ULAM SETS WHOSE INITIAL VECTORS ARE RANDOM

MAXIE D. SCHMIDT
SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GA 30332 USA
MAXIEDS@GMAIL.COM

1. Introduction to timing distributions

1.1. **Overview.** Kravitz and Stienerberger define generalized Ulam sets in higher dimensions for a set of initial starting vectors, $\{v_1, \ldots, v_k\} \subseteq \mathbb{R}^n_{\geq 0}$. We consider the case where we have two initial non-parallel vectors in the set $\{v_0, v_1\}$ typically drawn at random with real-valued entries in $[0, 1]$. That is, given two non-parallel vectors $\{v_0, v_1\}$ we define the next element of our generalized Ulam sets in 2D to be $v_0 + v_1$ and then we recursively generate the next element(s) of our Ulam set at the N^{th} step by choosing the vectors of the smallest two-norm (Euclidean norm) from the vectors that can be uniquely written as the sum of two distinct vectors already in the set at step $N - 1$.

In our special cases, we have by Theorem 1 in Section 2 of the article by Kravitz and Steinerberger that the Ulam sets with two initial vectors (chosen in some manner) correspond to the points of the form $v_0 + nv_1$ and $nv_0 + v_1$ for natural numbers $n \geq 0$ and the points on the inner lattice "*wedged-out*" by these two vectors of the form $mv_0 + nv_1$ for $m, n \geq 3$ both odd positive integers. Geometrically these points correspond to points on the boundary of the wedge formed by the two initial vectors, i.e., the points of the form $v_0 + nv_1$ and $nv_0 + v_1$ for $n \geq 0$, and the so-termed *inner lattice points*, i.e., the points of the form $mv_0 + nv_1$ for $m, n \geq 3$ both odd, which each lie on a distinct sloped line between two points on the bounding lines of the wedge of the form $v_0 + nv_1$ and $nv_0 + v_1$ for some $n \geq 3$. This leads to an infinite set of Ulam set points which are a subset of the lattice formed by multiples of the initial vectors and which are filled in *gradually* at different time steps $N \geq 1$.

1.2. **Questions about timing statistics and distributions.** Since the size of our Ulam sets are infinite as $N \to \infty$ we proceed to ask several natural questions about the times N where points at certain geometric positions within the set enter the full set of points. To the best of our knowledge such questions and conjectures relating to this subject matter have not yet been posed in the literature on generalized Ulam sets. In particular, we seek information about the timing distributions (over N) of the Ulam set points on a given inner liner segment bounded by the two vectors $v_0 + nv_1$ and $nv_0 + v_1$ for each $n \geq 2$. For example, given a list of times that the vectors on the n^{th} sloped inner line segments enter the set, we may consider the plots of summary statistics over n such as the minimum and maximum times for vectors on the line, arithmetic means of the times, geometric means, harmonic means, RMS of the times, the mode of the times, the median of the times, signal-to-noise of the times, and many, many other properties of these timing lines.

Other examples of summary statistics enabled in our code which we have not yet thoroughly plotted and examined in detail (see the examples in the next section) include the following: variations of Carleman means, Manhattan distance, variance, variation, standard deviation, absolute deviation, kurtosis, mean deviation, Pearson mode skewness, cubic means, interquartile ranges, contraharmonic means, averages of the log of the times to the base φ, coefficients of variation, interquartile means, several indexed k-L-moments, circular means, circular variances, circular standard deviations, and skewnesss. Needless to say, we have a large pool of conjectures we can make and analyses of the average cases when the initial vectors are drawn at random. Thus this is non-trivial topic matter we portend to study!

1.3. **Introducing the prototypical examples of our method.** We note that we normalize our statistics by the angle of the wedge formed by our two initial vectors to take into account that it takes longer to fill in the n^{th} line segment in the Ulam set lattice if the wedge angle is larger than if it is smaller in size. We perform this normalization to ideally lead to a limiting distribution of the plots we study which are (intended to be) independent of the two initial vectors in the average case. Prototypical examples of our method include the cases of the initial vectors

Date: Wednesday 28$^{\text{th}}$ June, 2017.

