

FRAUD RISK ASSESSMENT WITHIN BLOCKCHAIN TRANSACTIONS

PIERRE-O. GOFFARD,* *University of California, Santa Barbara*

Abstract

The probability of successfully spending twice the same bitcoins is considered. A double-spending attack consists in issuing two transactions transferring the same bitcoins. The first transaction, from the fraudster to a merchant, is included in a block of the public chain. The second transaction, from the fraudster to himself, is recorded in a block that integrates a private chain, exact copy of the public chain up to substituting the fraudster-to-merchant transaction by the fraudster-to-fraudster transaction. The double-spending hack is completed once the private chain reaches the length of the public chain, in which case it replaces it. The growth of both chains are modelled by two independent counting processes. The probability distribution of the time at which the malicious chain catches up with the honest chain, or, equivalently, the time at which the two counting processes meet each other, is studied. The merchant is supposed to await the discovery of a given number of blocks after the one containing the transaction before delivering the goods. This grants a head start to the honest chain in the race against the dishonest chain.

Keywords: Bitcoin blockchain; double-spending problem; risk theory; order statistic point processes; renewal processes; boundary crossing problems

2010 Mathematics Subject Classification: Primary 60G55

Secondary 60G40;12E10

1. Introduction

Bitcoin is a decentralized peer-to-peer (P2P) payment system that relies on proof of work (PoW). Electronic payments are performed by issuing transactions that transfers bitcoins (BTCs) between bitcoin peers. These transactions are broadcast to a network of bitcoin miners. These miners will compete to solve a cryptographic puzzle in order to build a block that contains the pending transactions. The first miner to solve the problem receives a certain number of BTCs, which is agreed upon by everyone in the network. At the time of the writing, this bounty is 12.5 BTCs; this value is halved every 210,000 blocks. The block is then included in the *blockchain* which plays the part of a public ledger recording all the transactions between bitcoin peers. Once a transaction enters the blockchain, it is considered validated. The only way to reverse the process, and, for instance, replace this transaction by another one, is to redo the work of the associated block and all the subsequent blocks. The blockchain allows in theory prevention from double spending the same BTCs. A double-spending attack consists in buying goods from a vendor and transferring the same bitcoins to oneself. Two conflicting transactions exist then in the network. The buyer-to-vendor transaction is included in the blockchain by

Received 27 February 2018; revision received 5 March 2019.

* Postal address: Institut de Science Financière et d'Assurances, Université Lyon 1, 50 Avenue Tony Garnier, 69007 Lyon, France.

the honest miners, while a group of colluding miners work on their own private branch, exact replication of the principal chain up to substituting the buyer-to-vendor transaction with the buyer-to-buyer transaction. In the presence of two versions of the blockchain, the network always opts for the longest because more computational effort has been put into it. If the conspiring miners' chain ever becomes as long as the honest chain, it will replace it. The double spending is then successful.

Nakamoto stressed in his whitepaper [27] that a successful double-spending attack is rather unlikely as long as the pool of honest miners retains the majority of the computing power. The vendor is advised to wait for a certain number of blocks, say $k \in \mathbb{N}$, to be added to the chain before delivering the goods. Assuming that, in the meantime, the attackers manage to discover $l < k$ blocks, then the honest chain is ahead by $z = k - l$. The $l \geq k$ case implies that double spending occurs right away. The derivation of the probability of a successful double-spending attack relies on an analogy with the one-sided gambler's ruin problem. Namely, the forthcoming block belongs to the honest chain with probability p or to the malicious chain with probability $q = 1 - p$. The difference between the length of the chains is then a random walk $\{Z_n, n \in \mathbb{N}\}$ on \mathbb{Z} defined as

$$Z_n = z + Y_1 + \cdots + Y_n \quad \text{for } n \in \mathbb{N},$$

where the Y_i are the independent and identically distributed (i.i.d.) steps of the random walk. Assuming that the honest miners have more resources implies that $p > q$, the probability that the malicious chain ever catches up with the honest one, given it is z blocks behind, is $(q/p)^z$. For a full treatment of the gambler's ruin problem, the reader is referred to [2].

The aim of this work is to refine the model underlying the double-spending problem. Counting processes are introduced to keep track of the number of blocks in the competing chains. These processes are generated by sequences of arrival times whose probability distribution reflect the block discovery frequency and the distribution of the computing power among honest and malicious miners. The probability distribution of the time at which the malicious chain overtakes (if it ever does) the honest chain is studied. Note that the probability mass function (PMF) of the stopping time

$$\tau_z = \inf\{n \in \mathbb{N}; Z_n = 0\}$$

in Nakamoto's framework is a consequence of the first hitting time theorem with

$$\mathbb{P}[\tau_z = n] = \frac{z}{n} \mathbb{P}[Z_n = 0] \quad \text{for } n \geq z;$$

see, for instance, [35, Theorem 1] and the references therein.

Let $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ be two independent counting processes governing the block arrival over time in the honest and the malicious chains, respectively. Assume that the honest chain is $z \geq 1$ blocks ahead of the malicious chain at $t = 0$. Consider the stopping time

$$\tau_z = \inf\{t \geq 0; M(t) = z + N(t)\},$$

at which the double-spending attack is successful. Denote by $\{T_k, k \geq 1\}$ and $\{S_k, k \geq 1\}$ the block arrival times in the honest and malicious chains, respectively. In Figure 1 we illustrate the race between the two processes. The distribution of τ_z is studied for different sets of assumptions over the counting processes $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$.

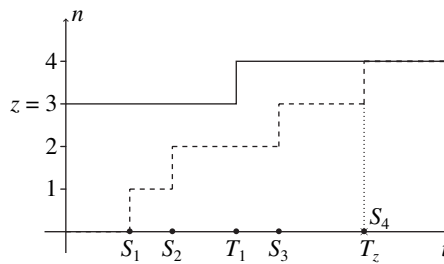


FIGURE 1: Time until the double-spending attack is completed. The solid line represents the length of the honest chain $\{z + N(t), t \geq 0\}$ and the dashed line represents the length of the malicious chain $\{M(t), t \geq 0\}$.

In Section 2 and 3, the length of the honest chain $\{z + N(t), t \geq 0\}$ is driven by an order statistic point process (OSPP), that is, provided that $N(t) = n$, the jump times T_1, \dots, T_n have the same distribution as the order statistics of a sample of n i.i.d.random variables concentrated on $[0, t]$ with distribution function F_t . In Section 2, the probability density function (PDF) of τ_z is derived in terms of Abel–Gontcharov (A-G) polynomials when the length of the malicious chain $\{M(t), t \geq 0\}$ is a renewal process (i.e.generated by i.i.d.interarrival times). In Section 3, the survival function (SF) of τ_z is expressed in terms of Appell polynomials in the case where $\{M(t), t \geq 0\}$ is an OSPP.

The probability of a successful double-spending attempt, defined by $\mathbb{P}[\tau_z < \infty]$, is further considered. An upper bound is derived in Section 4 when both $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are renewal processes. An exact expression is obtained when $\{N(t), t \geq 0\}$ is a Poisson process.

The formulae derived in this work hold for a fixed value z . In practice, because the length of the dishonest chain is unknown to the vendor at the time of the shipping, the delay should be modelled as an integer-valued random variable Z . Nakamoto [27] let Z be Poisson distributed with parameter zq/p (that is, the average number of blocks mined in the malicious chain when k blocks have been mined in the honest chain). Rosenfeld [32] stressed in his analysis that the right distribution is the negative binomial distribution. The distribution of the double-spending time follows from inserting the formulae given below in the law of total probability after conditioning upon the possible values of Z . The determination of the distribution of Z might not be an easy task in the general case and should be considered for further investigation.

Nakamoto [27] assumed that double spending occurs when the length of the malicious chain reaches exactly the length of the honest chain. The situation in which two chains of the same length coexist in the network is called a fork. The situation resolves as soon as a block is added to either one of the two chains. Some authors, including Rosenfeld [32] for instance, argued that the double-spending time should be defined as $\tau_z^+ = \inf\{t \geq 0; M(t) = N(t) + z + 1\}$. The results obtained in this paper apply in this context by noting that $\tau_z^+ = \tau_{z+1}$ almost surely.

Nakamoto [27] did not state explicitly that the block arrival is dictated by a homogeneous Poisson process. However, the probability of a successful double-spending attack, as we will see later, when the length of the public and private chains are governed by two Poisson processes of intensity λ and μ , is given by $(\mu/\lambda)^z$. Hence, the intensities play the same role as p and q as they reflect the hashrate of the miners. Every single result given in this work holds when the rival chains are modelled by a homogeneous Poisson process because the homogeneous Poisson process is the one and only renewal process enjoying the order statistic property. The formulae, which may seem involved in the general cases, simplify when the

arrival of blocks is Poisson. Now, is it worth considering more sophisticated models to track the growth of the blockchains?

A statistical study over the interblock times was conducted in a recent work of Bowden *et al.* [8]. The authors collected the timestamp information in the header of the blocks. As pointed out in [8], the timestamp information cannot be readily used, and preprocessing it represents quite a challenge in itself. The data after preprocessing is available on Rhys Bowden GitHub repository [7]. The empirical mean of the interblock time is 9.41 minutes while the standard deviation is 11.05 minutes. The high variance of the inter-block time is a known flaw which impedes the consistent flow of validated transactions. Renewal processes allow us to capture higher-order moments by choosing a more flexible probability distribution to fit the interblock time data. A propagation delay is sometimes added to the time required for the creation of a block. A newly discovered block is appended to the chain only once the word about that block has been spread to the entire network. If it is accepted that the block mining time is an exponential random variable, the actual time at which the block integrates the chain is an exponential random variable perturbed by another nonnegative noise. The result may or may not be exponentially distributed. One of the many takeaways of Bowden *et al.* [8] was that the rate at which blocks are discovered varies over time according to the adjustment of the cryptopuzzles difficulty. The authors of [8] proposed different models allowing for a time-dependent discovery rate with deterministic or even random difficulty adjustment. When the difficulty adjustment is deterministic then a nonhomogeneous Poisson process is suitable. This is fortunate as the nonhomogeneous Poisson process is a particular instance of OSPP. The point processes having the order statistic property were characterized a while ago by Puri [31]. The OSPPs are either mixed Poisson processes up to a timescale transformation or mixed sample processes. This class encompasses classical point processes such as the mixed Poisson process, the nonhomogeneous Poisson process, the linear birth process with immigration, and the linear death process.

