Evidence of persistent order in the primes ’last digit’ sequence driving a Schramm-Loewner process

Theophanes Raptis$^1$ and Alberto Fraile$^2$

$^1$Physical Chemistry Lab, Chemistry Dept. University of Athens
$^2$Nuclear Futures Institute, Bangor University, Bangor, LL57 1UT, UK

August 18, 2021

Abstract

We report on a peculiar effect regrading the use of the prime’s last digit sequence which is equivalent to a quaternary symbolic sequence. This was used as a driving sequence for the recently introduced Schramm-Loewner Evolution after trying two different possible binary encodings. We report on a clear deviation from the standard space-filling curves normally expected from such a process. We also contrast this behavior with others produced via a symbolic dynamics applied on standard noise sources as well as deterministic sequences produced by simple automata. Our findings include the well known, Morse-Thue sequence as the simplest model exhibiting such behavior apart from some strongly biased Levy walks. We conjecture on an analytical condition for reproducing the particular effect based on Carleman linearization as well as on the possibility of certain continuous Langevin diffusion process that could reproduce similar behavior for appropriate noise source. We discuss the possibility of further experiments with supercomputers further corroborating these results.

1 Introduction

In previous work [1], [2] a set of new geometrical methods were proposed for the study of the primes numbers distribution and their associated last digits sequence resulting from taking their $\pmod{10}$ congruences. The latter sequence is known to be equivalent to a quaternary symbolic sequence and has been shown in [2] to have some unusual characteristics when projecting it as a 2D random walk via a special binarization filter. In the present work we use additional numerical evidence for the presence of a possible automatic structure based on a
direct comparison with "1/f" noise generators when used as a "driver" for the so called, "Schramm-Loewner evolution" (SLE), a special sequence of conformal maps on the complex plane.

The SLE were first introduced in [?] where they were termed "Stochastic Loewner Evolutions" due to much earlier work on the Loewner equation [3] and they were proven to contain the unique scaling limits of several known models of complex lattice based planar physical systems such as the Ising and percolation model, as well as loop-erased and self-avoiding random walks. Due to the arbitrary character of the driving function, an SLE process can be used to produce a very large variety of possible paths on the complex plane including both smooth and stochastic ones. Some efficient algorithms for computing such paths have been prescribed by Kennedy in [5].

In this work, we report on a particular use of one such process where the driving function is given in terms of a symbolic sequence in a binary alphabet selecting between the two possible branches of one of the general solutions of the Loewner equation. Two classes of such binary sequences are contrasted, one being a deterministic set including the primes last digit and, in second place, some representatives of random ones. The later were obtained after applying symbolic dynamics on certain known noise sources given an adjustable threshold on their main random variable domain.

In section 2, we briefly introduce the SLE algorithm and we comment on certain observations from numerical experiments. Specifically, we report on a peculiar extremal behavior presented by the last digit sequence when compared with other standard noise sources. We also report on certain properties that appear persistent in such a case while, adjusting the symbolic dynamics threshold over the domain of a standard Levy walk which reveals two possible extremes approaching either the real or the imaginary axis indefinitely. We only present evidence from relatively small samples at least for the last digit sequence and we stress the need of some supercomputing infrastructure to explore the problem just a few orders of magnitude farther.

In section 3, we attempt to elucidate some of the peculiar characteristics found from the driven SLE based on certain statistical properties related to the block length statistics as obtained via a run length analysis [6], a well known compression method. We also contrast these results with other known deterministic sequences to find that the simplest model reproducing a similar behavior appears to be the well known, Thue-Morse sequence. Based on the analytical structure of this sequence as a toy model of the observed behavior, we also derive an analytical condition for the subset of sequences with similar behavior using the Carleman linearization method [7].
2 Schramm-Loewner evolution as a complexity descriptor

An SLE is generally defined via the solutions of a stochastic differential equation which when defined on the upper half plain takes the form of a Loewner stochastic process as

\[ \partial_t f_t(z) = 2(f_t(z) - \xi(t))^{-1} \]  