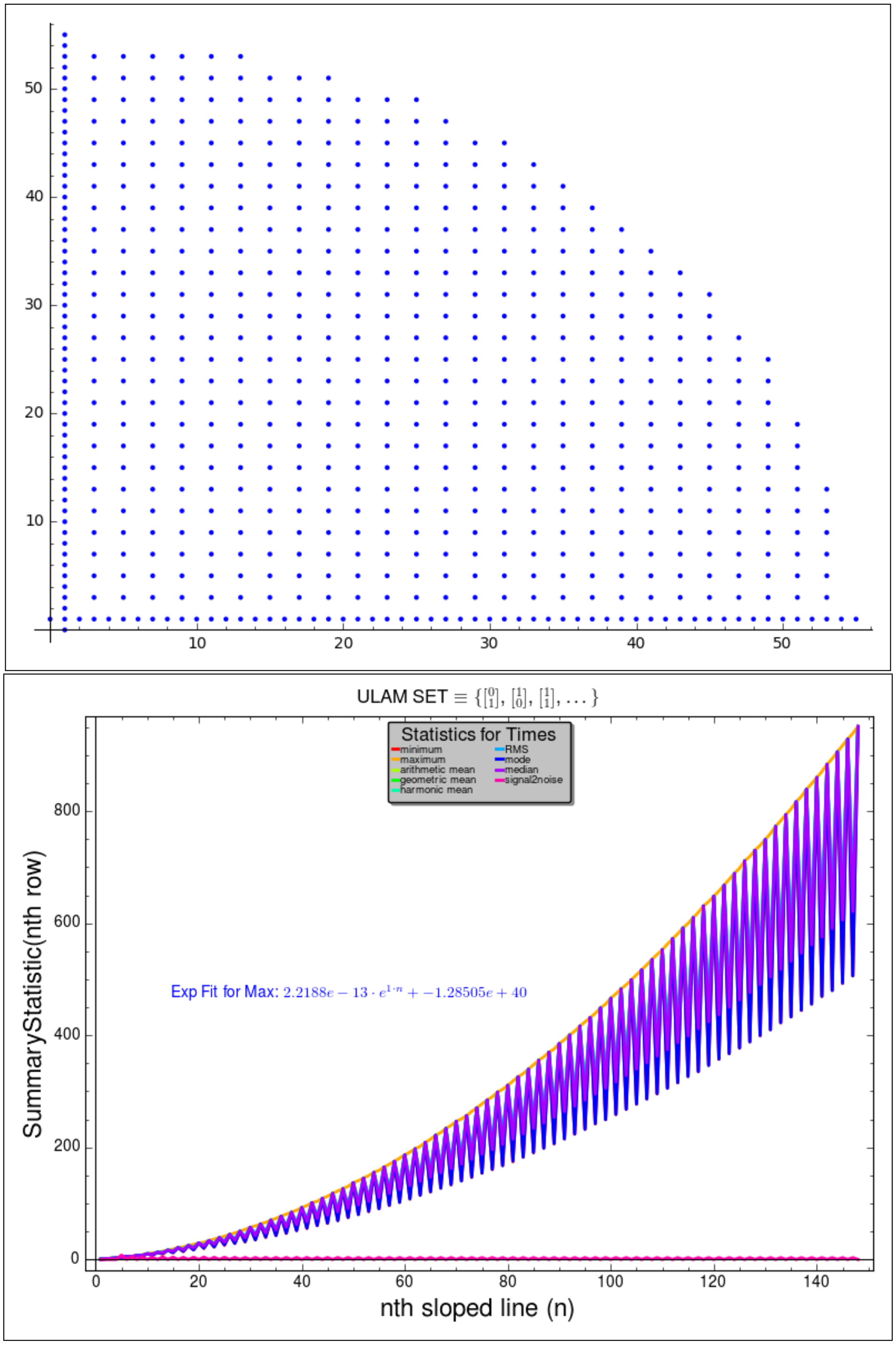


Figure 1.1. *The set arising from* $\{(0,1),(1,0)\}$ *when* $N := 250$ *and* $N := 1500$