From a mathematical standpoint, this work resembles an early work of Picard and Lefèvre [30] where the probability distribution of the first *rendez-vous* time between two counting processes was derived in terms of the Appell and A-G polynomials. The definition of the stopping time is slightly different in the case considered here. Plus, the reasoning differs as it relies extensively on the order statistic property and the connection between the aforementioned families of polynomials and the order statistics joint distribution. It is more in the spirit of Goffard and Lefèvre [17] where the crossing problem of an OSPP through a moving boundary was treated. These arguments are inspired from risk theory when solving the ruin problem in ordered risk models; see, e.g. [9], [15], [18], [20], and [24]. To the best of my knowledge, the closest link to queueing theory is the single-server queue with either work or customer removal introduced by Gelenbe *et al.* [13] and further considered in [5], [19], and [21] for different queues. The related risk process includes lump addition; see [6]. Perry *et al.* [28], [29] used renewal-type arguments to study the distribution of boundary crossing times of the difference between two Poisson processes and linear boundaries. Regarding the evaluation of the probability of a successful double-spending attack, $\mathbb{P}[\tau_z < \infty]$, the first step consists of swapping the role of time and space. Namely, a correspondence is established between the ruin times of two risk models with inverted characteristics. This trick is now standard in risk theory; see, for instance, [4], [10], [18], [26], and [34]. A classical martingale approach allows us to derive an expression of the probability of successfully spending twice the same BTCs.

The rest of the paper is organized as follows. In Section 2 a formula for the PDF of τ_z when $\{N(t), t \geq 0\}$ is an OSPP and $\{M(t), t \geq 0\}$ is a renewal process is derived in terms

of A-G polynomials. In Section 3 a formula for the SFof τ_z when both $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are OSPPs is provided in terms of Appell polynomials. Section 4 is concerned with the probability of the double-spending attack ever being successful. Section 5 is devoted to numerical illustrations.

2. The PDF of the double-spending time

In this section, the length of the honest chain $\{z + N(t), t \geq 0\}$ is governed by an OSPP. Its jump times, provided that $N(t) = n$, have the same joint distribution as a vector of order statistics. Namely, it holds that

$$[T_1, \dots, T_n \mid N(t) = n] \stackrel{D}{=} (V_{1:n}, \dots, V_{n:n}),$$

where ‘ $\stackrel{D}{=}$ ’ stands for equality in distribution and $V_{1:n}, \dots, V_{n:n}$ correspond to the order statistics of n i.i.d.random variables having a cumulative distribution function (CDF) $F_t(s)$, for $0 \leq s \leq t$. The length of the malicious chain is a renewal process generated by a sequence of i.i.d.interarrival times $\{\Delta_k^S, k \geq 1\}$ having a PDFdenoted by f_{Δ^S} . The sequence of arrival times $\{S_n, n \in \mathbb{N}\}$, with the convention $S_0 = 0$, corresponds to the partial sums of the interarrival times sequence. The PDFof S_n , for $n \in \mathbb{N}$, is given by

$$f_{S_n}(t) = f_{\Delta^S}^{*n}(t) \quad \text{for } t \geq 0,$$

where f^{*n} denote the n -fold convolution of f_{Δ^S} with itself. Let $z \geq 1$ be an integer, the following result gives a formula for the PDFof

$$\tau_z = \inf\{t \geq 0; M(t) = N(t) + z\},$$

the time at which the double-spending attack is completed.

Theorem 1. *If $\{N(t), t \geq 0\}$ is an OSPP and $\{M(t), t \geq 0\}$ is a renewal process, then the PDFof τ_z is given by*

$$f_{\tau_z}(t) = \mathbb{E}[(-1)^{N(t)} h_{N(t)}(t, z) f_{\Delta^S}^{*[N(t)+z]}(t)], \quad t \geq 0, \tag{1}$$

where

$$h_n(t, z) = \mathbb{E}\{G_n[0 \mid F_t(S_z), \dots, F_t(S_{n+z-1})] \mid S_{n+z} = t\}, \tag{2}$$

and $G_n(0 \mid \cdot)$ is an A-G polynomial such as defined in Appendix A.

Proof. The event $\{\tau_z \in (t, t + dt)\}$ for $t \geq 0$ corresponds to the exact time at which the double-spending attack is successful as the malicious chain takes over the honest chain. At time $t = 0$, the honest chain is ahead by $z \geq 1$ blocks. Assuming that later, at time $t > 0$, the honest miners manage to add $N(t) = n \in \mathbb{N}$ blocks to the chain, then the malicious chain must be of length $M(t^-) = n + z - 1$ at some time $t^- < t$ and jumps to the level $n + z$ exactly at t . Conditioning over the values of $\{N(t), t \geq 0\}$ yields

$$\{\tau_z \in (t, t + dt)\} = \bigcup_{n=0}^{+\infty} \{\tau_z \in (t, t + dt)\} \cap \{N(t) = n\}.$$

In the case where $N(t) = 0$, the only requirement is that the z th jump of $\{M(t), t \geq 0\}$ occurs at time t . It then follows that

$$\{\tau_z \in (t, t + dt)\} \cap \{N(t) = 0\} = \{S_z \in (t, t + dt)\} \cap \{N(t) = 0\}$$

and, consequently,

$$f_{\tau_z | N(t)=0}(t) = f_{\Delta S}^{*z}(t), \quad t \geq 0,$$

where $f_{\tau_z | N(t)=0}(t)$ denotes the conditional PDF of τ_z given that $N(t) = 0$. On the set $\{N(t) \geq 1\}$ we need to make sure that $\{M(t), t \geq 0\}$ behaves properly by constraining its jump times so that it does not reach $N(s) + z$ at any time $s < t$ and performs the $(n + z)$ th jump at t . Hence, it holds that

$$\{\tau_z \in (t, t + dt)\} \cap \{N(t) \geq 1\} = \bigcup_{n=1}^{+\infty} \bigcap_{k=1}^n \{T_k \leq S_{z+k-1}\} \cap \{S_{z+n} \in (t, t + dt)\} \cap \{N(t) = n\}.$$

Applying the law of total probability yields

$$\begin{aligned} & \mathbb{P}[\{\tau_z \in (t, t + dt)\} \cap \{N(t) \geq 1\}] \\ &= \sum_{n=1}^{+\infty} \mathbb{P}\left[\bigcap_{k=1}^n \{T_k \leq S_{z+k-1}\} \cap \{S_{z+n} \in (t, t + dt)\} \mid N(t) = n\right] \mathbb{P}[N(t) = n]. \end{aligned} \tag{3}$$

By virtue of the order statistic property, the successive jump times (T_1, \dots, T_n) are distributed as the order statistics $(V_{1:n}, \dots, V_{n:n})$ of a sample of n i.i.d.random variables with CDF $F_t(s)$, $0 \leq s \leq t$. The conditional probability in (3) may be rewritten as

$$\begin{aligned} & \mathbb{P}\left[\bigcap_{k=1}^n \{V_{k:n} \leq S_{z+k-1}\} \cap \{S_{z+n} \in (t, t + dt)\}\right] \\ &= \mathbb{P}\left[\bigcap_{k=1}^n \{U_{k:n} \leq F_t(S_{z+k-1})\} \cap \{S_{z+n} \in (t, t + dt)\}\right] \\ &= \mathbb{P}\left[\bigcap_{k=1}^n \{U_{k:n} \leq F_t(S_{z+k-1})\} \mid S_{z+n} \in (t, t + dt)\right] \mathbb{P}[S_{z+n} \in (t, t + dt)] \\ &= \mathbb{E}\{(-1)^n G_n[0 \mid F_t(S_z), \dots, F_t(S_{z+n-1})] \mid S_{z+n} \in (t, t + dt)\} \mathbb{P}[S_{z+n} \in (t, t + dt)], \end{aligned} \tag{4}$$

where $U_{1:n}, \dots, U_{n:n}$ denote the order statistics of a sample of n i.i.d.uniform random variables on $[0, 1]$, and $G_n(\cdot \mid \cdot)$ correspond to the sequence of A-G polynomials as defined in the Appendix A. Inserting (4) into (3) and letting dt be small enough yields

$$\begin{aligned} f_{\tau_z | N(t) \geq 1}(t) &= \sum_{n=1}^{+\infty} \mathbb{E}\{(-1)^n G_n[0 \mid F_t(S_z), \dots, F_t(S_{z+n-1})] \mid S_{z+n} = t\} \\ &\quad \times f_{S_{z+n}}(t) \mathbb{P}[N(t) = n]. \end{aligned}$$

The final step consists in noting that $G_0(x) = 1$ for every $x \in \mathbb{R}$, and writing

$$f_{\tau_z}(t) = \sum_{n=0}^{+\infty} \mathbb{E}\{(-1)^n G_n[0 \mid F_t(S_z), \dots, F_t(S_{z+n-1})] \mid S_{z+n} = t\} f_{\Delta S}^{*(z+n)}(t) \mathbb{P}[N(t) = n], \tag{5}$$

which is equivalent to (1). □

The stopping time τ_z may be interpreted as the first meeting time of the OSPP $\{N(t), t \geq 0\}$ and the lower randomized boundary defined by $\{M(t) - z, t \geq 0\}$. This remark explains why Theorem 1 is reminiscent of the results given in [17, Proposition 3.1] and [18, Theorem 3.1], where the first-meeting problem of an OSPP with a lower deterministic boundary was handled. The numerical evaluations of (1), to compute, for instance, the probability that the double-spending attack succeeds within a fixed time period, looks challenging. A method based on the truncation of the infinite series in (5) coupled with a numerical integration routine, in the same vein as proposed in [4], can be put together. The next result shows how (1) may be simplified when $\{N(t), t \geq 0\}$ is a mixed Poisson process.