(1)

with the additional condition \( f(\gamma(t)) = \xi(t) \) where \( \xi \) is a driving function taking values on the boundary and \( \gamma \) a planar curve. Additionally, the solution is supposed to satisfy a "hydrodynamic normalization" condition as \( f(z) \to z + cz^{-1} + O(z^{-2}) \) as \( z \to \infty \). We note that (1) can also adopt the form of a standard Langevin process via a change of variables \( w = f(z) - \xi \) as

\[ \partial_t w_t(z) = 2w_t(z)^{-1} - \xi(t) \]

(2)

There are two branches of such solutions given in the literature in the form

\[ f_{\pm} = (z \mp a)^a(z \pm b)^b \]

(3)

The exponents in (3) have a complicated dependence on a critical parameter given as

\[ a = \frac{(1 - \sqrt{16/(16 - \kappa)})}{2}, b = 1 - a \]

which can also be simplified as

\[ a = \frac{1 - c}{2}, b = 1 + c, c = 1/\sqrt{4 - \kappa/4} \]

where \( \kappa \) guides through different types of known processes like self-avoiding walks while for values greater than 8 it produces space filling curves over the plane.

An SLE interface is produced by the iterative application of the two possible conformal maps leading to sequences of the form \( f_{X_1} \circ f_{X_2} \circ ... \circ f_{X_n} \) where \( X_n \) stands for some dichotomic random variable choosing between branches and all paths start at \( z = 0 \). Each path is formed by a linearly increasing sequence of words or functional composites as

\[ f_{X_1} \]

\[ f_{X_2} f_{X_1} \]

\[ f_{X_3} f_{X_2} f_{X_1} \]

\[ ... \]

each contributing a new point in the resulting curve. It should be pointed out that similar curves can be obtained reading a sequence in reverse order which although irrelevant for certain random sequences it could influence some of the results from determinsitic ones like those tested in the second part of this work. Extensive testing with reversed sequences revealed no significant differences. To facilitate this iterative scheme it is possible to combine the two branches into a single expression utilizing the internal symmetry in the algebraic representation of \( f_{\pm} \) in the form
\[ f_\sigma = (z - \sigma a)^{\alpha} (z - \sigma a + \sigma)^{1-\alpha} \]  

using \( \sigma_n = 2X_n - 1 \) as the corresponding sign variable projecting from a Boolean value of \( X_n \) to the \( \pm 1 \) alphabet. An example Matlab code is already given in [8] which is easy to modify accordingly.

Here we treat an SLE as a possible indicator of underlying pattern complexity of the driving bitstreams entering (3) via the indices defining a separate branch each time. We then test the behavior of the resulting curves when driven from different noise sources versus the one given by the PLD sequence. The various noise generators are treated as arbitrary dynamical systems extracting from each a particular symbolic dynamics using a thresholding scheme via the assignment \( X_n \leftarrow r_n > t \) for some adaptable threshold \( t \) value.

In the case of the PLD sequence, we apply two alternative binarization protocols, the one given as a direct substitution of values in the set \{1, 3, 7, 9\} with the two block strings \{00, 10, 01, 11\} and a second one, where the original quaternary alphabet is split and the rule \{1, 3\} \rightarrow 0 and \{7, 9\} \rightarrow 1 is applied. We discriminate between the two referring to the second case as the reduced PLD sequence or RPLD. We observe no significant difference when altering the reduced protocol for all other possible assignments like say, \{1, 7\} \rightarrow 0 and \{3, 9\} \rightarrow 1.

The resulting curves were compared with four other possible noise sources appropriately thresholded, including a flat and a Gaussian noise source with thresholds 1/2 and 0 corresponding to unbiased, equiprobable density binary sequences. These were also contrasted against certain 1/f noise sources also known as "pink" or "flicker" noise and a thresholded Levy Walk (TLW) with a variable threshold taken in \([\epsilon, \infty)\) for some small \( \epsilon > 0 \). A standard implementation of a noise source for both white, Brownian and pink noise exists in the harmonic analysis package LTFAT[9]. Some immediate conclusions for values of \( \kappa \) in any range were as follows.