$\{v_0, v_1\} := \{(0,1),(1,0)\}$ and $\{v_0, v_1\} := \{(1,\varphi),(\varphi,1)\}$. The Ulam set points and timing distributions in each of these sets when $N := 250$ ($N := 1500$) is shown in Figure 1.1 and Figure 1.2. Other examples of Ulam sets with randomly drawn initial vectors when $N := 1500$ are shown in the subsequent figures below.

Conjecture 1. We make one initial conjecture, which is that the maximum of the times at step n is exponentially distributed over n.

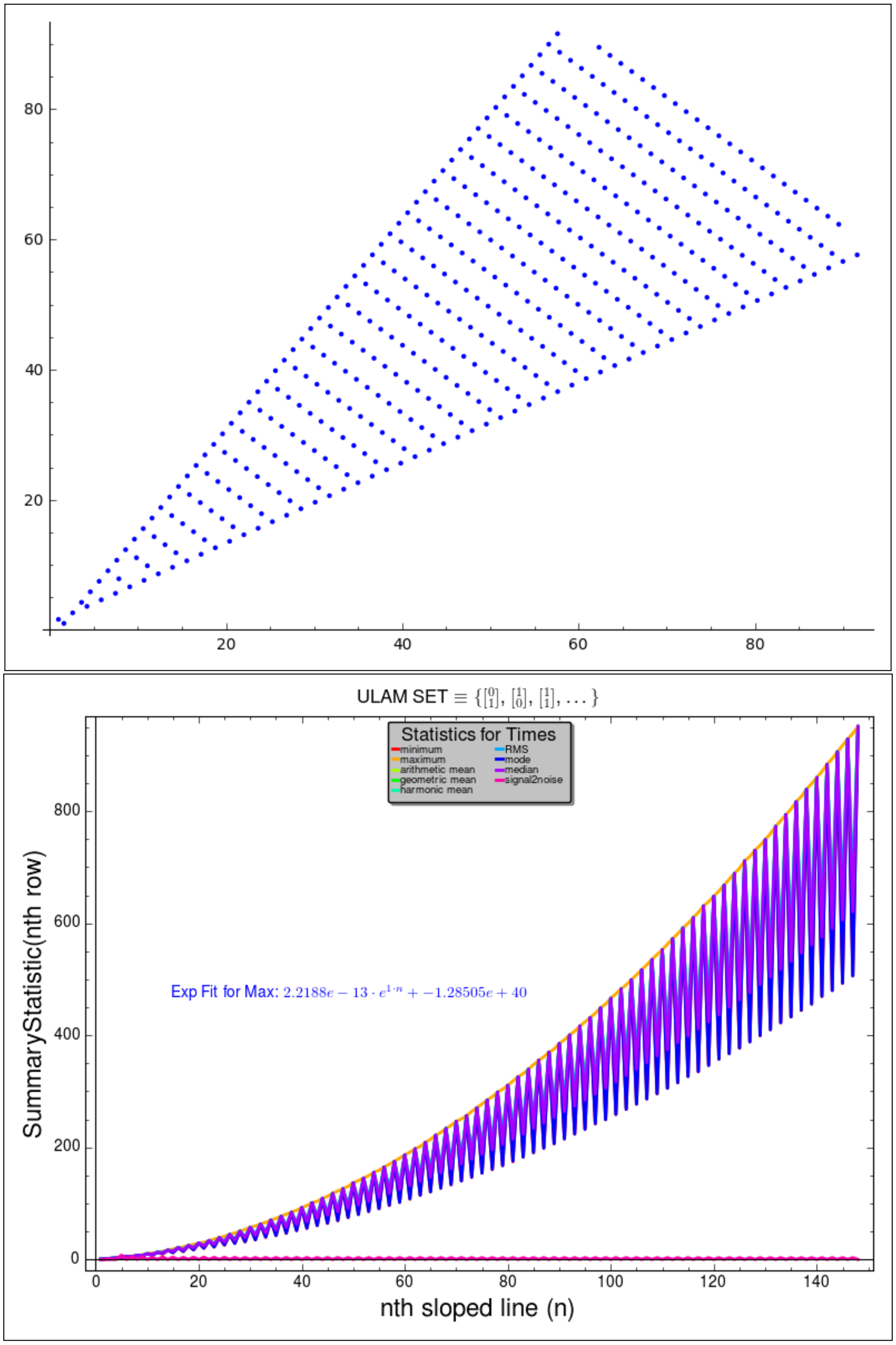
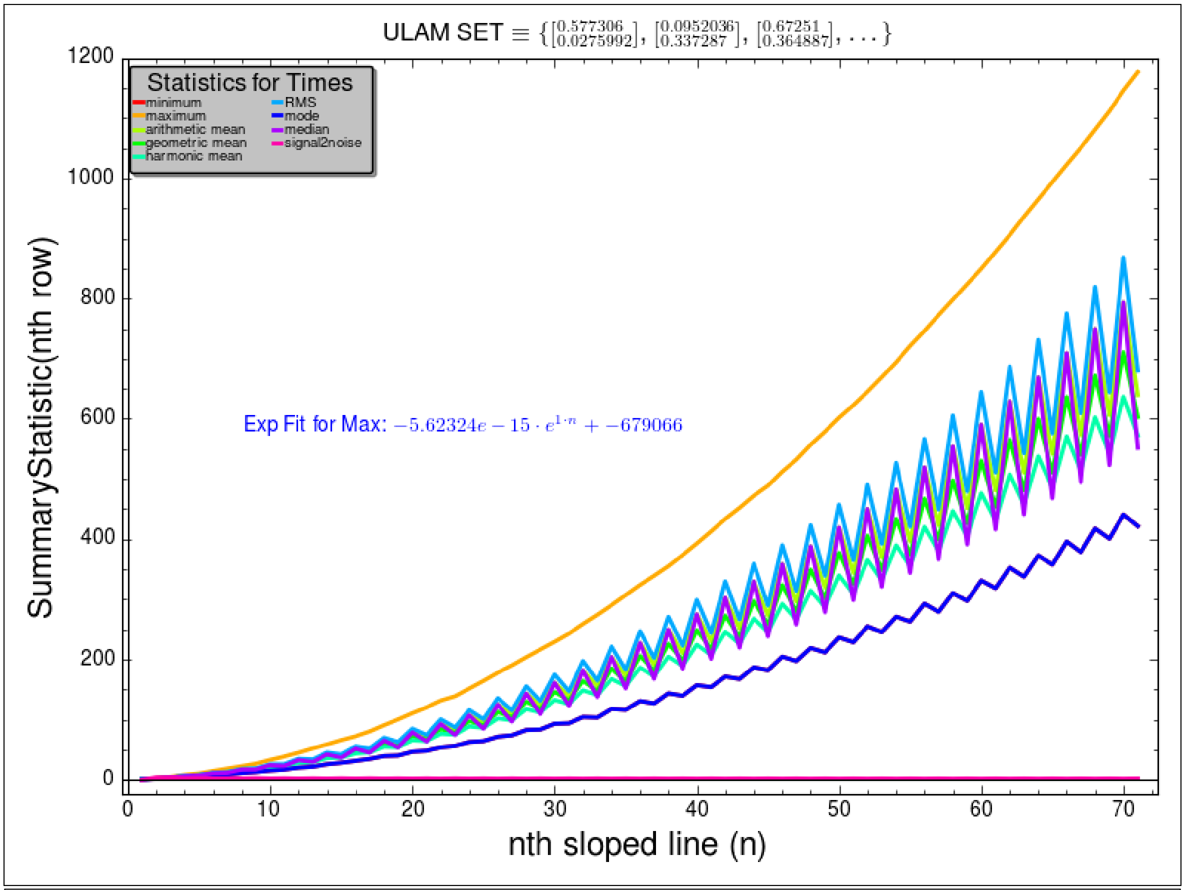