Corollary 1. *If $\{N(t), t \geq 0\}$ is a mixed Poisson process then the PDF of τ_z is given by*

$$f_{\tau_z}(t) = \mathbb{E} \left[\frac{z}{z + N(t)} f_{\Delta^S}^{*[N(t)+z]}(t) \right] \quad \text{for } t \geq 0. \tag{6}$$

Proof. As $\{N(t), t \geq 0\}$ is a mixed Poisson process then we can apply Theorem 1 replacing $F_t(s)$ by s/t for $s \leq t$. The function $h_n(t, z)$ defined in (2) becomes

$$h_n(t, z) = \mathbb{E} \left\{ \frac{1}{t^n} G_n(0 \mid S_z, \dots, S_{n+z-1}) \mid S_{n+z} = t \right\},$$

after applying identity (45). Conditioning upon the values of S_z , and applying successively the identities (45) and (47) leads to

$$\begin{aligned} h_n(t, z) &= \frac{1}{t^n} \mathbb{E} \{ \mathbb{E} [G_n(0 \mid S_z, \dots, S_{n+z-1}) \mid S_z, S_{n+z}] \mid S_{n+z} = t \} \\ &= \frac{1}{t^n} \mathbb{E} \left\{ \mathbb{E} \left[G_n \left(-S_z \mid 0, \Delta_{z+1}^S, \dots, \sum_{k=1}^{n-1} \Delta_{z+k}^S \right) \mid \sum_{k=1}^n \Delta_{z+k}^S = S_{n+z} - S_z \right] \mid S_{n+z} = t \right\} \\ &= \frac{1}{t^n} \mathbb{E} \{ (-S_z) [-S_z - (S_{n+z} - S_z)]^{n-1} \mid S_{n+z} = t \} \\ &= \frac{(-1)^n}{t^n} \mathbb{E} [S_z (S_{z+n})^{n-1} \mid S_{n+z} = t] \\ &= (-1)^n \frac{z}{n+z}. \end{aligned} \tag{7}$$

Substituting (7) into (1) yields (6). □

The formula given in Corollary 1 is reminiscent of the first-hitting time theorem and also the so-called Kendall identity (see, for instance, [3]), which gives the PDF of the first-meeting time of a spectrally negative Lévy process with a lower linear boundary. The following example leads to an expression for the PDF of τ_z that allows the evaluation of the probability of a successful double-spending attack $\mathbb{P}[\tau_z < \infty]$.

Example 1. Assume that the length of the chains $\{z + N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are governed by two homogeneous Poisson processes of intensity λ and μ , respectively. The interarrival times $\{\Delta_k^S, k \geq 1\}$ are i.i.d. exponential random variables with parameter μ and associated PDF

$$f_{\Delta^S}(x) = \mu e^{-\mu x} \quad \text{for } x \geq 0.$$

Applying Corollary 1 yields, after a couple of rearrangements, the following expression for the PDF of τ_z ,

$$f_{\tau_z}(t) = \sum_{n=0}^{+\infty} \binom{z}{z+n} \binom{2n+z-1}{n} \left(\frac{\mu}{\mu+\lambda}\right)^{n+z} \left(\frac{\lambda}{\mu+\lambda}\right)^n \frac{(\lambda+\mu)^{2n+z} t^{2n+z-1} e^{-t(\mu+\lambda)}}{\Gamma(2n+z)} \quad (8)$$

for $t \geq 0$. Assuming that $\lambda > \mu$ and integrating (8) with respect to t yields the probability of successful double-spending attack with

$$\begin{aligned} \mathbb{P}[\tau_z < \infty] &= \sum_{n=0}^{+\infty} \binom{z}{z+n} \binom{2n+z-1}{n} \left(\frac{\mu}{\mu+\lambda}\right)^{n+z} \left(\frac{\lambda}{\mu+\lambda}\right)^n \\ &= \left(\frac{\lambda+\mu}{\lambda}\right)^z \sum_{n=0}^{+\infty} \binom{2n+z-1}{n} \left(\frac{1}{z+n}\right) \left[\frac{\lambda\mu}{(\mu+\lambda)^2}\right]^{n+z} \\ &= \left(\frac{\lambda+\mu}{\lambda}\right)^z \sum_{n=0}^{+\infty} \binom{2n+z-1}{n} \int_0^{\lambda\mu/(\mu+\lambda)^2} t^{n+z-1} dt \\ &= \left(\frac{\lambda+\mu}{\lambda}\right)^z \sum_{n=0}^{+\infty} \binom{2n+z-1}{n} \int_0^{\lambda\mu/(\mu+\lambda)^2} t^{n+z-1} dt \\ &= \left(\frac{\lambda+\mu}{\lambda}\right)^z \int_0^{\lambda\mu/(\mu+\lambda)^2} t^{z-1} \sum_{n=0}^{+\infty} \binom{2n+z-1}{n} t^n dt \\ &= \left(\frac{\lambda+\mu}{\lambda}\right)^z \int_0^{\lambda\mu/(\mu+\lambda)^2} t^{z-1} \frac{C(t)^{z-1}}{\sqrt{1-4t}} dt, \end{aligned} \quad (9)$$

where

$$C(t) = \frac{1 - \sqrt{1-4t}}{2t}, \quad (10)$$

denotes the generating function of Catalan’s numbers; see, for instance, [1, Chapter 3]. Note that the last equality follows from an exercise in the textbook of Aigner [1, Exercise 3.25]. Inserting (10) into (9) yields, after straightforward integration,

$$\mathbb{P}[\tau_z < +\infty] = \left(\frac{\mu}{\lambda}\right)^z.$$

The result given above is consistent with Corollary 3; see Section 4.

3. The SF of the double-spending time

In this section, the length of the honest and the malicious chains are governed by two independent OSPPs. The order statistic property satisfied by $\{M(t), t \geq 0\}$ implies that

$$[S_1, \dots, S_m \mid M(t) = m] \stackrel{D}{=} (V_{1:m}^*, \dots, V_{m:m}^*),$$

where $V_{1:m}^*, \dots, V_{m:m}^*$ denote the order statistics of a sample V_1^*, \dots, V_m^* of m i.i.d. random variables having a CDF $F_t^*(s)$ for $0 \leq s \leq t$. Note that, regarding $\{N(t), t \geq 0\}$, the notation of Section 2 is preserved. The following result gives a formula in terms of Appell polynomials for the SF of $\tau_z = \inf\{t \geq 0, M(t) = N(t) + z\}$.

Theorem 2. If $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are two OSPPs, then the SFof τ_z is given by

$$\mathbb{P}[\tau_z > t] = \mathbb{E}(A_{M(t)}\{1 \mid 0, \dots, 0, F_t^*[V_{1:N(t)}], \dots, F_t^*[V_{M(t)+1-z:N(t)}]\} \mathbf{1}_{M(t) \leq N(t)+z-1}) \quad (11)$$

for $t \geq 0$, where $A_n(1 \mid \cdot)$ is an Appell polynomial such as defined in Appendix A.

Proof. If $M(t) \geq N(t) + z$, at time $t \geq 0$, then the double-spending attack already occurred. Consider the event $\{\tau_z > t\}$, conditioning upon the possible values of the counting processes leads to

$$\{\tau_z > t\} = \bigcup_{n=0}^{+\infty} \bigcup_{m=0}^{n+z-1} \{\tau_z > t\} \cap \{N(t) = n\} \cap \{M(t) = m\}.$$

The double-spending attack did not happen before time $t \geq 0$ if $M(t)$ is smaller or equal to $z - 1$, irrespective of the value of $N(t)$. If $M(t)$ falls between z and $N(t) + z$ then $N(t)$ must have jumped, at least once, and an investigation over the jump times of both point processes must be conducted. The event $\{\tau_z > t\}$ is further rewritten as

$$\{\tau_z > t\} = \bigcup_{m=0}^{z-1} \{M(t) = m\} \cup \bigcup_{n=1}^{+\infty} \bigcup_{m=z}^{n+z-1} \bigcap_{k=z}^m \{S_k > T_{k+1-z}\} \cap \{N(t) = n\} \cap \{M(t) = m\}.$$

The law of total probability yields

$$\begin{aligned} \mathbb{P}[\tau_z > t] &= \mathbb{P}[M(t) \leq z - 1] + \sum_{n=1}^{+\infty} \sum_{m=z}^{n+z-1} \mathbb{P}\left[\bigcap_{k=z}^m \{S_k > T_{k+1-z}\} \mid N(t) = n, M(t) = m\right] \\ &\quad \times \mathbb{P}[N(t) = n, M(t) = m]. \end{aligned} \quad (12)$$

Now, by the order statistic property, it holds that

$$[(T_1, \dots, T_{m+1-z}) \mid N(t) = n] \stackrel{D}{=} (V_{1:n}, \dots, V_{m+1-z:n})$$

and

$$[(S_z, \dots, S_m) \mid M(t) = m] \stackrel{D}{=} (V_{z:m}^*, \dots, V_{m:m}^*).$$

Therefore, the conditional probability in (12) may be rewritten as

$$\begin{aligned} \mathbb{P}\left[\bigcap_{k=z}^m \{V_{k:m}^* > V_{k+1-z:n}\}\right] &= \mathbb{P}\left[\bigcap_{k=z}^m \{U_{k:m} > F_t^*(V_{k+1-z:n})\}\right] \\ &= A_m[1 \mid 0, \dots, 0, F_t^*(V_{1:n}), \dots, F_t^*(V_{m+1-z:n})], \end{aligned} \quad (13)$$

where $U_{1:m}, \dots, U_{m:m}$ are the order statistics of a sample of m i.i.d. uniform random variables on $(0, 1)$ and $A_m(\cdot \mid \cdot)$ denote the Appell polynomials defined in Appendix A. Inserting (13) into (12) yields

$$\begin{aligned} \mathbb{P}[\tau_z > t] &= \mathbb{P}[M(t) \leq z] + \sum_{n=1}^{+\infty} \sum_{m=z}^{n+z-1} A_m[1 \mid 0, \dots, 0, F_t^*(V_{1:n}), \dots, F_t^*(V_{m+1-z:n})] \\ &\quad \times \mathbb{P}[N(t) = n, M(t) = m], \end{aligned} \quad (14)$$

which is the same as (11) after noting that $A_m(1 \mid 0, \dots, 0) = 1$ for every $m \in \mathbb{N}$. □

In Subsection 5.2 a Monte Carlo evaluation of the expectation of (11) is performed. This type of estimator has been studied in [17, Section 6] and named Appell polynomial Monte Carlo (APMC). The procedure entails a variance reduction in comparison to a crude Monte Carlo evaluation. The following result shows how (11) may be simplified by setting $z = 1$, when the OSPPs are similar in a sense detailed below.

Corollary 2. *Assume that $z = 1$. If $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are two OSPPs such that $F_t(s) = F_t^*(s)$ for every $s \leq t$, then the SFof τ_z is given by*

$$\mathbb{P}[\tau_z > t] = \mathbb{E}\left(\frac{N(t) - M(t) + 1}{N(t) + 1} \mathbf{1}_{M(t) \leq N(t)}\right) \quad \text{for } t \geq 0. \tag{15}$$

Proof. Let $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ be two OSPPs such that $F_t(s) = F_t^*(s)$ for every $s \leq t$. Applying Theorem 2, with $z = 1$, yields

$$\mathbb{P}[\tau_z > t] = \mathbb{E}(A_{M(t)}\{1 \mid U_{1:N(t)}, \dots, U_{M(t):N(t)}\} \mathbf{1}_{M(t) \leq N(t)}) \quad \text{for } t \geq 0.$$

Recall the probabilistic interpretation of the Appell polynomial in Proposition 1 with

$$\mathbb{P}[\tau_z > t] = \mathbb{E}(\mathbb{P}[U_{1:M(t)}^* > U_{1:N(t)}, \dots, U_{M(t):M(t)}^* > U_{M(t):N(t)}] \mathbf{1}_{M(t) \leq N(t)})$$

for $t \geq 0$. Applying Bertrand’s ballot theorem, allowing for ties, yields (15). □

The case treated in the numerical illustrations considers that the length of the chains are governed by two nonhomogeneous Poisson processes, which is consistent with the empirical study conducted in [8], as explained in the following example.

