Exploiting the Massey Gap

Andrei Tănăsescu and Pantelimon George Popescu*

Department of Computer Science and Engineering, University Politehnica of Bucharest, Splaiul Independentei 313 (6), 060042 Bucharest, Romania; andrei.tanasescu@mail.ru

*Correspondence: pgpopescu@yahoo.com

Received: 21 October 2020; Accepted: 7 December 2020; Published: 11 December 2020

Abstract: In this paper, we find new refinements to the Massey inequality, which relates the Shannon and guessing entropies, introducing a new concept: the Massey gap. By shrinking the Massey gap, we improve all previous work without introducing any new parameters, providing closed-form strict refinements, as well as a numerical procedure improving them even further.

Keywords: shannon entropy; guessing; massey inequality

1. Introduction

The guessing entropy associated with a probabilistic vector \( \mathbf{p} = (p_1, p_2, \ldots, p_n) \) with \( p_1 \geq \cdots \geq p_n > 0 \) is the expected value of the random variable \( G(\mathbf{p}) \) given by \( \Pr[ G(\mathbf{p}) = i ] = p_i \forall i \in \{1, n\} \), i.e., \( \mathbb{E} [ G(\mathbf{p}) ] = \sum_{i=1}^{n} i p_i \), and it corresponds to the optimal average number of binary questions required guessing the value of a random variable distributed according to \( \mathbf{p} \). Meanwhile, the Shannon entropy of \( \mathbf{p} \) is \( H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \log p_i \), and the first relation between them was provided by J. Massey in [1], namely that if \( H(\mathbf{p}) \geq 2 \), then \( \mathbb{E} [ G(\mathbf{p}) ] \geq 2 H(\mathbf{p}) - 2 + 1 \).

While the Massey inequality has been shown to have many applications such as password guessing [2] and encourages multiple developments such as moments inequalities (e.g., [3,4]), there are only two known direct refinements to the Massey inequality, both of which were published in 2019. In an ISITpaper, Popescu and Choudary [5] found that:

\[
\mathbb{E} [ G(\mathbf{p}) ] \geq 2^{2H(\mathbf{p})} + 2p_n - 2 + 1 - n \leq 1 / \inf \mathbf{p} < \infty.
\]

subject to the same conditions as the Massey inequality. Meanwhile, Rioul’s inequality [6], published in a CHESpaper [7], states that for all values of \( H(\mathbf{p}) \), we have:

\[
\mathbb{E} [ G(\mathbf{p}) ] > 2^{H(\mathbf{p})} / e
\]

a bound that is similar to the Massey inequality and refines it when \( H(\mathbf{p}) \geq \log \frac{e}{1-e/T} \).

Remarkably, the refinement in [5] is actually strict since \( \mathbf{p} \) has finite support, \( |\mathbf{p}| = n \leq 1 / \inf \mathbf{p} < \infty \).

In fact, applying the method from [5], it can be easily observed that Equation (4) could be further improved. For example, applying just the first step of the method presented in [5], we can obtain that for \( H(\mathbf{p}) \geq \log \frac{e}{2m} \), we have a new further refinement of Equation (4),

\[
\mathbb{E} [ G(\mathbf{p}) ] > 2^{H(\mathbf{p}) + p_n / e} - \frac{1}{2} p_n > 2^{H(\mathbf{p})} / e.
\]
Motivated by this insight, in this paper, we find and fully investigate a method to refine an inequality regarding finite support vectors using only itself. Here, we apply our analysis to the famous Massey inequality and explore multiple avenues of applying our method: shrinking the Massey gap.

2. Results

In this section, we present our results, as well as their proofs. In Section 2.1, we find a refinement over the initial result of [5], i.e., Equation (2), which we improve in Section 2.2 beyond even the strongest result of [5], i.e., Equation (1). Then, in Section 2.3, we introduce a new concept, the Massey gap, and use it to generalize the method in Section 2.2. Next, in Section 2.4, we show how using the Massey gap, we can almost always improve our refinement, by rebalancing the middle of the distribution, and finally, in Section 2.5, we find closed-form easy to compute refinements over the work of [5].