Figure 1.2. *The set arising from* $\{(1, \varphi), (\varphi, 1)\}$ *when* $N := 250$ *and* $N := 1500$



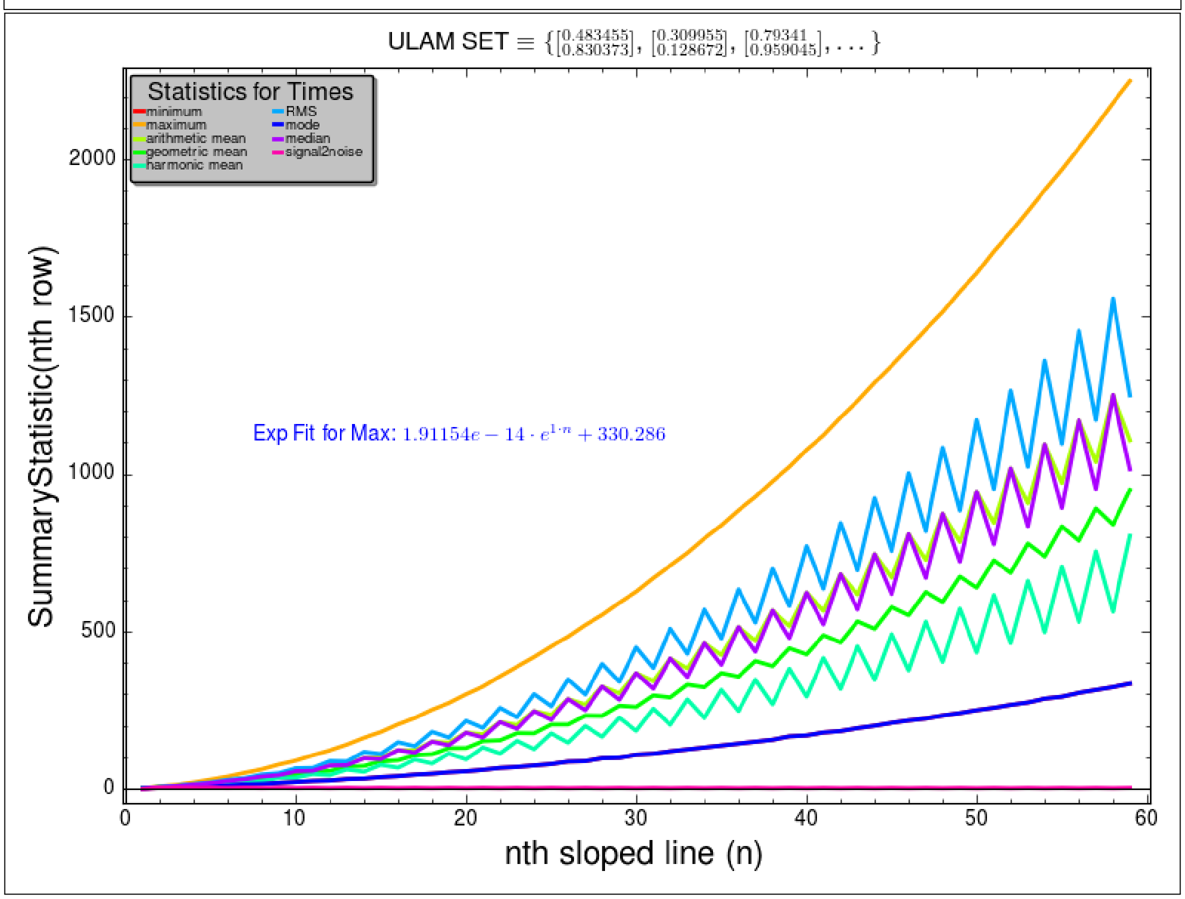


Figure 1.3. *Timing distributions for randomly drawn initial vectors when* $N := 1500$