Example 2. Bowden *et al.* [8] recommended modeling the arrival of blocks as a nonhomogeneous Poisson process with an intensity function designed to capture the evolution of the global hashrate on one hand and the difficulty adjustment of the cryptopuzzles on the other hand. The block number n is associated to a hash $f(n)$, which is a number drawn randomly from the lattice $\{0, 1, \dots, 2^{256} - 1\}$. Mining a block consists in computing the hash of the block until it is lower than a target L . The number of trials required is a geometric random variable $\text{Geom}(L \times 2^{-256})$ with associated PMF $(1 - L \times 2^{-256})^{k-1} L \times 2^{-256}$ for $k \geq 1$. The difficulty is adjusted by tuning the target L . Denote by $\{T_k, k \geq 1\}$ the sequence of arrival times of the blocks. The difficulty is adjusted every 2016 blocks to maintain an average of 1 block mined every 10 minutes. Mining 2016 takes about 2 weeks. This leads to the definition of a piecewise constant target function $L(t)$ as

$$L(t) = \begin{cases} L_0 & \text{for } t \in (0, T_{2016}), \\ L_k & \text{for } t \in (T_{2016k}, T_{2016(k+1)}) \text{ and } k > 0, \end{cases}$$

where the sequence of real numbers $\{L_k, k \geq 0\}$ is defined recursively as

$$L_k = \begin{cases} 2^{224} & \text{for } k = 0, \\ L_{k-1} \frac{T_{2016k} - T_{2016(k-1)}}{1209600} & \text{for } k > 0. \end{cases}$$

Note that the time unit is the second and 1 209 600 seconds correspond to 2 weeks. The number of trials relates to the mining time through the (global) hashrate function $H(t)$. The hashrate function corresponds to the number of hashes computed per second by the entire network

of miners. Hence, the instantaneous average mining time is given by $2^{256}/H(t)L(t)$ and the intensity function of the underlying nonhomogeneous Poisson process is given by

$$\lambda(t) = \frac{H(t)L(t)}{2^{256}} \quad \text{for } t \geq 0. \tag{16}$$

There are two main drivers of the hashrate. First, the improvement of the mining machines which enhances the computing power of the miners. Second, the number of miners in the network. The miners enter and exit the network according to how profitable mining BTCs is at the moment. This last factor depends on the price of the electricity and the value of the BTCs at a given point in time. Information regarding the target $L(t)$ can be collected from the header of the block. The hashrate $H(t)$ is retrieved from the knowledge of the difficulty and the timestamp data. Bowden *et al.* [8] proposed an exponential function of the form $H(t) = e^{at+b}$ arguing that the log hashrate is piecewisely linear over time. The values of a and b follow from the linear interpolation within successive time periods. Once the hashrate function has been determined, the length of the blockchain $\{N(t), t \geq 0\}$ is a nonhomogeneous Poisson process with intensity function $\lambda(t)$ defined in (16). Assuming that $N(t) = n$, the arrival times T_1, \dots, T_n are distributed as the order statistics of n i.i.d.random variables with associated CDF

$$F_t(s) = \frac{\Lambda(s)}{\Lambda(t)} \quad \text{for } s \leq t,$$

where $\Lambda(t) = \int_0^t \lambda(s) ds$. In the event of a double-spending attack, the difficulty of the puzzle is the same for the honest miners and the colluding miners. The difference between the two pools lies in their computing power and, thus, their hashrate function. We may assume that both the honest and dishonest miners contribute to the global hashrate of the network in an additive way. More specifically, let $H_1(t) = p \times H(t)$ be the hashrate of the honest miners and $H_2(t) = (1 - p) \times H(t)$, where $p \in (0, 1)$ represents the repartition of the computing resources among the miners. Theorem 2 is applicable and (11) simplifies to

$$\mathbb{P}[\tau_z > t] = \mathbb{E}(A_{M(t)}\{1 \mid 0, \dots, 0, U_{1:N(t)}, \dots, U_{M(t)+1-z:N(t)}\} \mathbf{1}_{M(t) \leq N(t)+z-1}), \tag{17}$$

because $F_t^*(s) = F_t(s)$ for every $s \leq t$. An evaluation via Monte Carlo simulation is possible by generating values for $M(t)$, $N(t)$, and $U_{1:N(t)}, \dots, U_{M(t)+1-z:N(t)}$. Appell polynomials do not usually admit a closed-form expression but can be computed recursively via the relations provided in Appendix A; see Proposition 2.

Note that the evaluation of (11) may be achieved through the truncation of the infinite series in (14) followed by numerical integration, in the same vein as performed in [11]. Another solution would be to resort to a fully recursive evaluation as in [22].

4. The probability of a successful double-spending attack

In this section we study the probability of a successful double-spending attack, $\mathbb{P}[\tau_z < +\infty]$, when the length of the chains $\{z + N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are modeled by independent renewal processes generated by their respective sequence of i.i.d.interarrival times denoted by $\{\Delta_k^T, k \geq 1\}$ and $\{\Delta_k^S, k \geq 1\}$. Assume that

(A1) $\mathbb{E}(\Delta^S) > \mathbb{E}(\Delta^T)$;

(A2) the equation

$$\log \mathbb{E}[e^{\theta(\Delta^T - \Delta^S)}] = 0 \tag{18}$$

has a unique nonnegative solution denoted by γ , referred to as the adjustment coefficient.

The stopping time $\tau_z = \inf\{t \geq 0; M(t) = z + N(t)\}$ coincides with the ruin time $\tau_z = \inf\{t \geq 0; R(t) = 0\}$ associated to the risk process

$$R(t) = z + N(t) - M(t), \quad t \geq 0. \tag{19}$$

Define the claim surplus process as

$$S(t) = M(t) - N(t), \quad t \geq 0. \tag{20}$$

In risk theory, processes such as (19) model the evolution of the net worth of an insurance company over time. Here, the insurance company holds an initial capital of amount z , its premium income is governed by $\{N(t), t \geq 0\}$ while $\{M(t), t \geq 0\}$ corresponds to its liability at time $t \geq 0$. Note that risk theory terminology is being used only to improve the presentation, I am not claiming that one should model the evolution of the financial reserves of any nonlife insurance company via (19). When studying the distribution of the ruin time is problematic, a simple trick consists in passing to a dual risk model. This approach is rather standard (see the references given in the introduction). For the sake of clarity, the idea is recalled hereafter and illustrated by Figure 2. In Figure 2(a) we display the ruin problem in model (19). The initial ruin problem is converted into another equivalent ruin problem. Increment the value of $\{M(t), t \geq 0\}$ by one unit and consider the risk model

$$\tilde{R}(t) = z + N(t) - [M(t) + 1], \quad t \geq 0.$$

Further define the ruin time

$$\tilde{\tau}_z = \inf\{t \geq 0; z + N(t) < [M(t) + 1]\},$$

which corresponds to the first-crossing time of $\{M(t) + 1, t \geq 0\}$ through the upper boundary $\{N(t) + z, t \geq 0\}$; see Figure 2(b). It holds that

$$\tau_z \stackrel{\text{a.s.}}{=} \tilde{\tau}_z,$$

where ‘ $\stackrel{\text{a.s.}}{=}$ ’ stands for equality almost sure as it is true for every trajectory. Then rotate Figure 2(b) by 90° anticlockwise to obtain Figure 2(c). Shifting the origin from $(0, 0)$ to $(z - 1, 0)$ finally leads to Figure 2(d). The ruin problem displayed in Figure 2(d) concerns a discrete-time risk model denoted by $\{R^*(n), n \geq 1\}$ and defined as

$$R^*(n) = S_{z-1} + \sum_{k=1}^n (\Delta_{k+z-1}^S - \Delta_k^T) \quad \text{for } n \in \mathbb{N}. \tag{21}$$

The initial capital is S_{z-1} , the sequence $\{\Delta_{k+z-1}^S, k \geq 1\}$ models the premium collected at each time period, and the sequence $\{\Delta_k^T, k \geq 1\}$ represents the total claim amounts incurred during each time period. The conventions $S_0 = 0$ and $T_0 = 0$ are adopted. The claim surplus process $\{S^*(n), n \geq 1\}$ is given by

$$S^*(n) = \sum_{k=1}^n (\Delta_k^T - \Delta_{k+z-1}^S). \tag{22}$$

The ruin time is defined as $\sigma_{S_{z-1}} = \inf\{n \in \mathbb{N}; R^*(n) \leq 0\}$ and relates to $\tau_z = \inf\{t \geq 0; R(t) = 0\}$ as

$$\tau_z \stackrel{\text{a.s.}}{=} \tilde{\tau}_z \stackrel{\text{a.s.}}{=} S_{\sigma_{S_{z-1}} + z - 1},$$