- The standard flat and Gaussian sources give the expected types of random curves resembling random walks on the complex plane. This is also reflected in fig. 1(b) with respect to a critical behavior of the imaginary part as explained below.

- The RPLD case behaves in a more erratic manner with an increasing value of the \( \kappa \) parameter while the simple PLD effect persists even beyond \( \kappa = 8 \) where one expects a space filling curve. In what follows we will restrict attention to the RPLD, the main reason being that we cannot at the moment work with very large samples and there may be anomalies in the standard PLD sequence that do not appear in the RPLD case at least for the sample sizes we were able to check.

- The case of both the PLD and RPLD sequences is separate from all other tested being uniquely characterized by the asymptotic condition \( \text{Im}(z) >> \text{Re}(z) \). In both cases the real part appears confined in a narrow strip around the real axis although this has not been verified for extremely large samples. In figure 1(a) we
show an example of a 10000 points SLE created from an equal number of primes using the RPLD sequence, showing the SLE curve growing rapidly along the imaginary axis. We only show a small sample for convenience of visualization.

- In both cases, the imaginary part always evolves approximately like $\sqrt{an}$ for some $a < 1$ where $n$ the iteration index for each new point. In figure 1(b) we show an overlap of the imaginary parts for all SLEs where the maximal blue curve corresponding to the RPLD sequence which exhibit the smoothest possible evolution with respect to all the others.

- The last two $1/f$ noise sources exhibit a complementary behavior which appears to be asymptotically "orthogonal" to the previous class satisfying $\text{Im}(z) \ll \text{Re}(z)$ growing fast towards either the positive or the negative real axis. The case of the TLW has a additional sensitive dependence on the value of the threshold parameter $\epsilon$. As the parameter varies towards zero it also falls near the real axis while moving beyond some value near 0.2 the resulting curves tend to follow almost parallel to the imaginary axis.

- The two complementary behaviors appear to exhibit a statistical imbalance since both the PLD and RPLD sequences are only two unique strings, fixed over the total of $2^L$ sampled by other noise sources for any length $L$. Even the case of highly biased TLWs only select a special subset of strings hence one of the two extremes appear to be part of a "thin" subset.

A simple way to position any such process in a way that makes the extremalities observed easily visible seems to be via the introduction of a hyperbolic model for characterising noise sources via their proximity to a set of geodesics. This is possible using a class of conformal maps of the upper half plane on the unit disk similar to the construction of Smith charts in RF engineering which carries the same properties with
Figure 2: Examples of a Cayley map of both the RPLD sequence and two TLW driven SLE with $\kappa > 0.2$

the Poincare disk model for hyperbolic geometry [10]. This is often given in terms of some complex impedance $z_0$ as

$$z_\Gamma \leftarrow (z - z_0)(z + z_0)^{-1}$$

The special case of $z_0 = i$ is known as the Cayley transform while in engineering applications one often chooses $z_0 = 1$ with (5) representing the voltage reflection coefficient or $\Gamma$ and its inverse $z = (1 + \Gamma)(1 - \Gamma)^{-1}$ standing for the load impedance. In figure 3 the projections of the SLE for both the RPLD sequence and a typical $1/f$ noise source are shown for comparison. Due to the inevitable distortions of any such projection we normalized each SLE with the maximal element of its least varying part, thus taking $z_{RPLD} \leftarrow z_{RPLD}/\max(Re(z_{RPLD}))$ and $z_{TLW} \leftarrow z_{RPLD}/\max(Im(z_{RPLD}))$. The rationale behind this is that whenever we have a condition like $\text{Im}(z) \gg \text{Re}(z)$ and vice versa, then the same holds true for maximal values so that the tendency towards one of the two geodesics gets more pronounced.

To check the possibility of a hidden symmetry between $1/f$ processes and the deterministic RPLD sequence, we also checked a possible inversion under axes reversal which corresponds to the involution $\hat{L}(z) = iz^*, \hat{L}^2 = Id$ where star denotes transposition, effecting an exchange of real and imaginary parts. While this is not a symmetry of the original solution (3), our purpose was to check whether mixing of terms could result in an asymptotic condition like $|f(\hat{L}(z)) - \hat{L}(f(z))| < \epsilon$ for some small $\epsilon$. Results were negative for all noise sources and arbitrary values of the $\kappa$ parameter.