2.1. An Improved Refinement to the Massey Inequality

In this section, we find a new relation between the Shannon and guessing entropy, dependent on the minimal probability of a given distribution, refining the initial bound, Equation (2), from the work of [5].

Consider a decreasing probabilistic vector \( p = (p_1, p_2, \ldots, p_n) \) with \( H(p) \geq 2 \). Inspired by the trick in [5], we construct the new probabilistic vector \( q = (p_1, p_2, \ldots, p_{n-1}, (1 - \alpha)p_n, \alpha p_n) \), which is decreasing and strictly positive if and only if \( \alpha \in (0, 1/2) \). From the grouping property of entropy,

\[
H(q) = H(p) + p_n h(\alpha),
\]

where \( h(\alpha) = -\alpha \log\alpha - (1 - \alpha) \log(1 - \alpha) \) is the binary entropy function. Moreover, we calculate:

\[
\mathbb{E}[G(q)] = \mathbb{E}[G(p)] + \alpha p_n.
\]

Thus, because \( H(q) \geq 2 \), we can apply the Massey inequality for \( q \) to obtain:

\[
\mathbb{E}[G(p)] = \mathbb{E}[G(q)] - \alpha p_n \geq 2^{H(q) - 2} + 1 - \alpha p_n = 2^{H(p) + p_n h(\alpha) - 2} + 1 - \alpha p_n.
\]

Now, we note that \( h(0) = 0 \) and \( h(1/2) = 1 > \frac{1}{2} \ln 2 \approx 0.7213 \); thus, by the strict concavity of \( h \), we have \( h(\alpha) > \alpha \ln \frac{1}{2} \). Using the inequality \( 2^x > 1 + x \ln 2 \) for \( x = p_n h(\alpha) > 0 \), we find that:

\[
0 \leq p_n (h(\alpha) \ln 2 - \alpha) < 2^{p_n h(\alpha) - 1} - \alpha p_n
\]

\[
\leq 2^{H(p) - 2} \left( 2^{p_n h(\alpha) - 1} - \alpha p_n \right) = \left( 2^{H(p) + p_n h(\alpha) - 2} + 1 - \alpha p_n \right) - (2^{H(p) - 2} + 1),
\]

i.e., our lower bound is always strictly tighter than Massey’s,

\[
\mathbb{E}[G(p)] \geq 2^{H(p) + p_n h(\alpha) - 2} + 1 - \alpha p_n > 2^{H(p) - 2} + 1 \quad \forall \alpha \in (0, 1/2).
\]

For the particular value \( \alpha = 1/2 \), the bound above produces the initial bound, Equation (2), of [5]. Thus, we conclude this subsection with the following refinement to Equation (2) summarizing the above:

**Theorem 1.** For strictly positive descending probabilistic vectors \( p \in \mathbb{R}^n \) with \( H(p) \geq 2 \), we have:

\[
\mathbb{E}[G(p)] \geq \sup_{\alpha \in [0, 1/2]} 2^{H(p) + p_n h(\alpha) - 2} + 1 - \alpha p_n \geq 2^{H(p) + p_n - 2} + 1 - \frac{1}{2} p_n
\]

\[
> \inf_{\alpha \in [0, 1/2]} 2^{H(p) + p_n h(\alpha) - 2} + 1 - \alpha p_n = 2^{H(p) - 2} + 1.
\]
Proof. The first inequality follows taking the supremum over $\alpha$ in Equation (7). The second and third inequality follow taking $\alpha = 1/2$, which yields the right-hand side in Equation (2), where the strictness of the last inequality follows from the strictness of (7), the right-hand side of which is the bound obtained taking $\alpha = 0$. \qed

Up to this point, we found a direct refinement over the initial result of [5]. In the next section, we will leverage Theorem 1 in order to refine the strongest result, the bound in Equation (1).

### 2.2. Refinement through Generalization

In this section, we extend the reasoning of [5] to improve our bound from Theorem 1 beyond (1). We leverage a technique previously used in [8], namely applying the Massey inequality to increasingly complicated perturbations of the input in order to obtain ever tighter bounds for the original input.

Given an initial decreasing probabilistic vector $p$, we construct a list of probabilistic sequences $\{Q_k\}$, which we recursively define according to the procedure in the previous subsection.