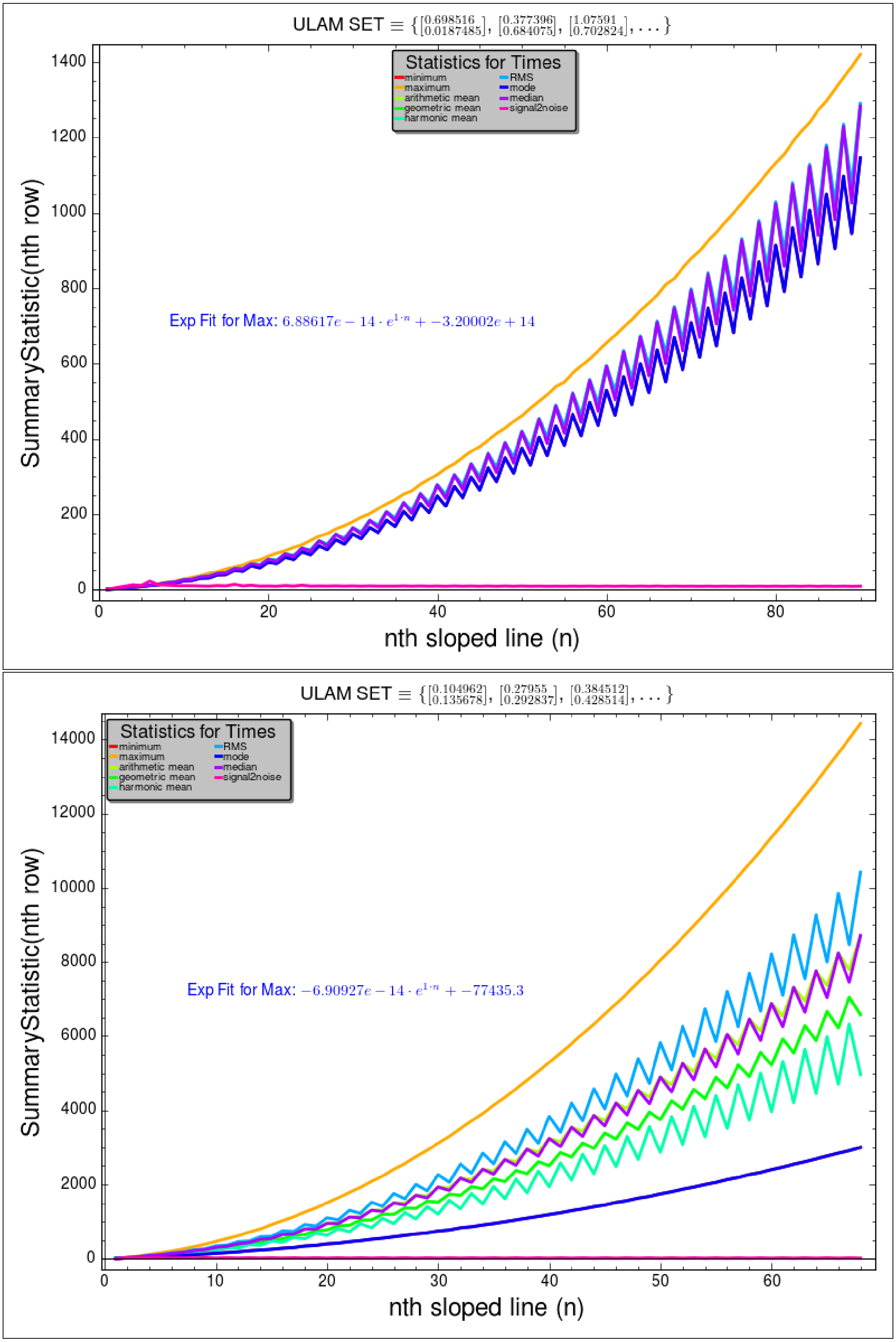


Figure 1.4. *Timing distributions for randomly drawn initial vectors when* $N := 1500$