This shows the origin of the two complementary types of behavior to be rather different in origin. While we do not currently have
an analytical explanation of where the persistence of the PLD/RPLD sequences effect originates, we can make a comparison of the spectrograms using LTFAT’s short-time Fourier transform facilities. Figures 3(a - b) shows two samples, the first being a typical "white" noise source and the second being RPLD exhibiting a "lighter" colormap than a standard flat distribution thus implying an overall modulation over the spectral range. In the next section we discuss the possible origin of this effect in the overall block length statistics, usually associated with the run length analysis of symbolic sequences.

3 Searching for possible underlying order

The case of TLW offers more intermediate possibilities due to the potential of lowering the threshold asymptotically towards zero thus effecting also a drop of the overall curve obtained inside a small strip above the real axis. Hence, the particular class of TLW appears to be able to cover a broad spectrum of behaviors between the two extremes found in the previous section. Motivated by this we also examined an increasing bias on other both, white, Gaussian and Brownian noise generators by moving the threshold values appropriately the result being always one of the two extreme behaviors always restricting the curves obtained near the real axis. This leads deducing the existence of some rare, "thin" subset of possible binary strings among the plethora of $2^N$ total for $N >> 1$ of whose the specific block structures reflect the particular properties to which an SLE process responds similarly with the PLD/RPLD effect.

Another way to discriminate between sequences with different pattern structures is given in terms of it’s run length analysis by measuring the block lengths of consecutive one and zero blocks or alternating $\{ \pm 1 \}$ symbols. This always results in another alternating sign sequence of higher integer values for any binary input sequence. Making a pre-

Figure 3: Spectrograms for (a) white noise sequence and (b) the RPLD sequence.
Figure 4: Comparison of frequencies for consecutive zero or one block lengths for the symbolic dynamics of various noise sources vs the RPLD sequence.

A precise integer histogram of the results for a large sample of both the RPLD with 78,500 primes and other noise sources including the TLW for threshold values adapted in the two extreme classes shows similar statistics as in figure 4, apart from two excess values due to proliferation of single symbol blocks.

We should stress for the reader that the apparent appearance of a power law in figure 4, is not related to some so called, "Zipf’s Law" [11] which associates frequencies with sub-word lengths. Instead, the frequencies appearing there denote repetitiveness of same symbols. In particular, the negative X axis was used for the frequencies of same "zero" blocks and the positive axis for same "ones" blocks. The over-abundance of single symbol blocks for the RPLD sequence, suggests the origin of the extremalities observed to be related with noise sources that are singular near the origin. One can readily verify that any constructible sequence with large blocks of same symbols leaving the trajectory on the same branch tend to create curves that slowly fill the plane while, deliberately constructed sequences with rapidly changing small blocks prevent the space filling character of the resulting curves.

To understand this effect, we also tried a very special noise source known as the "Mittag-Leffler noise" [12]. This comes from a special transformation of a uniform random variable leading to distributions that are non-singular near the origin while still exhibiting the $1/f$ fat tails asymptotically. Indeed in such a case, the extremalities of the standard TLW case disappear, allowing us to conjecture that what the SLE effectively classifies is the presence of a singular distribution of the sampled noise source via its symbolic dynamics. The two remaining questions from the previous section concern the appearance of the persistent $\sqrt{n}$ behavior of the imaginary part and the deeper connection
between deterministic and sampled noisy sequences.

The deterministic PLD/RPLD case must then be contrasted not only with the symbolic dynamics of noise sources but also certain sequences produced by some much simpler automata based on substitution rules. A thorough analysis of the properties of such automatic sequences has been presented in a classical monography by Allouche [13]. We used some of the most known examples, including the Morse-Thue (MT), the Baum-Sweet (BS) and the Rudin-Shapiro (RS) sequences for which we know the existence of simple production rules thus being of lesser complexity than the PLD/RPLD case. Surprisingly, all three types of behavior are then being observed. Specifically, the RS sequence is the one exhibiting the space filling character expected from high values of the $\kappa$ parameter, while the BM and the MT ones, correspond to the opposite sides of the spectrum with the later exhibiting the most smooth evolution of the imaginary part close to the $\sqrt{n}$ curve while the real part exhibits a kind of dissipative oscillation, asymptotically approaching zero as shown in figure 5.