We begin by fixing an arbitrary parameter $\alpha \in (0, 1/2]$ as in the section above. Now, we explicate our construction. Denoting by $Q_{k,i}$ the $i$th component of the sequence $Q_k$, we define the terms of the list $\{Q_k\}$ as follows. We let the support of the first term coincide with $p$, i.e., $Q_0 = (p_0, p_1, \ldots, p_n, 0, 0, \ldots, 0, \ldots)$, and we define the other terms by recurrence:

$$Q_{k+1} = (Q_{k,0}, Q_{k,1}, \ldots, Q_{k,n+k-1}, (1 - \alpha) Q_{k,n+k}, \alpha Q_{k,n+k}, 0, 0, \ldots, 0, \ldots).$$

Now, we expand the recursion relation for the Shannon entropy, $H(Q_{k+1})$, and guessing entropy $H[G(Q_{k+1})]$, which are analogous to Equations (5) and (6), respectively, to obtain:

$$H(Q_{k+1}) = H(Q_k) + Q_{k,n+k} h\left(\frac{1}{2}\right) + \alpha^k p_n h\left(\frac{1}{2}\right) = H(p) + p_n h\left(\frac{1}{2}\right) \left(1 + \alpha^k\right),$$

$$H[G(Q_{k+1})] = H[G(Q_k)] + \alpha Q_{k,n+k} = H[G(Q_k)] + \alpha^{k+1} p_n = H(G(p)) + p_n \sum_{j=0}^{k} a^j = H(G(p)) + p_n a \left(\frac{1 - \alpha^{k+1}}{1 - \alpha}\right),$$

so, at each step $k$, we can use $Q_{k+1}$ to obtain an approximation of $H[G(Q_k)]$ that is tighter than Massey’s. This procedure allows us to obtain ever tighter approximations of the guessing entropy.

For example, fixing an arbitrary mixing rate $\alpha > 0$ and step number $k > 0$, we apply the Massey inequality for the distribution $Q_{k+1}$ and find that as per Equation (7),

$$H[G(Q_k)] = H[G(Q_{k+1})] - \alpha Q_{k,n+k} \geq 2^{H(Q_{k+1}) - 2} + 1 - \alpha Q_{k,n+k} > 2^{H(Q_k) - 2} + 1.$$

Now, this refinement trickles down producing ever tighter bounds on the guessing entropy of lower index elements of the list up to the initial element $p = Q_0$,

$$H(G(p)) = H(Q_0) - p_n a \left(\frac{1 - \alpha^k}{1 - \alpha}\right) = H[G(Q_k)] + \sum_{j=0}^{k-1} \left(H[G(Q_j)] - H[G(Q_{j+1})]\right) \geq 2^{H(Q_k) - 2} + 1 + \sum_{j=0}^{k-1} \left(H[G(Q_j)] - H[G(Q_{j+1})]\right) > 2^{H(Q_k) - 2} + 1 + \sum_{j=0}^{k-2} \left(H[G(Q_j)] - H[G(Q_{j+1})]\right) = \cdots = 2^{H(Q_0) - 2} + 1 = 2^{H(p) - 2} + 1,$$
where the tightest of the enumerated bounds is:

\[
\mathbb{E} [G(p)] \geq 2^{H(Q_k) - 2} + 1 + \sum_{j=0}^{k-1} (\mathbb{E} [G(Q_j)] - \mathbb{E} [G(Q_{j+1})]) \\
= 2^{H(p) + h(\alpha) \frac{k}{1-\alpha} - 2} + 1 - p_n \frac{\alpha}{1-\alpha},
\]

which as we have shown increases with \( k \) up to the limit:

\[
\mathbb{E} [G(p)] \geq 2^{H(p) + 2p_n - 2} + 1 - p_n > 2^{H(p) - 2} + 1.
\]

**Proof.** The first inequality follows taking the supremum over \( \alpha \) in Equation (8) and the rest taking \( \alpha = 1/2 \). \( \square \)

Furthermore, note that Theorem 2 is a refinement over the strongest result of [5], namely Equation (1).

### 2.3. The Massey Gap

In this section, we reevaluate the scaffolding we used to prove Theorem 2 and re-frame our previous technique as an optimization process applied to an objective function we name the Massey gap, the function \( \mathcal{M}(x) \) described below.

