All suffixes of the sequence are said to be *trapped* by p_q—namely, premiums will be based on p_q. The premium assigned for a length-m sequence trapped by p_q is

$$2 f_{p_q, \frac{6\eta/2}{2\pi^2 n^2}} \left(F_{p_q}^{-1} \left(1 - \frac{6\eta/2}{2\pi^2 n^2} \right) \right).$$

c) Φ enters with probability 1: First, we verify that the scheme enters with probability 1, no matter what distribution $p \in \mathcal{P}$ is in force, observe that the alphabet of the sequence eventually stablizes since $\mathcal{P}$ contains finite support distributions. Subsequently, for all large enough n, and types in I_p, it follows that (3) will hold and the scheme enters with probability 1.

d) Probability of bankruptcy $\le \eta$: We now analyze the scheme. Consider any $p \in \mathcal{P}$. Among sequences on which Φ has entered, we will distinguish between those that are in *good* traps and those in *bad* traps. If x is trapped by p' such that $p \in B_{p'}$, p' is a good trap. Conversely, if $p \notin B_{p'}$, p' is a bad trap.

(Good traps) Suppose a length-n sequence x^n is in a good trap, namely, it is trapped by a distribution p' such that $p \in B_{p'}$. Recall that the premium assigned is

$$2 f_{p', \delta'/2} \left(F_{p'}^{-1} (1 - \delta'/2) \right)$$

where $\delta' = 6\eta/2\pi^2 n^2$, which in turn is greater than $2F_p^{-1}(1-\delta'/2)$. Therefore, the scheme is bankrupted with probability at most $\delta' = 6\eta/2\pi^2 n^2$ in the next step. Therefore, sequences in good traps contribute at most $\eta/2$ to the probability of bankruptcy.

(Bad traps) We will show that the probability with which sequences generated by p fall into bad traps $\le \eta/2$. Pessimistically, the conditional probability of bankruptcy given a sequence falls into a bad trap is 1. Thus the contribution to bankruptcy by sequences in bad traps is at most $\eta/2$.

Let q be a length-n type with alphabet A. Say it is trapped by $\tilde{p}$ with reach $\tilde{\epsilon}$ such that $p \notin B_{\tilde{p}}$, we obtain from Lemma 9 that

$$\mathcal{J}(p,q) \ge \frac{\tilde{\epsilon}^2 \ln 2}{16},$$

and therefore,

$$D(q||p) \ge \frac{1}{2\ln 2} |p-q|_1^2 \ge \mathcal{J}^2(p,q) \frac{\ln 2}{2} = \frac{\tilde{\epsilon}^4 (\ln 2)^3}{512}$$

where the second inequality follows from Lemma 8. Noting that the probability of the length-n type q is at most $2^{-nD(q||p)}$, we use (3) to conclude that for $p \in \mathcal{P}^\infty$, if q falls in a bad trap,

$$p(\text{ length-}n \text{ type } q\ ) \le \frac{1}{(n+1)^{|A|}} \frac{\eta}{2i(A)^2 n^2} \frac{36}{\pi^4}.$$

There are at most $(n+1)^{|A|}$ length-n types with alphabet A. Therefore, the probability of sequences falling into bad traps

$$\le \sum_{n\ge 1}\sum_{A\in 2^{\mathbb{N}}} \frac{1}{(n+1)^{|A|}}\frac{\eta}{2i(A)^2n^2}\frac{36}{\pi^4} \le \eta/2$$

since

$$\sum_{A\in 2^{\mathbb{N}}} \frac{1}{i(A)^2} = \sum_{n\ge 1}\frac{1}{n^2} = \frac{\pi^2}{6},$$

The theorem follows. □

V. APPENDIX

Lemma 7. For distributions p and q over some support $\mathcal{A}$, with $\mathcal{J}(p,q) \le \epsilon$. For some $S \subset \mathcal{A}$ and $\alpha < 1 - \ln 2 = .30685$, if $p(S) \ge 1-\alpha$, then

$$q(S) \ge 1 - 2\epsilon - 2h(\alpha).$$

Proof Let $p(S)$ ($q(S)$ respectively) denote a binary distribution, the two probabilities of which correspond to $[p(S), 1-p(S)]$ ($[q(S), 1-q(S)]$ respectively). Now,

$$\begin{aligned}\epsilon \ge \mathcal{J}(p,q) &\ge D\left(p(S)||\frac{q(S)+p(S)}{2}\right)\\ &= p(S)\log\frac{2}{p(S)+q(S)} - h(p(S))\\ &\ge (1-\alpha)\log\frac{2}{p(S)+q(S)} - h(\alpha).\end{aligned}$$