This last coincidence as well as the smootheness of the observed curve for the imaginary part enhance the rough explanation of the previous paragraphs around figure 4, since the run length statistics of the MT sequence is entirely different than that of all previous cases with only blocks of single and double symbols, either ones or zeros and with a perfectly symmetric histogram. On the other hand, it may allow a different interpretation for the origin of the RPLD persistence effect as a superposition of two processes, one automatic and the other stochastic.

We note in passing that recent intensive research by Ninagawa suggesting the association of $1/f$ statistics and computational universality
or "Turing-completeness" were presented in [14], [15]. Since several different types of universal automata appear to exhibit similar behavior, one conjectures that the general class of such noise sources may include Turing complete mechanisms as a special subclass. This then implies the possibility that deterministic sequences like the PLD/RPLD may hide a special hierarchy of subwords corresponding to some such hidden automaton associated at least with large fragments of this sequence. This idea of coexistence is also inspired by recent work [16] where concrete proofs have been given regarding the possibility of turning in principle uncomputable properties into computable ones in the presence of injected noise.

At the moment we conjecture that the MT case could serve as a toy model of the observed RPLD behavior apart from some additional contribution of unknown origin. Additionally, the abstract definition of the MT sequence is given in terms of one of the most fundamental integer sequences characterising the binary representations of the integers with well known fractal characteristics, the so called, "Digit-Sum" [17] or Hamming weights sequence with lots of applications in combinatorics and computer science [18]. Given a polynomial representation of every natural number we then write this sequence with the aid of a local bit value map as

\[ w_H(\nu) = \sum_i \left\lfloor \frac{\nu}{2^i} \right\rfloor \] (6)

where \( L = \lceil \log_2(\nu) + 1 \rceil \) the maximal power of 2 present. Given this sequence, the MT sequence is directly obtained as \((-1)^{w_H(\nu)} \cong \exp(i\pi w_H(\nu))\) in the \(\{\pm 1\}\) alphabet or simply \(\text{mod}(w_H(\nu), 2)\) in the standard binary one. As a matter of fact, the latter is also identical with the binary parity of bit strings in common use for computer science applications including error correction [19].

In previous work [20], it was pointed out by one of the authors that many sequences computed by simple automata acting on a lexicographically ordered powerset of all strings of constant length, also termed a "Hamming Space", may inherit an internal self-similarity out of the periodicities involved which can be seen as a type of semi-direct product between a set of dilations and a set of automorphisms of the \(\mathbb{Z}_2\) group. This then allows an arithmetized form of any such based on an iterated list concatenation system utilising much simpler "reproducing maps" of lower complexity in the form

\[ L_n \leftarrow [L_{n-1}, M(L_{n-1})] \] (7)

The correspondence between (6) and (7) reduces to the choices \(M(x) = x + 1, L_0 = 0\) where \(M(x)\) is to be applied pointwise to every element of the previous list. For the associated MT sequence in the \(\{\pm 1\}\) alphabet one simply has to use \(M(x) = -x, L_0 = -1\). Based on this one can write any formal sequence of conformal map compositions as
Simple inspection of the list concatenation scheme in (7) shows that any symbol sequence follows as $M^{w_H(i)}$ where $i$ the composites enumeration index. Since the action of $M$ for the MT sequence is identical with that of the $Z_2$ group, in the above we have a case of a group modulation similar to the case of a semi-direct product between the set of automorphisms of $Z_2$ and the discrete conformal group [21].