The first observation that we make is that Theorem 2 also applies to descending probabilistic infinite sequences \( p \), by taking \( n \) to be the size of its support, \(|p|\). Moreover, we notice that the core of the argument was not the precise construction of the list \( \{Q_k\} \), but rather the fact that:

\[
2^{H(Q_k) - 2} + 1 + \mathbb{E} [G(Q_{k-1})] - \mathbb{E} [G(Q_k)] > 2^{H(Q_{k-1}) - 2} + 1 \quad \forall k \geq 1,
\]

i.e., the sequences \( \{Q_k\} \) incrementally optimize the objective function \( \mathcal{M}(x) = \mathbb{E} [G(x)] - 2^{H(x) - 2} - 1 \), with tightness being achieved if and only if \( \lim_{k \to \infty} \mathcal{M}(Q_k) = 0 \).

Motivated by this insight, we now take an alternative view and focus not on the precise sequence of distributions, but rather the objective function \( \mathcal{M}(x) \), which we name the Massey gap.

This view leads us to the following useful result:

**Remark 1.** Let \( \{X_k\} \) be a list of probabilistic sequences with \( H(X_k) \geq 2 \) and \( \mathcal{M}(X_k) \geq \mathcal{M}(X_{k+1}) \geq 0 \) for all values of \( k \). Then, the sequence \( y_k = 2^{H(X_k) - 2} + 1 + \mathbb{E} [G(X_0)] - \mathbb{E} [G(X_k)] \) is increasing.

Notice that this remark provides a sequence \( \{y_k\} \) of ever tighter lower bounds on \( \mathbb{E} [G(X_0)] \), the weakest of which, \( y_0 \), reduces to the Massey inequality, while the strongest is given by \( \lim_{k \to \infty} y_k \).

As an example, considering the list \( X_k := Q_k \), this limit coincides with the result of Theorem 2.
2.4. Mixing Probabilities

In the previous section, we presented a refinement to (1) by generalizing the tail rebalancing method of [5]. In this section, we investigate whether we can obtain even better results by applying the same rebalancing technique in the middle of the distribution, rather than at the tail.

In the previous section, we showed how Theorem 2 follows from Remark 1 for some particular \( \{X_k\} \). However, this option is not necessarily unique. In this section, we consider alternative methods of optimizing the Massey gap, that is we fix an arbitrary intermediary configuration \( X_k \), and we find conditions on the existence of available \( X_{k+1} \) that decrease the Massey gap, i.e., \( M(X_k) \geq M(X_{k+1}) \).

In order to simplify the notation, for the remainder of this section, we denote the available configuration \( X_k \) by \( \mathbf{p} \) and the potential next configuration \( X_{k+1} \) by \( \mathbf{q} \).

Now, we are ready for a trick similar to the one in Section 2.2: we fix an arbitrary integer \( i \in \mathbb{N} \) such that \( p_i \neq 0 \) and an arbitrary mixing rate \( \alpha \in (0, 1/2) \) and consider \( \mathbf{q} \) of the form:

\[
\mathbf{q} = (p_0, p_1, \ldots, p_{i-1}, (1-\alpha) p_i + \alpha p_{i+1}, \alpha p_i + (1-\alpha) p_{i+2}, p_{i+3}, \ldots),
\]

with the same support as \( \mathbf{p} \).

For \( \mathbf{q} \) as defined above in Equation (9), we find that \( \mathbb{E}[G(\mathbf{q})] = \mathbb{E}[G(\mathbf{p})] + \alpha (p_i - p_{i+1}) \) while:

\[
H(\mathbf{q}) = H(\mathbf{p}) + (p_i + p_{i+1}) \left( h \left( \frac{q_i}{q_i + q_{i+1}} \right) + \frac{1}{p_i + p_{i+1}} \log \left( 1 + \frac{\alpha (p_i - p_{i+1})}{2^{H(\mathbf{p})}} \right) \right),
\]

thus, the condition \( M(\mathbf{p}) \geq M(\mathbf{q}) \) is equivalent to:

\[
h \left( \frac{q_i}{q_i + q_{i+1}} \right) \geq \frac{1}{p_i + p_{i+1}} \log \left( 1 + \frac{\alpha (p_i - p_{i+1})}{2^{H(\mathbf{p})}} \right),
\]

which, as we have seen in the previous section, always holds at least for some \( i \), namely \( i = |\mathbf{p}| \).