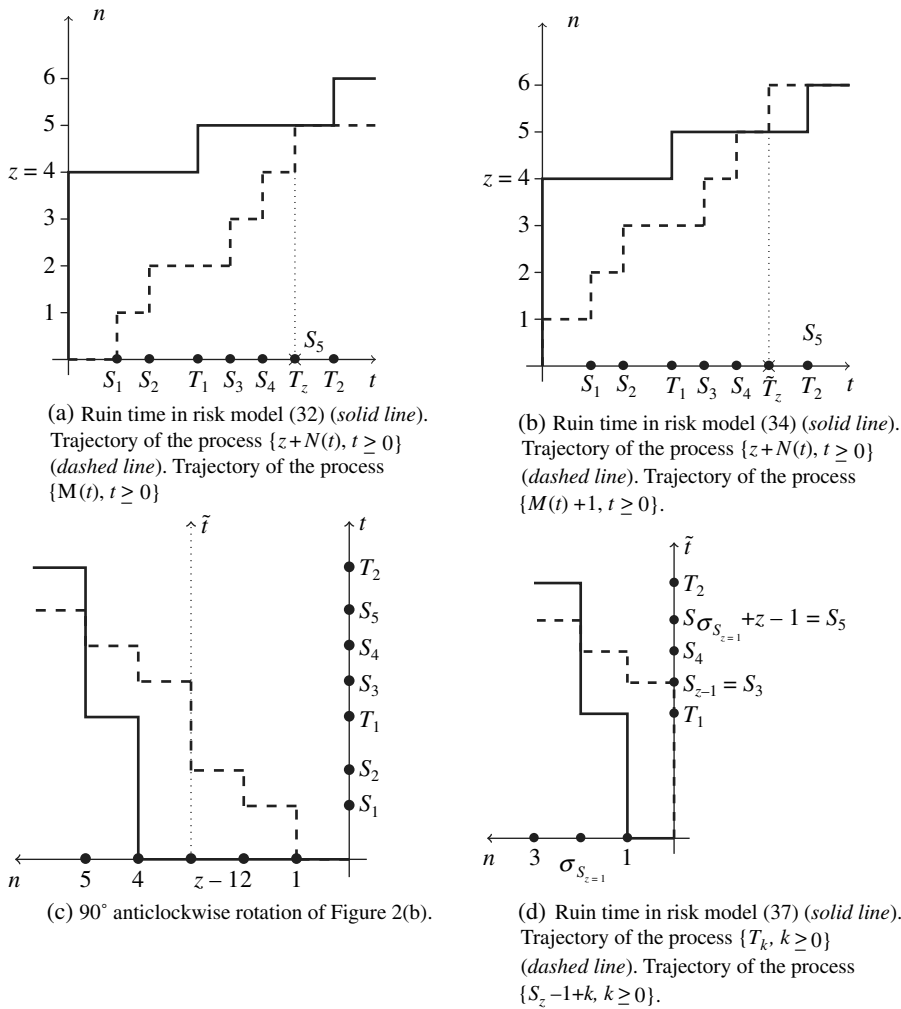


FIGURE 2: Boundary crossing problems in the various risk models.

which implies that

$$\mathbb{P}[\tau_z < \infty] = \mathbb{P}[\sigma_{S_{z-1}} < \infty]. \tag{23}$$

Again the one-to-one correspondence between the trajectories leading to ruin in the multiple risk models entail the equality almost surely. The following result provides, *inter alia*, an upper bound for the probability of a successful double-spending attack.

Theorem 3. *If $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are two independent renewal processes such that (A1) and (A2) hold then*

$$\mathbb{P}[\tau_z < \infty] = \frac{[\mathbb{E}(e^{-\gamma \Delta^S})]^{z-1}}{\mathbb{E}[e^{\gamma \xi(S_{z-1})} \mid \tau_z < \infty]}, \tag{24}$$

where $\xi(S_{z-1}) = S(\sigma_{S_{z-1}}) - S_{z-1}$ denotes the overshoot, immediately after ruin, in model (21). The following Cramér–Lundberg upper bound holds:

$$\mathbb{P}[\tau_z < \infty] \leq [\mathbb{E}(e^{-\gamma \Delta^S})]^{z-1}. \tag{25}$$

Proof. The claim surplus process $\{S^*(n), n \geq 1\}$ in (22) is a random walk. Assumption (A2) implies that the process $\{e^{\gamma S^*(n)}, n \geq 1\}$ is a martingale as a consequence of [2, Theorem 1.1]. Note also that $S^*(n) \xrightarrow{\text{a.s.}} -\infty$, where ‘a.s.’ stands for convergence almost surely, follows from assumption (A1) and the law of large numbers. Let the initial reserves $S_{z-1} = s \geq 0$ be fixed in (21). The application of [2, Proposition 3.1] allows us to rewrite the ultimate ruin probability as

$$\mathbb{P}[\sigma_s < \infty] = \frac{e^{-\gamma s}}{\mathbb{E}[e^{\gamma \xi(s)} \mid \sigma_s < \infty]},$$

where $\xi(u) = S^*(\sigma_s) - s$ denotes the overshoot given that ruin occurred in model (21). Thanks to the connection (23), by conditioning on the values of S_{z-1} , it holds that

$$\mathbb{P}[\tau_z < \infty] = \frac{\mathbb{E}[e^{-\gamma S_{z-1}}]}{\mathbb{E}[e^{\gamma \xi(S_{z-1})} \mid \tau_z < \infty]} = \frac{\mathbb{E}[e^{-\gamma \Delta^S}]^{z-1}}{\mathbb{E}[e^{\gamma \xi(S_{z-1})} \mid \tau_z < \infty]}.$$

The upper bound (25) follows from noting that $\mathbb{E}[e^{\gamma \xi(S_{z-1})} \mid \sigma_s < \infty] > 1$. □

The next result specifies the expression for the probability $\mathbb{P}[\tau_z < \infty]$ in the case where $\{N(t), t \geq 0\}$ is a Poisson process.

Corollary 3. *If $\{N(t), t \geq 0\}$ is a homogeneous Poisson process of intensity λ and $\{M(t), t \geq 0\}$ is a renewal process, independent from $\{N(t), t \geq 0\}$, such that (A1) and (A2) holds then*

$$\mathbb{P}[\tau_z < \infty] = \frac{\lambda - \gamma}{\lambda} \mathbb{E}[e^{-\gamma \Delta^S}]^{z-1}. \tag{26}$$

If $\{M(t), t \geq 0\}$ is also a homogeneous Poisson process of intensity $\mu < \lambda$ then

$$\mathbb{P}[\tau_z < \infty] = \left(\frac{\mu}{\lambda}\right)^z. \tag{27}$$

Proof. As $\{N(t), t \geq 0\}$ is a homogenous Poisson process, it is a renewal process and (24) holds. The sequence of interarrival times $\{\Delta_k^T, k \geq 1\}$ is formed by i.i.d.exponential random variables with parameter λ , which implies that the overshoot $\xi(S_{z-1})$ is also exponentially distributed with parameter λ by virtue of the lack of memory of the exponential distribution. It follows that

$$\mathbb{E}[e^{\gamma \xi(S_{z-1})} \mid \tau_z < \infty] = \frac{\lambda}{\lambda - \gamma}.$$

Substituting in (24) yields (26). Now, assume that $\{M(t), t \geq 0\}$ is also a Poisson process with intensity $\mu < \lambda$. The sequence of interarrival times $\{\Delta_k^S, k \geq 1\}$ is made of i.i.d.exponential random variables with parameter μ . Equation (18) is equivalent to

$$\log \left(\frac{\lambda \mu}{(\lambda - \theta)(\mu + \theta)} \right) = 0$$

and admits $\gamma = \lambda - \mu$ as only nonnegative solution. Substituting $\gamma = \lambda - \mu$ into (26) yields (27). □

This result allows confirmation of Corollary 1; see Example 1. Note that the probability of a successful double-spending attack when the length of the chains are two independent Poisson processes may be retrieved without using the duality argument as shown in the following remark.

Remark 1. Assume that $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are two independent homogeneous Poisson processes with respective intensity λ and μ such that $\lambda > \mu$. The claim surplus process, already defined in (20) as

$$S(t) = M(t) - N(t) \quad \text{for } t \geq 0,$$

is the difference between two independent Poisson processes and, thus, a Lévy process. Theorem 1.2 of [2] then implies that the process defined by

$$e^{\theta S(t) - t\kappa(\theta)} \quad \text{for } t \geq 0,$$

where $\kappa(\theta) = \log \mathbb{E}[\theta S(1)]$, is a martingale. The equation $\kappa(\theta) = 0$ is equivalent to

$$\mu e^\theta + \lambda e^{-\theta} - (\lambda + \mu) = 0$$

and admits a unique nonnegative solution $\gamma = \log(\lambda/\mu)$. Consequently, the process $\{e^{\gamma S(t)}, t \geq 0\}$ is a martingale. Moreover, the condition $\lambda > \mu$ entails $S(t) \xrightarrow{\text{a.s.}} -\infty$. Applying [2, Proposition 3.1] yields

$$\mathbb{P}[\tau_z < \infty] = \frac{e^{-\gamma z}}{\mathbb{E}[e^{\gamma \xi(z)} \mid \tau_z < \infty]}, \tag{28}$$

where $\xi(z) = S(\tau_z) - z$ denotes the overshoot after ruin occurred in model (19). In the case considered here there is no overshoot as $S(\tau_z) = z$. Substituting $\gamma = \log(\lambda/\mu)$ into (28) yields (26).

5. Numerical illustrations

The numerical results presented here are based on the data collected by Bowden [7] and the analysis conducted in [8]. In Subsection 5.1 data is used to fit the interblock time distribution within the public chain. The lack of data regarding the growth of the malicious chain is circumvented by assuming that the interblock times in the malicious chain are defined as a transformation of the interblock times in the public chain. This allows us to illustrate the results given in Sections 2 and 4. Subsection 5.2 focuses on the case where the lengths of the chains are governed by nonhomogeneous Poisson processes which seems to be the most suitable model according to Bowden *et al.* [8]. It allows us to illustrate the results derived in Section 3.

5.1. Length of the chains as renewal process

In this subsection the length of the chains $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are assumed to be governed by renewal processes. The first task consists in studying the fit of the interblock time distribution to the data provided in [7]. In Figure 3 we display the interblock times chronologically. The distribution of the interblock times of the first few blocks presents a few spikes (to the magnitude of the day) before reaching stationarity around the 200 000th block; see Figure 3(a). If we limit our analysis to the latest blocks, starting, for instance, from the 400 000th, then the data admits fewer outliers, with a maximum of 2 hours; see Figure 3(b). The inference of the block arrival is therefore based on the interarrival times starting from the 400 000th block onward which still represents 96 628 data points. The empirical mean is

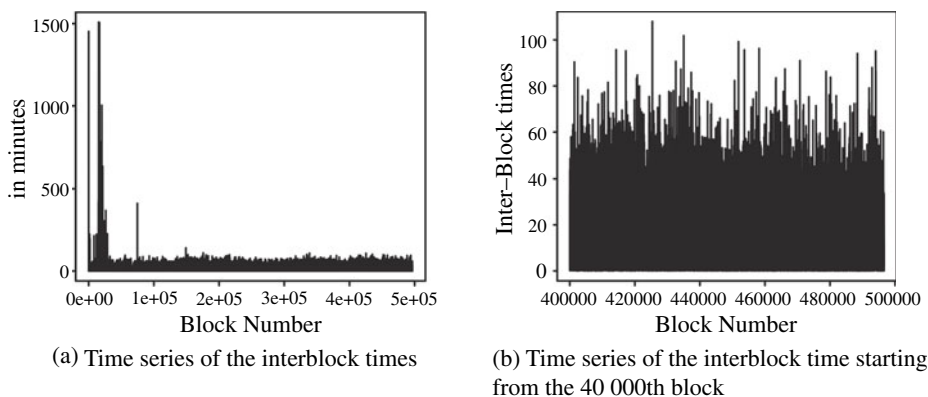


FIGURE 3: Chronological series of the interblock times.