Direct comparison of MT with the RPLD sequence comes from the smoothened form of the $\sqrt{n}$ behavior of the imaginary part which persists in both cases. It may then help to split the functional form of (4) into a generic map and a perturbed variable to gain a better understanding of the influence of the driving sequence as

$$f_\sigma = (z - \sigma a + \sigma) \left(1 - \frac{\sigma}{z - \sigma a + \sigma}\right)^a = Z_\sigma \left(1 - \sigma \frac{Z_\sigma}{|Z_\sigma|^2}\right)^a \quad (8)$$

$$Z_\sigma = z + \sigma (1 - a)$$

This way the original can be rewritten via the auxiliary function $G(z) = z(1 - \sigma/z)^a$ as $f_\sigma = G(Z_\sigma)$. Use of the intermediate variable $Z_\sigma$ helps to reanalyze the dynamics in (8) as a system with both additive and multiplicative noise in the update rule

$$Z_\sigma \leftarrow G_\sigma(Z_\sigma, \sigma) + \sigma (1 - a)$$

changing also the initial condition to $Z_\sigma^{(0)} \leftarrow 0 + \sigma (1 - a) = \pm b$. We notice that there is a set of fixed points of $G$ due to the roots of $G(z) = z \sim 1^{1/b} \approx \exp(2\pi i/b)$ which is dense on the unit circle. However, when perturbed by the externally driven variable $Z_\sigma$ these sets become unstable. According to the observational evidence for the MT sequence, the observed extremes must correspond to a pair of conditions such that the real part of the $Z_\sigma$ sequence become contractive and the imaginary part follows a $\sqrt{n}$ curve.

An effective method for linearizing recursions was first proposed by Carleman [7] and later expanded by Jabotinsky [22], [23]. The method was rediscovered independently in the analytic solution of the logistic map recursion [24], [25] and only later it was recognised that the two derivations were identical [26]. In the original method, any finite $N$-dimensional vector field defining a dynamical system is transformed into an infinite dimensional linear system by expanding over a set of products of auxilliary variables which is formally solvable. The original solution is then obtained via a projection from the infinite dimensional space over the first $N$ variables. In case of discrete systems functional composition is exchanged with infinite matrix representations following a transfer function logic, each matrix providing a symbolic representation of the original function when multiplied with a right row vector of all powers of $x$. These are abstractly defined via Taylor expansions and satisfy $\mathcal{T}(f \circ g) = \mathcal{T}(f) \cdot \mathcal{T}(g)$ for arbitrary $C^n$ functions $\{f, g\}$. Similarly, a recursion over any function $f$ can be written as a power $\mathcal{T}^n$. 

11
Derivation of such matrices requires special generating functions for the rows and in our case we have the additional difficulty of fractional exponents forcing infinite expansions in all orders.

It is still possible to extract a necessary condition for the MT behavior using only abstract properties. Specifically, for a discrete map like $G(z)$ applied to the SLE algorithm of the previous section, successive words represent a sequence of linearly increasing functional composites producing each point in the resulting trajectory. Hence, each of these words can be replaced with an infinite hierarchy of Krylov subspaces [27] from successive powers of the corresponding Carleman matrix. By convention all such matrices start with a row of the form $[1, 0, ...]$. Projection to a specific value for any map trajectory is done via a left row vector $e = [0, 1, 0, ...]$ isolating all elements of the second row of each $T^n$ sequence. Thus one can always construct a new total matrix formed by these rows as $T_i \leftarrow e \cdot T^i$. This now is in a one to one correspondence with the functional composite leading to each point of the original trajectory. In the case of the SLE using the $G(z, \sigma)$ version, we may interpret the noise terms as limiting cases of a unimodular variable $e^{i\phi}$. By fixing the input vector we can write a generic condition as

$$\pm b\left(T \cdot 1\right) \rightarrow f^i + i\sqrt{a\pi}$$

(9)

where $f^i$ stands for some rapidly decaying oscillatory function for the real part as observed in the MT case. The above is to be read as a set of limiting paths for the additional parameter $\phi$ leading to the desired behavior.