Now, we check whether alternative values for \( i \) may be eligible, i.e., whether we may mix probabilities in the middle of the distribution. Using the inequality ( Tighter bounds may also be used.)

\[
1 - h \left( \frac{y}{2} \right) \leq \frac{y^2}{2^{h(y)}}
\]

with \( y = \frac{q_i - q_{i+1}}{q_i + q_{i+1}} = (1 - 2\alpha) \frac{p_i - p_{i+1}}{p_i + p_{i+1}} \), we find that:

\[
h \left( \frac{q_i}{q_i + q_{i+1}} \right) \geq 1 - (1 - 2\alpha)^2 \left( \frac{p_i - p_{i+1}}{p_i + p_{i+1}} \right)^2
\]

Thus, to satisfy \( M(\mathbf{p}) \geq M(\mathbf{q}) \), corresponding to a decrease in the Massey gap, it is sufficient that:

\[
1 - (1 - 2\alpha)^2 \left( \frac{p_i - p_{i+1}}{p_i + p_{i+1}} \right)^2 \geq h \left( \frac{p_i}{p_i + p_{i+1}} \right) + \frac{1}{p_i + p_{i+1}} \log \left( 1 + \frac{(p_i - p_{i+1})}{2^{H(\mathbf{p})}} \right)
\]

i.e.,

\[
\alpha \geq \frac{1}{2} - \frac{1}{2} \frac{p_i + p_{i+1}}{p_i - p_{i+1}} \left[ 1 - h \left( \frac{p_i}{p_i + p_{i+1}} \right) - \frac{1}{p_i + p_{i+1}} \log \left( 1 + \frac{(p_i - p_{i+1})}{2^{H(\mathbf{p})}} \right) \right]^{1/2} \geq 0,
\]

so depending on the value of \( H(\mathbf{p}) \), multiple indices \( i \) and mixing ratios \( \alpha \) may be selected for rebalancing, as outlined in the following

**Theorem 3.** Let \( \mathbf{p} \) be a descending probabilistic sequence with \( H(\mathbf{p}) \geq 2 \) and \( i \in \mathbb{N} \) be such that \( p_i \neq 0 \) and \( \alpha \in (0, 1/2) \) satisfy Equation (10) or, sufficiently, Equation (11). Then, defining \( \mathbf{q} \) as in Equation (9), we have:

\[
\mathbb{E}[G(\mathbf{p})] \geq 2^{H(\mathbf{q})-2} + 1 - \alpha (p_i - p_{i+1}) \geq 2^{H(\mathbf{p})-2} + 1.
\]
\textbf{Proof.} We have shown that Equation (11) is a sufficient condition for us to have $\mathcal{M}(p) \geq \mathcal{M}(q)$; thus, applying Remark 1 to the sequence $X = (p, q, q, \ldots)$, the result follows. \qed

We conclude this section by remarking that Theorem 3 has the potential to further refine bounds obtained using Theorem 1 depending on the form of the distribution $p$.

2.5. \textit{Strict Refinement and Optimal Mixing Rate}

In this section, we consider the problem of finding the optimal mixing rate $\alpha$.

We notice that $\alpha$ must be a critical point of the difference between the sides of (10), which is:

$$f(\alpha) = \frac{p_i - p_{i+1}}{p_i + p_{i+1}} \left[ \log \frac{q_i}{q_{i+1}} - \frac{1}{\alpha (p_i - p_{i+1}) + 2^H(p) - 2} \right].$$

For example, we notice that if $H(p) \geq 2$, then the right denominator is greater than one and:

$$f(1/2) = -\frac{p_i - p_{i+1}}{p_i + p_{i+1}} \frac{1}{\alpha (p_i - p_{i+1}) + 2^H(p) - 2} < 0$$

so the mixing rate $\alpha = 1/2$ is never optimal; thus, Theorem 2 is a strict refinement over the work of [5].