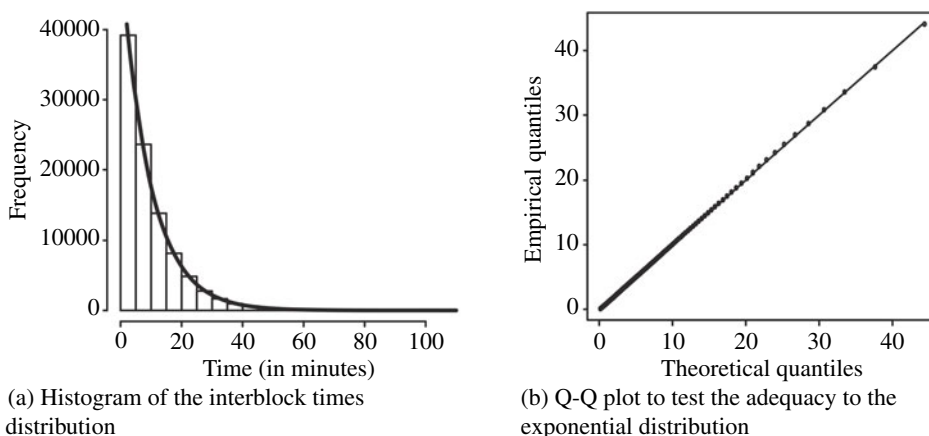


FIGURE 4: Distribution of the interblock times.

equal to 9.57 minutes while the standard deviation is significantly lower than the overall one with 9.56 minutes. In Figure 4 we show the histogram of the interblock times. The PDF of the exponential distribution $\text{Exp}(\hat{\lambda})$, where $\hat{\lambda} = 1/9.57$ corresponds to the method of the moment estimator, matches reasonably well the histogram. In Figure 4(b) the empirical quantiles are plotted against the quantiles of the exponential distribution $\text{Exp}(\hat{\lambda})$. The points overlap the diagonal $y = x$, which indicates a superb fit. This analysis leads us to model the number of blocks in the honest chain $\{N(t), t \geq 0\}$ by a homogeneous Poisson process of intensity $\hat{\lambda}$.

Regarding the block arrival in the malicious chain, no data is available. The only *a priori* information is that the interblock time should be larger to account for the unbalanced repartition of the computing power in favour of the honest miners. The growth of $\{M(t), t \geq 0\}$ should be slower than the growth of $\{N(t), t \geq 0\}$. In the sequel, two definitions of the interarrival times $\{\Delta_k^S, k \geq 1\}$ that generate $\{M(t), t \geq 0\}$ are compared in terms of the risk of a double-spending attack.

1. Define

$$\Delta^S \stackrel{D}{=} \frac{\Delta^T}{r},$$

where $r > 1$. The interarrival times $\{\Delta_k^S, k \geq 1\}$ are i.i.d.exponential random variables and the process $\{M(t), t \geq 0\}$ is a homogeneous Poisson process with intensity $\widehat{\lambda}/r$. Corollary 3 applies and the probability of a successful double-spending attack is given by

$$\mathbb{P}[\tau_z < \infty] = \left(\frac{1}{r}\right)^z. \tag{29}$$

The PDF of the double-spending time f_{τ_z} follows from applying Corollary 1 and is given in (8) after substituting $\lambda = \widehat{\lambda}$ and $\mu = \widehat{\lambda}/r$.

2. Define

$$\Delta^S \stackrel{D}{=} \Delta_1^T + \dots + \Delta_r^T,$$

where $r > 1$ is integer valued. The interarrival times $\{\Delta_k^S, k \geq 1\}$ are i.i.d.gamma random variables $\text{Gam}(r, \lambda)$ with associated PDF

$$f_{\Delta^S}(t) = \frac{e^{-\lambda t} \lambda^r t^{r-1}}{\Gamma(r)} \quad \text{for } t \geq 0.$$

The process $\{M(t), t \geq 0\}$ is a renewal process and Corollary 3 applies. The probability of a successful double-spending attack is given by

$$\mathbb{P}[\tau_z < \infty] = \frac{\lambda - \gamma}{\lambda} \left[\frac{\lambda}{\lambda + \gamma} \right]^{r(z-1)}, \tag{30}$$

where γ is the only nonnegative solution to

$$\log \left[\frac{\lambda^r}{(\lambda - \theta)(\lambda + \theta)^{r-1}} \right] = 0. \tag{31}$$

Note that the root in (31) is derived numerically using the `uniroot` built-in function in R. The PDF f_{τ_z} of the double-spending time follows from the application of Corollary 1 and reduces, after a couple of rearrangements, to

$$f_{\tau_z}(t) = \sum_{n=0}^{+\infty} \binom{z}{z+n} \frac{\Gamma[r(n+z)+n]}{\Gamma(n+1)\Gamma[r(n+z)]2^{r(n+z)+n}} \frac{(2\lambda)^{r(n+z)+n} t^{r(n+z)+n-1} e^{-2\lambda t}}{\Gamma[r(n+z)+n]} \tag{32}$$

for $t \geq 0$.

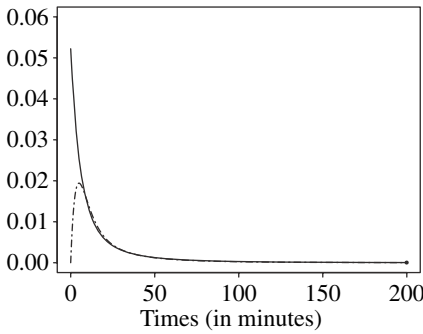
In Table 1 we report the values of (29) and (30) for $r = 2, 3, 4, 5$ and $z = 1, 2, 3, 4, 5$. Although definitions 1 and 2 both mean that the building of blocks is r times slower on average in the malicious chain, the probability of a successful double-spending attack is much smaller when $\{M(t), t \geq 0\}$ is a renewal process with gamma-distributed interarrival times. It shows the influence of the shape of the distribution of the interarrival times on the likelihood of a double-spending attack. In Figure 5 we display the PDF and CDF of the double-spending time

$$\tau_1 = \inf\{t \geq 0; N(t) + 1 = M(t)\}$$

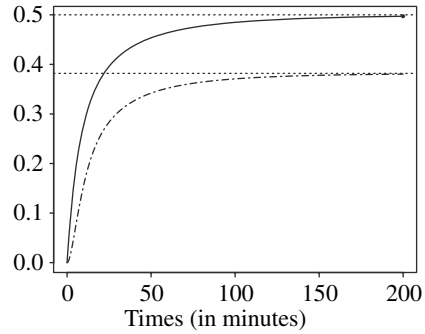
for $r = 2$ along with the reference horizontal lines $y = \mathbb{P}[\tau_1 < \infty]$. Note that the infinite series in (8) and (32) are truncated to the order $K = 50$. In both cases, the merchant who is waiting for

TABLE 1: Probability of a successful double-spending attempt

z	$\Delta^S \sim \text{Exp}(\lambda/r)$				$\Delta^S \sim \text{Gamma}(r, \lambda)$			
	$r=2$	$r=3$	$r=4$	$r=5$	$r=2$	$r=3$	$r=4$	$r=5$
1	0.5000	0.3333	0.2500	0.2000	0.3819	0.1608	0.0724	0.0342
2	0.2500	0.1111	0.0625	0.0400	0.1459	0.0258	0.0052	0.0012
3	0.1250	0.0370	0.0156	0.0080	0.0557	0.0042	0.0004	0.0000
4	0.0625	0.0123	0.0039	0.0016	0.0213	0.0007	0.0000	0.0000
5	0.0312	0.0041	0.0010	0.0003	0.0081	0.0001	0.0000	0.0000



(a) PDF of τ_1



(b) CDF of τ_1 as well $\mathbb{P}(\tau_1 < \infty)$

FIGURE 5: PDF and CDF of the double-spending time τ_1 : $\{M(t), t \geq 0\}$ is a homogeneous Poisson process with intensity $\hat{\lambda}/2$ (solid line), $\{M(t), t \geq 0\}$ is a renewal process with gamma $\text{Gam}(r=2, \hat{\lambda})$ distributed interarrival times (dash-dot line).

two hours is not taking much risk as the CDF reaches the barrier $\mathbb{P}[\tau_1 < \infty]$. Namely, if $\mathbb{P}[\tau_1 < \infty] - \mathbb{P}[\tau_1 < t] = \varepsilon$, for some $t \geq 0$ and $\varepsilon > 0$, then the probability of the double-spending ever being successful, conditioning upon $\{\tau_z > t\}$, may be written as

$$\mathbb{P}[\tau_z = +\infty \mid \tau_z > t] = \frac{\mathbb{P}[\tau_z = +\infty]}{\mathbb{P}[\tau_z = +\infty] + \varepsilon}.$$

This probability becomes close to 1 as we let ε tend to 0. The R code is accessible online at [16] for the sake of reproducibility.

5.2. Length of the chains as a nonhomogeneous Poisson process

In this subsection the length of the chains $\{z + N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are assumed to be governed by two nonhomogeneous Poisson processes. The intensity function of $\{N(t), t \geq 0\}$ is given by

$$\lambda(t) = \frac{pH(t)L(t)}{2^{256}},$$

where $H(t)$ denotes the global hashrate function, $L(t)$ is the target function, and $p \in (0, 1)$ reflects the repartition of the computing power between honest and dishonest miners. We refer the reader to Example 2 for a definition of these quantities. The intensity function of $\{M(t), t \geq 0\}$ is given by

$$\lambda^*(t) = \frac{(1-p)H(t)L^*(t)}{2^{256}}.$$

The global hashrate is assumed to admit a parametric form with $H(t) = e^{at+b}$, where the values of a and b are selected from [8, Table 1]. The difficulty is assumed to be the same for all the miners so that $L^*(t) = L(t)$. More specifically, let us consider a time span during which the difficulty is constant, equal to L say. This is true for time periods that are about two weeks long as it corresponds to the average time to discover 2016 blocks. The intensity function $\lambda(t)$ associated to $\{N(t), t \geq 0\}$ becomes

$$\lambda(t) = \frac{pe^{at+b}L}{2^{256}},$$

and may be integrated as

$$\Lambda(t) = \int_0^t \lambda(s) ds = \frac{pe^bL}{2^{256}a}(e^{at} - 1). \tag{33}$$

Note that we equivalently have

$$\lambda^*(t) = \frac{(1-p)e^{at+b}L}{2^{256}}$$

and

$$\Lambda^*(t) = \int_0^t \lambda^*(s) ds = \frac{(1-p)e^bL}{2^{256}a}(e^{at} - 1). \tag{34}$$