The last inequality follows because the condition $p(S) \ge 1-\alpha \ge \frac{1}{2}$ implies $h(p(S)) \le h(1-\alpha)$. Therefore,

$$\log\frac{2}{p(S)+q(S)} \le \frac{h(\alpha)+\epsilon}{(1-\alpha)},$$

implying that

$$\begin{aligned}\frac{1+q(S)}{2} \ge \frac{p(S)+q(S)}{2} &\ge 2^{-\frac{h(\alpha)+\epsilon}{1-\alpha}}\\ &\ge 1-(h(\alpha)+\epsilon),\end{aligned}$$

where the last inequality follows since $\ln 2 \le 1-\alpha$. □

Lemma 8. For any two distributions p, q,

$$\frac{1}{4\ln 2}|p-q|_1^2 \le \mathcal{J}(p,q) \le \frac{1}{\ln 2}|p-q|_1$$

and for any three distributions p, q, r,

$$\mathcal{J}(p,q) + \mathcal{J}(q,r) \ge \mathcal{J}^2(p,r)\frac{\ln 2}{8}. \qquad \square$$

Proof The lower bound follows since

$$D\left(p||\frac{p+q}{2}\right) \ge \frac{1}{2\ln 2}\frac{1}{4}|p-q|_1^2$$

and similarly for $D\big(q||\frac{p+q}{2}\big)$. The upper bound follows since

$$\begin{aligned}\mathcal{J}(p,q)\ln 2 &\le \sum_x p(x)\left(\frac{p(x)-q(x)}{p(x)+q(x)}\right) + \sum_x q(x)\left(\frac{q(x)-p(x)}{p(x)+q(x)}\right)\\ &\le |p-q|_1.\end{aligned}$$

The triangle-like inequality follows from routine manipulations. □

Lemma 9. Suppose a type q is trapped by p_0 with reach ϵ_0. For all $p \in \mathcal{P}$ with $\mathcal{J}(p,p_0) \ge \epsilon_0$,

$$\mathcal{J}(p,q) \ge \frac{\epsilon_0^2 \ln 2}{16}.$$

Proof Since q is trapped by p_0, it follows that

$$|p_0 - q| \le \frac{\epsilon_0^2(\ln 2)^2}{16}$$

and therefore

$$\mathcal{J}(p_0,q) \le \frac{\epsilon_0^2 \ln 2}{16}.$$

Furthermore, since $\mathcal{J}(p,p_0) \ge \epsilon_0$, Lemma 8 then implies that

$$\mathcal{J}(p,q) + \frac{\epsilon_0^2\ln 2}{16} \ge \mathcal{J}(p,q) + \mathcal{J}(p_0,q) \ge \frac{\epsilon_0^2\ln 2}{8}. \quad \square$$

ACKNOWLEDGMENTS

N. Santhanam was supported by a startup grant from the University of Hawaii and NSF Grant CCF-1018984. V. Anantharam was supported by the ARO MURI grant W911NF-08-1-0233, ”Tools for the Analysis and Design of Complex Multi-Scale Networks”, by the NSF grants CCF-0500234, CCF-0635372, CNS-0627161, and CNS-0910702, by Marvell Semiconductor Inc., and by the U. C. Discovery program.

REFERENCES

[1] K. Englund and A. Martin-Löf. *Statisticians of the Centuries*, chapter on Ernst Filip Oskar Lundberg, pages 308–311. New York: Springer, 2001.

[2] S. Shakkottai and R. Srikant. Economics of network pricing with multiple ISPs. *IEEE/ACM Transactions on Networking*, pages 1233–1245, Dec 2006.

[3] A. M. Odlyzko. Paris metro pricing for the internet. In *Proceedings of the ACM Conference on Electronic Commerce*, pages 140–147, 1999.

[4] H. Cramer. Historical review of Filip Lundberg's work on risk theory. *Skandinavisk Aktuarietidskrift (Suppl.)*, 52:6–12, 1969. Reprinted in The Collected Works of Harald Cramér edited by Anders Martin-Löf, 2 volumes Springer 1994.

[5] F. M. J. Willems, Y. M. Shtarkov, and Tj. J. Tjalkens. The context-tree weighting method: basic properties. *IEEE Transactions on Information Theory*, 41(3):653–664, 1995.

[6] A. Orlitsky, N.P. Santhanam, and J. Zhang. Universal compression of memoryless sources over unknown alphabets. *IEEE Transactions on Information Theory*, 50(7):1469—1481, July 2004.

[7] A. Barron, J. Rissanen, and B. Yu. The minimum description length principle in coding and modeling. *IEEE Transactions on Information Theory*, 44(6):2743—2760, oct 1998.

[8] J.C. Kieffer. A unified approach to weak universal source coding. *IEEE Transactions on Information Theory*, 24(6):674—682, nov 1978.