Investigating whether $q$ actually decreases the Massey gap, we now consider the optimality of $\alpha = 0$. In this case, $q_i = p_i$ and $q_{i+1} = p_{i+1}$; thus, when $p_i \geq p_{i+1} 2^{2-H(p)}$, then $f(0) > 0$, so $\alpha = 0$ is certainly sub-optimal as $\mathcal{M}(q) - \mathcal{M}(p)$ is increasing around zero.

Moreover, we note that the simple upper bound:

$$f(\alpha) \leq \frac{p_i - p_{i+1}}{p_i + p_{i+1}} \left[ \log \frac{q_i}{q_{i+1}} - \frac{1}{(p_i - p_{i+1}) + 2^H(p) - 2} \right]$$

imposes that the optimal mixing ratio $\alpha^*$ satisfy:

$$\frac{p_{i+1} + p_i}{q_{i+1}} - 1 = \frac{q_i}{q_{i+1}} \geq 2((p_i - p_{i+1}) + 2^H(p) - 2)^{-1}$$

i.e.,

$$q_{i+1} \leq \frac{p_i + p_{i+1}}{1 + 2((p_i - p_{i+1}) + 2^H(p) - 2)^{-1}}$$

and hence:

$$\alpha^* \leq \frac{1}{p_i - p_{i+1}} \left[ -p_{i+1} + \frac{p_i + p_{i+1}}{1 + 2((p_i - p_{i+1}) + 2^H(p) - 2)^{-1}} \right] \leq \frac{1}{2},$$

which is a non-trivial bound.

For example, to apply Equation (12) to the particular case of mixing only at the tail of the distribution, i.e., Theorem 2, we take $i = n$, hence $p_{i+1} = 0$, and substituting in Equation (12), we obtain the requirement:

$$\alpha^* \leq 1 / \left( 1 + 2^{1/(p_n + 2^{H(p) - 2})} \right) < 1/2$$

on the mixing rate realizing the supremum, so we have a closed-form strict refinement over Equation (2):

\textbf{Theorem 4.} Let $p$ be a descending probabilistic sequence with $H(p) \geq 2$ and $i \in \mathbb{N}$ be such that $p_i \geq 2^{-H(p)} p_{i+1}$, and let $\alpha = 1 / \left( 1 + 2^{1/(p_n + 2^{H(p) - 2})} \right)$. Then, defining $q$ as in Equation (9), we have:

$$\mathbb{E}[G(p)] \geq 2^{H(q) - 2} + 1 - \alpha (p_i - p_{i+1}) > 2^{H(p)} p_i p_{i+1}^{-2} + 1 - \frac{1}{2} p_{i+1} \geq 2^{H(p) - 2} + 1.$$
We end this section by remarking that Theorem 4 can be used in conjunction with Theorem 2, complementing the construction of the list of sequences defined in the proof of Theorem 2.

3. Discussion

In this section, we discuss in a more detailed manner the context of our work, evaluate the saturation conditions of the discussed bounds, and present prospects for future work.

We begin our discussion by taking a closer look at the most prevalently used form of the Massey inequality, namely Equation (3). Notably, this inequality only holds for the case that $H(p) \geq 2$, suggesting that its nice shape is a case of form vs. function, as illustrated below.

Massey [1] found that the set of probabilistic vectors of arbitrary length, which have a fixed guessing entropy $A$, are convex for any given $A$; thus, by strict convexity, the Shannon entropy attains its maximum over this set at an interior point. Using Lagrange multipliers, Massey found that the critical $p$ must be a geometric distribution, and imposing the condition $E[G(p)] = A$, he found

$$p_i = \frac{1}{A-1} \left(1 - \frac{1}{A}\right)^i,$$

hence:

$$H(p) = \log(A - 1) + \log \left(\left(1 - \frac{1}{A}\right)^{-A}\right),$$

(13)

Because this distribution maximizes the Shannon entropy, for arbitrary distributions $p$, we have:

$$H(p) \leq \log(E[G(p)]) - 1 + \log \left(1 - \frac{1}{E[G(p)]}\right)^{-E[G(p)]},$$

(14)

an unwieldy form that is saturated only for geometric distributions. Thus, the bound in Equation (14) is only ever saturated when $|p| = \infty$; therefore, the Massey inequality is never tight for distributions encountered in most applications such as cryptography.