In view of these assumptions, the conditional distribution of the block arrival times given the length of the chain is the same in the honest and the malicious chains. Namely, it holds that $F_t(s) = F_t^*(s)$ and the probability $\mathbb{P}[\tau_z > t]$ of the double-spending attack being unsuccessful before t may be estimated via (17). The practical evaluation is handled via Monte Carlo simulations; note that the numerical value of Appell polynomials of any order may be computed recursively using the relations given in Appendix A (see Proposition 2). Let the time unit be one second, and consider a two week time period, which corresponds to 1 209 600 seconds. Let us set the parameters of the hashrate function to $a = -9.44 \times 10^{-9}$ and $b = 27.1$ according to the first row of Table 1 of [8]. The difficulty is assumed to be constant, equal to

$$L = 2016 \times \frac{2^{256}a}{e^{b(e^{a*1209600} - 1)}},$$

in order to have on average 2016 blocks discovered by the end of the time horizon (i.e.two weeks). In Figure 6 we display the integrated intensity functions (33) and (34) over time. The parametrization entails a linear growth (on average) of the chains which makes our example close to the homogeneous Poisson arrival situation. Equation (11) is difficult to use for risk management purposes without the knowledge of the mass of probability associated to $\tau_z = \infty$. This issue is addressed by assuming that an attacker gives up his double-spending attempt if not completed within three hours (10 800 seconds). It makes little sense in practice for an attacker to carry on an attack for two weeks. We then investigate the probability of a successful double-spending attack attempted every three hours over the course of two weeks. Namely, denote by

$$t_k = k \times 10800 \quad \text{for } 0 \leq k < 112,$$

the sequence of time steps and by

$$p_{z,k} = \mathbb{P}[\tau_{z,k} < 10800] \tag{35}$$

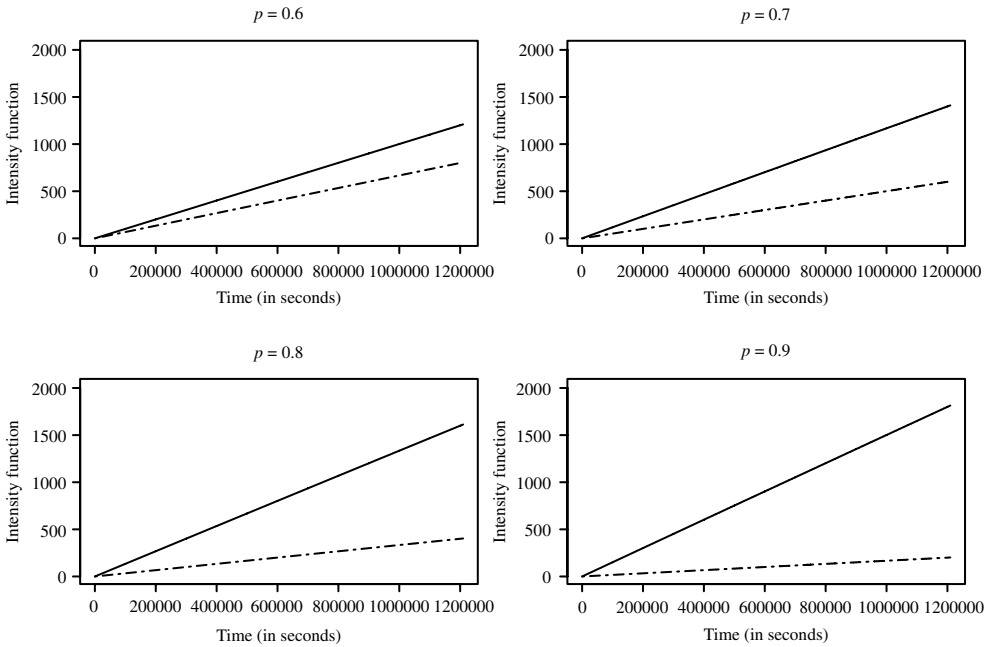


FIGURE 6: Integrated intensity functions over time: $\Delta(t)$ associated to $\{N(t), t \geq 0\}$ (solid line), $\Delta^*(t)$ associated to $\{M(t), t \geq 0\}$ (dash-dot line).

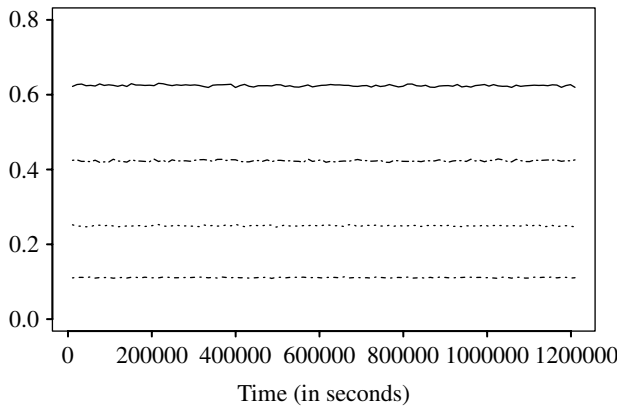


FIGURE 7: Evolution of the probability of performing a successful double-spending attack in the course of two weeks for various values of p : $p = 0.6$ (solid line), $p = 0.7$ (dashed line), $p = 0.8$ (dotted line), $p = 0.9$ (dash-dot line).

the probabilities of interest, where

$$\tau_{z,k} = \inf\{t \in (t_k, t_{k+1}); M(t) = z + N(t) \mid M(t_k) = N(t_k) = 0\}.$$

Let us assume that the honest chain is one block ahead, which in turn allows us to use (15) and alleviate the computational burden associated to the recursive evaluation of Appell polynomials. In Figure 7 we display the value of the probabilities (35) over the two weeks of operations for various repartition $p \in \{0.6, 0.7, 0.8, 0.9\}$ of the hash power. Note that the

evaluation is based on 10 000 trajectories of $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$. The probabilities $p_{z,k}$ are constant over time, which was expected as the arrival of blocks is almost time homogenous due to the parametrization of the global hashrate function. The source code is available online [16] and the reader is invited to experiment the effect of modifying the parameters a and b on the double-spending probabilities.

6. Concluding remarks

In this paper, the model, initially proposed by Nakamoto [27], to comprehend the double-spending issue is refined. Assuming that the lengths of the competing blockchains are governed by counting processes leads to interesting boundary crossing problems. This refinement offers more flexibility to reflect accurately the block discovery frequency as well as the distribution of the computing power among honest and dishonest miners through the calibration of the arrival times that generate the aforementioned counting processes. Theorem 3 is useful to advise the merchant on how many subsequent blocks should be added to the chain before shipping the goods. Theorems 1 and 2 enable us to determine the time at which the double-spending attack is most likely to occur. This is helpful to provide merchants with guidelines on how long they should wait before shipping goods, which compliments the advice on the number of blocks.

Throughout this paper, the lag $z > 0$ between the honest and malicious chains is assumed to be deterministic. In practice, because the length of the dishonest chain is unknown to the vendor at the time of the shipping, the delay should be assumed to be stochastic. From a mathematical point of view, the question is when do we initialize the counting processes $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$? If the starting time is the time at which the goods are shipped then Z and the counting processes are independent. Thus, we can apply the law of total probability together with the formulae derived above. We have an atom at 0 with probability mass $\mathbb{P}[Z \leq 0]$. If we initialize the counting processes at the time of issuance of the transactions then Z and the counting processes are linked. The vendor is asked to wait for say $k \in \mathbb{N}$ blocks to be added to the blockchain before shipping the goods. The honest chain is then ahead by $Z = k - M(T_k)$, where T_k is the arrival time of the k th block in the honest chain. The time interval $[0, T_k]$ is a burn-in period during which it does not matter which chain is leading. If $\{N(t), t \geq 0\}$ and $\{M(t), t \geq 0\}$ are Poisson processes then T_k has a gamma distribution, the stopped processes at T_k are again Poisson processes, and $M(T_k)$ is governed by a negative binomial distribution which is consistent with Rosenfeld's findings [32]. The general case will be studied in upcoming research.

The success of the blockchain method has resulted in bitcoin becoming increasingly popular and inspiring other electronic payment methods. It is worth mentioning that the results derived in this paper may be relevant to understanding other systems where similar blockchain policies are used.

Selfish mining, described, for instance, in [12], [14], and [33], is another type of miners' misconduct. Nowadays, it is no longer feasible to mine BTCs in isolation. Empirical evidence shows that BTC miners behave strategically by gathering in pools. All members contribute to the solution of each cryptopuzzle, and share the rewards proportionally to their contribution. Selfish mining is a strategy that can be used by a minority pool to obtain more revenue. The key idea is for the pool of selfish miners to keep its discovered blocks private while honest nodes continue to mine on the public chain. Assuming that the private chain has the lead over the public chain, when the public branch approaches the selfish miners' chain, the private chain is released publicly. It results in a waste of resources for all the miners but Eyal and Siren [12] showed that the revenue of the selfish miners goes beyond the revenue expected

when following the usual protocol given their share of the total computing power. The results presented here may be relevant in the context of a selfish mining attack as it boils down again to the race between two counting processes.

Appendix A. Appell and A-G polynomials

Appell and A-G polynomials are well known in mathematics. They can be used to solve various problems in statistics and applied probability. A short presentation is provided below. We refer the reader to, e.g. [25] for a review with applications in risk modeling. Let $U = \{u_i, i \geq 1\}$ be a sequence of real numbers, nondecreasing in our context. To U is attached a (unique) family of Appell polynomials of degree n in x , $\{A_n(x | U), n \geq 0\}$, defined as follows. Starting with $A_0(x | U) = 1$, the $A_n(x | U)$ satisfy the differential equations

$$A_n^{(1)}(x | U) = nA_{n-1}(x | U), \tag{36}$$

with the border conditions

$$A_n(u_n | U) = 0, \quad n \geq 1. \tag{37}$$

So, each $A_n, n \geq 1$, has the integral representation

$$A_n(x | U) = n! \int_{u_n}^x \left[\int_{u_{n-1}}^{y_n} \cdots \int_{u_1}^{y_1} dy_2 \right] dy_n. \tag{38}$$

In parallel, to U is attached a (unique) family of A-G polynomials of degree n in x , $\{G_n(x | U), n \geq 0\}$. Starting with $G_0(x | U) = 1$, the $G_n(x | U)$ satisfy the differential equations

$$G_n^{(1)}(x | U) = nG_{n-1}(x | \mathcal{E}U), \tag{39}$$

where $\mathcal{E}U$ is the shifted family $\{u_{i+1}, i \geq 1\}$, with the border conditions

$$G_n(u_1 | U) = 0, \quad n \geq 1. \tag{40}$$

So, each $G_n, n \geq 1$, has the integral representation

$$G_n(x | U) = n! \int_{u_1}^x \left[\int_{u_2}^{y_1} \cdots \int_{u_n}^{y_{n-1}} dy_n \right] dy_1. \tag{41}$$

Note that both polynomial families are sometimes defined without the factor $n!$ in (38) and (41). Of course, these polynomials are related through the identity

$$G_n(x | u_1, \dots, u_n) = A_n(x | u_n, \dots, u_1), \quad n \geq 1.$$

However, the two families (i.e. considered for all $n \geq 0$) are distinct and enjoy quite different properties. From (38) and (41), we may see that the polynomials A_n and $G_n, n \geq 1$, can be interpreted in terms of the joint distribution of the order statistics $(U_{1:n}, \dots, U_{n:n})$ of a sample of n independent uniform random variables on $(0, 1)$.