While the above observations may serve for the abstract analysis of the MT driven SLE we may also provide an alternative scheme for the passage from the MT toy model to the PLD/ RPLD case. Based on the idea of symbolic dynamics and the deterministic character of the RPLD, we may extend the idea of superposition of two sources by an appropriate redefinition of the original expression in (6). For this we utilise an alternative representation expressing the same as a sampling of an underlying continuous set of dynamical systems otherwise known as a "lacunary" set of Fourier modes with an exponential modulation of periods also termed the "Rademacher System"[28]. Given a sampling operator $\hat{S}$ over any harmonic function such that the sign is interpreted as a unique symbol for each half-period one can also write

$$w_M(\nu) = \sum_i \hat{S}(\sin(\pi 2^i t))$$

(10)

We can then associate the above with a corresponding ODE system of independent oscillators $\ddot{x}_i = -(\omega_i^2/2)x_i$ with $\omega_i = \pi 2^i$ and assume the existence of a frictionless Langevin system with some arbitrary noise source in the form
\[ dx_i = p_i dt \]  
\[ dp_i = -\left(\frac{\omega_i^2}{2}\right)x_i dt + \sqrt{\kappa}\sigma_i(t) \]

One can then conjecture the existence of appropriate noise sources \( \{\sigma_i(t)\} \) such that the resulting symbolic dynamics of the collective variable \( < S(y_i) > (mod2) \) as represented by (6) could reproduce the crucial characteristics of the RPLD sequence or the strongly biased TLW responsible for the persistence effect observed. In principle, we assume an infinity of stochastic oscillators in (11) and (12) but the onset of the asymptotically persistent behavior seems to suggest that a sufficiently large number of them could be sufficient. This is a case of a rather difficult inverse problem for Langevin diffusion which does not appear to have a known general solution.

4 Discussion and Conclusions

Motivated by recent research in properties of the primes "last digit" sequence (PLD), we proceeded exploring its unusual properties based on the idea of using different dynamical systems as possible classifiers of complexity when driven by such a sequence. The case of the Schramm-Loewner evolution (SLE) was chosen based on both the possibility of driving via a dichotomic noise source and its sensitivity in any external modulating signal manifested in the geometric characteristics of the resulting path. We uncovered an unexpected type of behavior where the deterministic PLD randomness or some strongly biased Levy walks show a persistent tendency of aligning with one of the two axes in the upper half plane. A most striking characteristic was the appearance of a \( \sqrt{n} \) curve for the imaginary parts in all cases where alignment was towards the imaginary axis. We also contrasted this behavior with other simpler deterministic sequences of lesser complexity which revealed that the so called, Thue-Morse (MT) sequence may serve as a model for this type of behavior with the most clear sign of the square root shaped curve for the imaginary part.

We should stress the fact that our findings are only indicative at least regarding such deterministic sequences like the PLD, since we only applied this method for samples not exceeding some hundred thousands while in previous work as in [2] where supercomputers were made available, it was possible to check billions of primes. It is then still possible that such behavior is similar to a very large transient. This though, may not hold for the MT model sequence since its structure is much simpler based on an internal set of symmetries reflecting the associated sum-of-digits sequence. This is known to be associated with an increasing set of exponential intervals as \([0, ..., 2^n - 1]\) such that the resulting type of modulation over the SLE should not be expected to show any significant qualitative differences over increasing sample lengths.

The similarity of the observed persistent effect which strongly deviates from standard space filling curves usually met in SLEs was only
observed for the two cases of deterministic sequences, the PLD and the MT cases which does not allow to fully explore the nature of the special subset of sequences sharing similar properties. The only other case of strongly biased Levy walks was compared with other types of $1/f$ noise sources for which this effect fails allowing us to conjecture that the main characteristic associated with this effect appears to be the singular nature of their distributions near zero.

To deepen our understanding of the type of behavior observed, we also proposed an analytical condition for the SLE that should be generally applicable. This was based on the well known Carleman embedding of finite dimensional dynamical systems into an infinite dimensional linear system which allows rewriting any original system as a projection via a sequence of transfer matrix products. We postponed further examination due to the need for special symbolic software given the limited space of this report. We also conjectured the possibility of some continuous Langevin processes of whom the symbolic dynamics that could give rise to a special subset of sequences with similar statistical properties as for the MT model system.

References

[1] A. Fraile, R. Martinez, D. Fernandez, *Jacob’s Ladder: Prime numbers in 2d*, Math. Comput. Appl., 25, 5 (2020)

[2] A. Fraile, O. Kinouchi, et al., *Prime numbers and random walks in a square grid*, Accepted in ...