To find a simpler form of Equation (14), Massey noticed that when $A \geq 2$, the second term of Equation (13) is at most two, with equality if and only if $A = 2$. Thus, when $H(p) \geq 2$, $E[G(p)] \geq 2^{H(p)-2} + 1$. However, the simplicity of this form impacts its tightness, as we can remark:

**Remark 2.** While the well-known form of Massey’s lower bound Equation (3) is only saturated when $p_i = 1/2^i$, his original inequality Equation (14) is saturated by all geometric distributions, i.e., for all values of $E[G(p)]$.

From this remark, we conclude that any inequality derived from an application of the Massey inequality is saturated if and only if the distribution for which the Massey inequality is ultimately applied is the geometric distribution of ratio 1/2. It follows that for the initial result of [5], Equation (2), as well as for our own initial result from Theorem 1, the bounds are saturated only for the geometric distribution of the ratio 1/2, even if they are tighter than the Massey inequality.

However, when those inequalities are applied after an unbounded number of steps, it is possible to obtain bounds saturated by more inputs. Indeed, for the stronger result of [5], as well as our own stronger result from Theorem 2, the bounds are saturated for any truncation of the geometric distribution of the ratio 1/2, i.e., when the original distribution has the form $p = (1/2, 1/4, \ldots, 1/2^{n-2}, 1/2^{n-1}, 1/2^{n-1})$ for some $n \geq 2$, i.e., for a countable range of $E[G(p)]$.

It is then natural to ask how the tightness improvement in Theorems 1 and 2 is reflected in saturation conditions. To this end, we notice that the limiting factor is the requirement that the limit of the list of probabilistic distributions is required to be the geometric distribution of the ratio 1/2; hence, for saturation, we necessarily have $a = 1/2$, i.e., this is a direct consequence of using the nice form of the Massey inequality, Equation (3). As such, we propose the open question of finding a similar technique exploiting the Massey gap with respect to the original form of the Massey inequality, Equation (14).
We continue our discussion by comparing the many refinements to the Massey inequality that we have presented on a simple numerical example. Let \( \mathbf{p} = (p_1, \ldots, p_{10}) \) be the descending probabilistic vector containing the first nine terms of the probabilistic geometric sequence of the ratio 0.8,

\[
\mathbf{p} = \left( \frac{1}{5}, 4, 25, 1953125, 16, 64, 256, 3125, 1024, 4096 \right)^	op,
\]

with Shannon entropy \( H(\mathbf{p}) \approx 3.32516 > \log \frac{10}{1-0.8^2} > 2 \) and guessing entropy \( E[\mathbf{G}(\mathbf{p})] \approx 4.02939 \). The state-of-the-art provides various lower bounds on the guessing entropy: 3.18126 from the Massey inequality, 3.21581 from the weaker result Equation (2) of [5], 3.25157 from their stronger result Equation (1), and 3.20977 from Rioul’s inequality [6,7]. Our new lower bounds for this example are: 3.21767 as the supremum in Theorem 1 taking \( \alpha \approx 0.38977 \) (improving Equation (2)), 3.25157 as the supremum in Theorem 2 taking \( \alpha = 0.5 \) (matching Equation (1)), and 3.21751 as the explicit value in Theorem 4 taking \( i = 10 \) (improving Equation (2)). In this particular case, Theorem 3 cannot be applied for \( i \neq 10 \) as \( p_i/p_{i+1} \leq 1.2 \), while \( 2^{2^{-H(\mathbf{p})}} \approx 1.3189 \). We notice that for this example, Theorem 2 matches the stronger result of [5].