Proposition 1. For $0 \leq u_1 \leq \dots \leq u_n \leq x \leq 1$,

$$\mathbb{P}[U_{1:n} \geq u_1, \dots, U_{n:n} \geq u_n \text{ and } U_{n:n} \leq x] = A_n(x | u_1, \dots, u_n),$$

while for $0 \leq x \leq u_1 \leq \dots \leq u_n \leq 1$,

$$\mathbb{P}[U_{1:n} \leq u_1, \dots, U_{n:n} \leq u_n \text{ and } U_{1:n} \geq x] = (-1)^n G_n(x | u_1, \dots, u_n).$$

These representations play a key role in the first-meeting problems discussed in the paper. Numerically, it will be necessary to evaluate some special values of the polynomials. To this end, it is convenient to use the following recursive relations.

Proposition 2. *We have*

$$A_n(x | U) = \sum_{k=0}^n \binom{n}{k} A_{n-k}(0 | U)x^k, \quad n \geq 1, \tag{42}$$

where the $A_n(0 | U)$ are obtained recursively from

$$A_n(0 | U) = - \sum_{k=1}^n \binom{n}{k} A_{n-k}(0 | U)u_n^k, \quad n \geq 1. \tag{43}$$

The A-G polynomials are computed through the recursion

$$G_n(x | U) = x^n - \sum_{k=0}^{n-1} \binom{n}{k} u_{k+1}^{n-k} G_k(x | U), \quad n \geq 1. \tag{44}$$

Equations (42) and (43) follow from Taylor’s expansion of A_n , using also (36) and (37). Equation (44) follows from an Abelian expansion of x^n based on (39) and (40). Details are omitted here. Of course, the computing time increases with the degree of the polynomials. Note that

$$A_n(x | a + bU) = b^n A_n\left(\frac{x - a}{b} \mid U\right), \quad n \geq 1, \tag{45}$$

with the same identity for G_n . An important particular case in our study is when the parameters in U are random and correspond to partial sums of exchangeable random variables.

Proposition 3. *Let $\{X_n, n \geq 1\}$ be a sequence of exchangeable random variables of partial sums $S_n = \sum_{k=1}^n X_k$ with $S_0 = 0$. Then, for $n \geq 1$,*

$$\mathbb{E}[A_n(x | S_1, \dots, S_n) | S_n] = x^{n-1}(x - S_n), \tag{46}$$

$$\mathbb{E}[G_n(x | S_0, \dots, S_{n-1}) | S_n] = x(x - S_n)^{n-1}. \tag{47}$$

Proof. The identity (46) was derived in Proposition A.1 of [23], while the identity (47) follows from [18]. □

Acknowledgements

I want to thank the anonymous referee for carefully reading my work, making useful suggestions, and pointing out the work of Rosenfeld [32]. I am thankful to Stéphane Loisel, Alexis Bienvenue, and Patrick Laub for helping me improving the presentation of my paper. This work was initiated while I was visiting the Department of Statistics and Applied Probability at UC, Santa Barbara. I am thankful for the warm welcome I received. Special thanks to Andrea Angiuli who got me interested in cryptocurrencies and Jean-Pierre Fouque for his insightful remarks and comments.

References

- [1] AIGNER, M. (2007). *A Course in Enumeration* (Graduate Texts Math. **238**). Springer, Berlin.
- [2] ASMUSSEN, S. AND ALBRECHER, H. (2010). *Ruin Probabilities*. World Scientific, Hackensack, NJ.
- [3] BOROVKOV, K. AND BURQ, Z. (2001). Kendall's identity for the first crossing time revisited. *Electron. Commun. Probab.* **6**, 91–94.
- [4] BOROVKOV, K. A. AND DICKSON, D. C. M. (2008). On the ruin time distribution for a Sparre Andersen process with exponential claim sizes. *Insurance Math. Econom.* **42**, 1104–1108.
- [5] BOUCHERIE, R. J. AND BOXMA, O. J. (1996). The workload in the M/G/1 queue with work removal. *Prob. Eng. Inf. Sci.* **10**, 261–277.
- [6] BOUCHERIE, R. J., BOXMA, O. J. AND SIGMAN, K. (1997). A note on negative customers, GI/G/1 workload, and risk processes. *Prob. Eng. Inf. Sci.* **11**, 305–311.
- [7] BOWDEN, R. Intersection of the longest increasing subsequences. <https://github.com/rhysbowden/LIS>.
- [8] BOWDEN, R., KEELER, H. P., KRZESINSKI, A. E. AND TAYLOR, P. G. (2018). Block arrivals in the bitcoin blockchain. Preprint. Available at <https://arxiv.org/abs/1801.07447>.
- [9] DIMITROVA, D. S., IGNATOV, Z. G. AND KAISHEV, V. K. (2017). On the first crossing of two boundaries by an order statistics risk process. *Risks* **5**, 14pp.
- [10] DIMITROVA, D. S., KAISHEV, V. K. AND ZHAO, S. (2015). On finite-time ruin probabilities in a generalized dual risk model with dependence. *Europ. J. Operat. Res.* **242**, 134–148.
- [11] DIMITROVA, D. S., KAISHEV, V. K. AND ZHAO, S. (2016). On the evaluation of finite-time ruin probabilities in a dependent risk model. *Appl. Math. Comput.* **275**, 268–286.
- [12] EYAL, I. AND SIRER, E. G. (2014). Majority is not enough: Bitcoin mining is vulnerable. In *Financial Cryptography and Data Security: 18th International Conference*, Springer, Berlin, pp. 436–454.
- [13] GELENBE, E., GLYNN, P. AND SIGMAN, K. (1991). Queues with negative arrivals. *J. Appl. Prob.* **28**, 245–250.
- [14] GÖBEL, J., KEELER, H. P., KRZESINSKI, A. E. AND TAYLOR, P. G. (2016). Bitcoin blockchain dynamics: The selfish-mine strategy in the presence of propagation delay. *Performance Evaluation* **104**, 23–41.
- [15] GOFFARD, P.-O. (2017). Two-sided exit problems in the ordered risk model. *Methodology Comput. Appl. Probab.* 2pp.
- [16] GOFFARD, P.-O. (2018). *Online accompaniment for "Fraud risk assessment within blockchain transactions"*. Available at <https://github.com/LaGauffre/FraudRiskBlockchainTransaction>.
- [17] GOFFARD, P.-O. AND LEFÈVRE, C. (2017). Boundary crossing of order statistics point processes. *J. Math. Anal. Appl.* **447**, 890–907.
- [18] GOFFARD, P.-O. AND LEFÈVRE, C. (2018). Duality in ruin problems for ordered risk models. *Insurance Math. Econom.* **78**, 44–52.
- [19] HARRISON, P. AND PITEL, E. (1996). The M/G/1 queue with negative customers. *Adv. Appl. Prob.* **28**, 540–566.
- [20] IGNATOV, Z. G. AND KAISHEV, V. K. (2016). First crossing time, overshoot and Appell–Hessenberg type functions. *Stochastics* **88**, 1240–1260.
- [21] JAIN, G. AND SIGMAN, K. (1996). Generalizing the Pollaczek–Khintchine formula to account for arbitrary work removal. *Prob. Eng. Inf. Sci.* **10**, 519–531.
- [22] LEFÈVRE, C. AND LOISEL, S. (2009). Finite-time ruin probabilities for discrete, possibly dependent, claim severities. *Methodology Comput. Appl. Probab.* **11**, 425–441.
- [23] LEFÈVRE, C. AND PICARD, P. (2011). A new look at the homogeneous risk model. *Insurance Math. Econom.* **49**, 512–519.
- [24] LEFÈVRE, C. AND PICARD, P. (2014). Ruin probabilities for risk models with ordered claim arrivals. *Methodology Comput. Appl. Probab.* **16**, 885–905.
- [25] LEFÈVRE, C. AND PICARD, P. (2015). Risk models in insurance and epidemics: A bridge through randomized polynomials. *Prob. Eng. Inf. Sci.* **29**, 399–420.
- [26] MAZZA, C. AND RULLIÈRE, D. (2004). A link between wave governed random motions and ruin processes. *Insurance Math. Econom.* **35**, 205–222.
- [27] NAKAMOTO, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Available at <https://bitcoin.org/bitcoin.pdf>.
- [28] PERRY, D., STADJE, W. AND ZACKS, S. (2002). Boundary crossing for the difference of two ordinary or compound Poisson processes. *Ann. Operat. Res.* **113**, 119–132.
- [29] PERRY, D., STADJE, W. AND ZACKS, S. (2005). A two-sided first-exit problem for a compound Poisson process with a random upper boundary. *Methodology Comput. Appl. Probab.* **7**, 51–62.
- [30] PICARD, P. AND LEFÈVRE, C. (2003). On the first meeting or crossing of two independent trajectories for some counting processes. *Stoch. Process. Appl.* **104**, 217–242.
- [31] PURI, P. S. (1982). On the characterization of point processes with the order statistic property without the moment condition. *J. Appl. Prob.* **19**, 39–51.

- [32] ROSENFELD, M. (2014). Analysis of hashrate-based double spending. Preprint. Available at <https://arxiv.org/abs/1402.2009>.
- [33] SAPIRSHTAIN, A., SOMPOLINSKY, Y. AND ZOHAR, A. (2016). Optimal selfish mining strategies in bitcoin. In *Financial Cryptography and Data Security: 20th International Conference*, Springer, Berlin, pp. 515–532.
- [34] SHI, T. AND LANDRIAULT, D. (2013). Distribution of the time to ruin in some Sparre-Andersen risk models. *ASTIN Bull.* **43**, 39–59.
- [35] VAN DER HOFSTAD, R. AND KEANE, M. (2008). An elementary proof of the hitting time theorem. *Amer. Math. Monthly* **115**, 753–756.