

as the base increases. This is a finite size effect, because for large scales most values are not represented in short strings, so their probability is considered as zero. In very long strings, however, this effect is less noticeable, so long random strings have a high entropy for different scales ( $I_b \approx 1$ ), as shown in Table 2.

Table 1: Example of a 32-bit string. Grouping  $b$  bits, it can be scaled to different bases.

$b$	base	string	$I_b$
1	2	0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 1 1 0 0 1 0 0 0 1 1 0 0 1 0 0 0	0.89603821
2	4	0 0 2 0 2 2 0 1 3 0 2 0 3 0 2 0	0.8246987
4	16	0 8 10 1 12 8 12 8	0.5389098
8	256	8 161 200 200	0.1875

Table 2: Example of  $I_b$  for a long pseudorandom string at different scales.

$b$	base	length	$I_b$
1	2	$2^{20}$	0.9999998
2	4	$2^{19}$	0.9999997
4	16	$2^{18}$	0.9999956
8	256	$2^{17}$	0.9998395

$E$ ,  $S$ ,  $C$ , and  $H$  for different scales can be obtained converting binary strings to higher bases and normalizing their information with Eq. 10. The behavior is similar for all except  $H$ , since the Hamming distance measures the percentage of different symbols between strings. For binary strings ( $b = 1$ ), uncorrelated strings have a Hamming distance  $d \approx 0.5$ , since the probability of having the same symbol in each position is 0.5. For base 4 strings ( $b = 2$ ), this will be halved to  $d \approx 0.25$ , because the probability of having the same symbol in each position is also halved to 0.25. Thus, the lowest expected  $H$  (uncorrelated states) will decrease with increasing  $b$  in the form of  $\frac{1}{2b}$ .

Since  $I$  can change drastically depending on the scale at which it is observed (e.g. a string 10101010 has  $I_1 = 1$  but  $I_2 = 0$ , since the string in base 4 becomes 2222), all the proposed measures can also have drastic changes with a change of scale.