The final topic of our discussion addresses a direction that was previously left unaddressed in the literature, rebalancing the initial distribution. We take an example and use it to showcase the effectiveness of the approach in Theorem 3. Considering the initial distribution \( \mathbf{p} = (5/8, 1/8, 1/8, 1/2^{16}, 1/2^{16}, \ldots, 1/2^{16}) \), we notice that \( H(\mathbf{p}) \geq 1/8 \log 2^{16} = 2 \), and moreover, \( p_1/p_2 = 5 \geq 2^{2^{-H(\mathbf{p})}} \). Hence, according to Section 2.4, we may rebalance our distribution, even at its head. Explicitly taking \( i = 1 \) and \( \alpha = 1/5 < 1/ (1 + 2^{1/2^{16}; \cdot; 2^{16}(\mathbf{p})^{-2}}) \), we construct \( \mathbf{q} = (21/40, 19/40, 1/8, 1/2^{16}, 1/2^{16}, \ldots, 1/2^{16}) \) such that \( \mathcal{M}(\mathbf{p}) \geq \mathcal{M}(\mathbf{q}) \geq 0 \). From Remark 1, it follows that each such rebalancing can be prepended to the list construction in Section 2.2.

However, continuing our example, we note that \( q_1/q_2 = 2^{13} \), so taking \( i = 3 \) and \( \alpha = 1/5 \), we may once again apply this procedure to obtain \( \mathbf{r} = (21/40, 19/40, (1 + 2^{-15})/10, 1/40 + 1/2^{14}, 1/2^{16}, \ldots, 1/2^{16}) \) with \( \mathcal{M}(\mathbf{p}) \geq \mathcal{M}(\mathbf{q}) \geq \mathcal{M}(\mathbf{r}) \geq 0 \). Noting that now \( r_2/r_3 \geq 19/5 = 4.25 \) and \( r_4/r_5 \geq 2^{16}/40 = 1638.4 \), the procedure can continue, gradually trickling down mass from the head up to the tail.

We emphasize that the indices eligible for rebalancing change at every step and the indices that were once ineligible (such as 2 and 4) can become eligible at future steps. In fact, if we rebalance at Index 4, we notice that the fourth position in the newly created vector would be smaller than \( (1 + 2^{-8})/50 \), so Index 3 would once again become eligible. It follows that some indices can become eligible for rebalancing more than once.

In conclusion, in this paper, we strictly refine the work of [5] presenting a generalized approach in the form of Theorem 2, as well as a data-dependent numerical procedure resulting in tighter refinements. The rebalancing technique in Section 2.4 and Theorem 3 provides a significant improvement to the procedure in Theorem 2; thus, for future work, we propose a precise study of the convergence of this method on real data, as well as a follow-up work showing how the same method can be applied to other inequalities, such as Rioul’s inequality, Equation (4).

**Author Contributions:** Conceptualization, A.T. and P.G.P.; Investigation, A.T. and P.G.P.; Supervision, P.G.P.; Writing—original draft, A.T.; Writing—review & editing, A.T. and P.G.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** We thank Olivier Rioul who showed us Equation (4), a gem well hidden within the pages of [7].

**Conflicts of Interest:** The authors declare no conflict of interest.
References

1. Massey, J.L. Guessing and entropy. In Proceedings of the 1994 IEEE International Symposium on Information Theory, Trondheim, Norway, 27 June–1 July 1994; IEEE: Piscataway, NJ, USA, 1994; p. 204.

2. Murray, H.; Malone, D. Convergence of Password Guessing to Optimal Success Rates. *Entropy* **2020**, *22*, 378. [CrossRef] [PubMed]

3. Arikan, E. An inequality on guessing and its application to sequential decoding. *IEEE Trans. Inf. Theory* **1996**, *42*, 99–105. [CrossRef]

4. Weinberger, N.; Shayevitz, O. Guessing with a Bit of Help. *Entropy* **2020**, *22*, 39. [CrossRef] [PubMed]

5. Popescu, P.G.; Choudary, M.O. Refinement of Massey Inequality. In Proceedings of the 2019 IEEE International Symposium on Information Theory (ISIT), Paris, France, 7–12 July 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 495–496.

6. Rioul, O. On Guessing. Unpublished Note, 2013.

7. de Chérisey, E.; Guilley, S.; Rioul, O.; Piantanida, P. Best Information is Most Successful. *Iacr Trans. Cryptogr. Hardw. Embed. Syst.* **2019**, *49–79*. [CrossRef]

8. Tăpuș, N.; Popescu, P. A new entropy upper bound. *Appl. Math. Lett.* **2012**, *25*, 1887–1890. [CrossRef]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).