A PROOFS

This appendix contains the proofs of the theorems from Section 3, which are adapted from Saad et al. (2020, Section 3) and included here for completeness.

Proposition A.1 (Proposition 3.1 in main text)**.** *For integers k and l with $0 \le l \le k$, define $Z_{kl} := 2^k - 2^l \mathbf{1}_{l<k}$. Then*

$$\mathbb{B}_{kl} = \left\{ \frac{0}{Z_{kl}}, \frac{1}{Z_{kl}}, \dots, \frac{Z_{kl}-1}{Z_{kl}}, \frac{Z_{kl}}{Z_{kl}} \mathbf{1}_{l<k} \right\}.$$

Proof. For $l = k$, the number system $\mathbb{B}_{kl} = \mathbb{B}_{kk}$ is the set of dyadic rationals less than one with denominator $Z_{kk} = 2^k$. For $0 \le l < k$, any $x \in \mathbb{B}_{kl}$ when written in base 2 has a (possibly empty) non-repeating prefix and a non-empty infinitely repeating suffix, so that x has binary expansion $(0.b_1 \dots b_l \overline{s_{l+1} \dots s_k})_2$. Now,

$$2^l (0.b_1 \dots b_l)_2 = (b_1 \dots b_l)_2 = \textstyle\sum_{i=0}^{l-1} b_{l-i} 2^i$$

and

$$\begin{aligned}(2^{k-l} - 1)(0.\overline{s_{l+1} \dots s_k})_2 &= (s_{l+1} \dots s_k)_2 \\ &= \textstyle\sum_{i=0}^{k-(l+1)} s_{k-i} 2^i\end{aligned}$$

together imply that

$$\begin{aligned}x &= (0.b_1 \dots b_l)_2 + 2^{-l}(0.\overline{s_{l+1} \dots s_k})_2 \\ &= \frac{(2^{k-l}-1)\sum_{i=0}^{l-1} b_{l-i} 2^i + \sum_{i=0}^{k-(l+1)} s_{k-i} 2^i}{2^k - 2^l}.\end{aligned}$$ ☐

Remark A.2. When $0 \le l \le k$, we have $\mathbb{B}_{kl} \subseteq \mathbb{B}_{k+1,l+1}$, since if $x \in \mathbb{B}_{kl}$ then Proposition A.1 furnishes an integer c such that $x = c/(2^k - 2^l \mathbf{1}_{l<k}) = 2c/(2^{k+1} - 2^{l+1}\mathbf{1}_{l<k}) \in \mathbb{B}_{k+1,l+1}$. Further, for $k \ge 2$, we have $\mathbb{B}_{k,k-1} \setminus \{1\} = \mathbb{B}_{k-1,k-1} \subseteq \mathbb{B}_{kk}$, since any repeating suffix with exactly one digit can be folded into the prefix (except when the prefix and suffix are all ones).

Theorem A.3 (Theorem 3.2 in main text)**.** *Let T be an entropy-optimal DDG tree with a non-degenerate output distribution $(p_i)_{i=1}^n$ for $n > 1$. The depth of T is the smallest integer k such that there exists an integer $l \in \{0, \dots, k\}$ for which all the p_i are integer multiples of $1/Z_{kl}$ (hence in $\mathbb{B}_{kl}$).*

Proof. Suppose that T is an entropy-optimal DDG tree and let k be its depth (note that $k \ge 1$, as $k = 0$ implies p is degenerate). Assume $n = 2$. From Theorem 2.1, for each $i = 1, 2$, the probability p_i is a rational number where the number of digits in the shortest prefix and suffix of the binary expansion (which ends in $\bar{0}$ if dyadic) is at most k. Therefore, we can express the probabilities p_1, p_2 in terms of their binary expansions as

$$\begin{aligned}p_1 &= (0.b_1 \dots b_{l_1} \overline{s_{l_1+1} \dots s_k})_2, \\ p_2 &= (0.w_1 \dots w_{l_2} \overline{u_{l_2+1} \dots u_k})_2,\end{aligned}$$

where l_i and $k - l_i$ are the number of digits in the shortest prefix and suffix, respectively, of the binary expansions of each p_i.