[3] O. Schramm, *Scaling limits of loop-erased random walks and uniform spanning trees*, Israel J. Math., 118: 221-288 (2000)

[4] C. Loewner, *Untersuchungen über schlichte konforme Abbildungen des Einheitskreises. I*, Math. Ann., 89 (1-2): 103-121 (1923).

[5] T. Kennedy, *Numerical computations for the Schramm-Loewner Evolution*, J. Stat. Phys. 137, 839-856 (2009)

[6] I. M. Pu, (2006) Chap. 3 in *Fundamental Data Compression*, Butterworth-Heinemann

[7] K Kowalski, W-H Steeb, (1991) *Nonlinear Dynamical Systems and Carleman Linearization*, World Sci.

[8] S. Boutiche

https://github.com/xsources/Matlab-simulation-of-Schramm-Loewner-Evolution

[9] http://ltfat.org

[10] M. S. Gupta *Escher’s art, Smith chart, and hyperbolic geometry* IEEE Microwave Magazine, 7-5, 66 -76 (2006)

[11] D. M. W. Powers, *Applications and Explanations of Zipf’s Law* Assoc. Comp. Linguistics: 151-160 (1998) in *New Methods in Language Processing and Computational Natural Language Learning*

[12] H.J. Haubold, A.M. Mathai, R.K. Saxena, *Mittag-Leffler Functions and Their Applications* J. Appl. Math., 2011, 298628 (2010)
[13] J-P. Allouche (2003) *Automatic Sequences* Cambridge University Press.

[14] S, Ninagawa, *Computational Universality and 1/f Noise in Elementary Cellular Automata* IEEE Xplore, Int. Conf. on Noise and Fluctuations (2013)

[15] S, Ninagawa, J. Martinez, *1/f Noise in the Computation Process by Rule 110*, J. Cell. Autom. 12(1-2): 47-61 (2016)

[16] M. Braverman, A. Grigo, C. Rojas, *Noise vs computational intractability in dynamics*, 128-141, 3rd Conference on Innovations in Theoretical Computer Science, ITCS 2012, Cambridge.

[17] L-E. Bouche, *An asymptotic formula for the average sum of the digits of integers* Am. Math. Monthly, 47 (3): 154-156 (1940)

[18] T. T. Thompson (1983) *From Error-Correcting Codes through Sphere Packings to Simple Groups*. The Carus Math. Monographs, Math. Assoc. America

[19] R. E. Ziemer, W. H. Tranter (2014) *Principles of communication: systems, modulation, and noise* Hoboken, NJ

[20] T. E. Raptis, *'Viral’ Turing Machines, Computation from Noise and Combinatorial Hierarchies* Chaos, Solitons, Fractals, 104, 734-740 (2017)

[21] V. Nekrashevych (2005) *Self-Similar Groups* Math. Surveys and Monographs, 117, Amer. Math. Soc.

[22] E. Jabotinsky, *Representation of Functions by Matrices. Application to Faber Polynomials* Proc. Amer. Math. Soc., 4(4): 546-553 (1953)

[23] E. Jabotinsky, *Analytic Iteration* Trans. Amer. Math. Soc., 108 (3): 457-477 (1963)

[24] S. Rabinovich, G. Berkolaiko, S. Havlin, *Solving nonlinear recursions* J. Math. Phys., 37(11) :5828 (1996)

[25] S. Rabinovich et al., *Analytical Solution of the Logistic Equation* Int. J. Bifurcatons, Chaos, 7(4): 837-838 (1995)

[26] G. Berkolaiko, S. Rabinovich, S. Havlin, *Analysis of Carleman Representation of Analytical Recursions* J. Math. Anal. Appl., 224:81-90 (1998)

[27] S. Valeria, (2015) *Krylov Subspaces* jn N. J. Higham, *he Princeton Companion to Applied Mathematics*, 113-114, Princeton University Press.

[28] S. Ashtashkin (2020) *The Rademacher System in Function Spaces*, Birkhäuser