If $l_1 = l_2$ then the conclusion follows from Proposition A.1. If $l_1 = k-1$ and $l_2 = k$ then the conclusion follows from Remark A.2 and the fact that $p_1 \ne 1$, $p_2 \ne 1$. Now, from Proposition A.1, it suffices to establish that $l_1 = l_2 =: l$, so that p_1 and p_2 are both integer multiples of $1/Z_{kl}$. Suppose for a contradiction that $l_1 < l_2$ and $l_1 \ne k-1$. Write $p_1 = a/c$ and $p_2 = b/d$ where each summand is in reduced form. By Proposition A.1, we have $c = 2^k - 2^{l_1}$ and $d = 2^k - 2^{l_2}\mathbf{1}_{l_2<k}$. Then as $p_1 + p_2 = 1$ we have $ad + bc = cd$. If $c \ne d$ then either b has a positive factor in common with d or a with c, contradicting the summands being in reduced form. But $c = d$ contradicts $l_1 < l_2$.

The case where $n > 2$ is a straightforward extension of this argument. ☐

Theorem A.4 (Theorem 3.4 in main text)**.** *Suppose p is defined by $p_i = a_i/m$ $(i = 1, \dots, n)$, where $\sum_{i=1}^n a_i = m$. The depth of any entropy-optimal sampler for p is at most $m-1$.*

Proof. By Theorem 3.2, it suffices to find integers $k \le m-1$ and $l \le k$ such that Z_{kl} is a multiple of m, which in turn implies that any entropy-optimal sampler for p has a maximum depth of $m-1$.

Case 1: Z is odd. Consider $k = m-1$. We will show that m divides $2^{m-1} - 2^l$ for some l such $0 \le l \le m-2$. Let ϕ be Euler's totient function, which satisfies $1 \le \phi(m) \le m-1 = k$. Then $2^{\phi(m)} \equiv 1 \pmod{m}$ as $\gcd(m, 2) = 1$. Put $l = m - 1 - \phi(m)$ and conclude that m divides $2^{m-1} - 2^{m-1-\phi(m)}$.

Case 2: m is even. Let $t \ge 1$ be the maximal power of 2 dividing m, and write $m = m'2^t$. Consider $k = m' - 1 + t$ and $l = j + t$ where $j = (m'-1) - \phi(m')$. As in the previous case applied to m', we have that m' divides $2^{m'-1} - 2^j$, and so m divides $2^k - 2^l$. We have $0 \le l \le k$ as $1 \le \phi(m) \le m-1$. Finally, $k = m' + t - 1 \le m'2^t - 1 = m - 1$ as $t < 2^t$. ☐

Theorem A.5 (Theorem 3.5 in main text)**.** *Let p be as in Theorem A.4. If m is prime and 2 is a primitive root modulo m, then the depth of an entropy-optimal DDG tree for p is $m-1$.*

Proof. Since 2 is a primitive root modulo m, the smallest integer a for which $2^a - 1 \equiv 0 \pmod{m}$ is precisely $\phi(m) = m - 1$. We will show that for any $k' < m - 1$ there is no exact entropy-optimal sampler that uses k' bits of precision. By Theorem A.4, if there were such a sampler, then $Z_{k'l}$ must be a multiple of m for some $l \le k'$. If $l < k'$, then $Z_{k'l} = 2^{k'} - 2^l$. Hence $2^{k'} \equiv 2^l \pmod{m}$ and so $2^{k'-l} \equiv 1 \pmod{m}$ as m is odd. But $k' < m - 1 = \phi(m)$, contradicting the assumption that 2 is a primitive root modulo m. If $l = k'$, then $Z_{k'l} = 2^{k'}$, which is not divisible by m since we have assumed that m is odd (as 2 is not a primitive root modulo 2). $\square$