On the Additive Sensitivity of LZ77 Under Consecutive Edits

Jeremiah Blocki[1] and Seunghoon Lee[2]

[1] Purdue University, West Lafayette, USA
[2] University of Waterloo, Waterloo, Canada
jblocki@purdue.edu, seunghoon.lee@uwaterloo.ca

Abstract. We revisit the problem of mitigating information leakage in the widely used but insecure compress-then-encrypt paradigm. While encryption hides message contents, the ciphertext length is directly related to the length of the compressed message, which may, in turn, leak information about the *content* of the message itself. Recent work of Blocki et al. (TCC 2025) proposed an (ε, δ)-differentially private approach that adds randomized padding calibrated to the global sensitivity of the compression algorithm, and showed that LZ77 has global sensitivity $O(W^{2/3} \log n)$ for input length n and sliding window size W.

However, prior analysis focused only on sensitivity with respect to single-character edits, which leads to limited privacy guarantees when protecting longer substrings such as passwords, passphrases, cookies, or confidential user records. A natural attempt to handle longer secrets is to appeal to group privacy, but for approximate differential privacy, this leads to very poor parameter degradation: in particular, the effective value of δ can grow exponentially with the group size g. In this work, we introduce and study the sensitivity of compression schemes under block edits. Specifically, we define two strings to be g-neighbors if they differ only within a contiguous interval of length g.

Our main technical contribution is a nearly tight characterization of the g-consecutive sensitivity of LZ77. We show that the g-consecutive sensitivity of LZ77, both with and without self-referencing, is at most $O\left((W^{2/3} + g + \sqrt{Wg}) \log n\right)$. In particular, when $g \leq W^{1/3}$, the bound simplifies to $O(W^{2/3} \log n)$, matching the known bound for single-character edits. Combined with the framework of Blocki et al., our bound yields (ε, δ)-differential privacy for g-neighbors with the same asymptotic padding scale as for single-character edits. We provide matching lower bounds to demonstrate that our upper bound is tight, e.g., when $n = W = \Theta(g^2)$, the g-consecutive sensitivity of LZ77 is at least $\tilde{\Omega}(g^{3/2})$, matching the $\sqrt{Wg} = \Theta(g^{3/2})$ term from our upper bound up to a logarithmic factor.

1 Introduction

Data compression is a fundamental tool for reducing the size of data prior to storage or transmission by exploiting redundancy in the input. Practical lossless compression schemes are ubiquitous, including formats such as ZIP (Deflate [Deu96]) and modern web compression standards like Brotli [AS16]. A common approach in practice is to combine compression with encryption: given a message w, one first applies a compression algorithm to obtain a shorter representation $y = \texttt{Compress}(w)$, and then encrypts y to produce a ciphertext $c = \mathsf{Enc}_K(y)$. This approach aims to simultaneously reduce communication costs and protect the confidentiality of the data. Unfortunately, this compress-then-encrypt framework is insecure as the ciphertext can leak information about the length of the underlying plaintext (e.g., an AES-GCM [BM08] ciphertext reveals the *exact* length of the corresponding plaintext) and the length of a compressed message y can itself depend on the *content* of w.

One of the concrete motivating examples given by Blocki et al. [BLG25] was to consider the message $w =$ "`The top secret password is: password.`". An attacker who observes $|\texttt{Compress}(w)|$ can test guesses of the form $w' =$ "`The top secret password is: [guess].`" and eliminate any guess for which $|\texttt{Compress}(w')| \neq |\texttt{Compress}(w)|$. Several real-world exploits

such as CRIME (Compression Ratio Info-leak Made Easy) [RD12] and BREACH (Browser Reconnaissance and Exfiltration via Adaptive Compression of Hypertext) [GHP13] similarly utilize this leakage. Despite its clear security issues, the "Compress-then-Encrypt" framework continues to be used as a tool to reduce bandwidth overhead. Defenses include static analysis and program transformation (DeBreacher [PSPW20]), randomized padding of response size (Heal the Breach [PFPSÚGDZ22]), and formal length-hiding mechanisms (Degabriele's Gaussian-padding framework [Deg21]).

Recent work of Blocki et al. [BLG25] proposed a Differentially Private Compress-then-Encrypt mechanism where one randomly pads the compressed file y before encryption to provide the mathematically rigorous guarantee of (ϵ, δ)-differential privacy. In their framework, the noise (randomized padding) is calibrated to the global sensitivity of the compression algorithm. In particular, let $f(w) = |\texttt{Compress}(w)|$ denote the length of the compressed file $y = \texttt{Compress}(w)$. Then the global sensitivity of the compression algorithm `Compress` is defined as

$$\max_{w \sim w'} \left| f(w) - f(w') \right| ,$$

where the maximum is taken over all pairs of neighboring strings w and w' which differ on a single character (i.e., $w[i] = w'[i]$ for all $i \neq j$ and $w[j] \neq w'[j]$, which we denote $w \sim w'$). Intuitively, the larger the global sensitivity, the more noise must be added; thus, the framework is only useful when the global sensitivity of the compression algorithm is $o(n)$, since otherwise the padding length would exceed the size of the original file and eliminate the benefits of compression. Intuitively, differentially private padding prevents an attacker from distinguishing between neighboring messages w and w' which will help to thwart the password attacker described above.

Blocki et al. [BLG25] upper bounded the global sensitivity of the popular LZ77 compression algorithm as $O(W^{2/3} \log n)$ where n is the length of the input string w and W is the sliding window size. In particular, if we don't constrain the LZ77 window, then $W = n$ and the sensitivity is $O(n^{2/3} \log n)$. Blocki et al. [BLG25] provided a nearly matching lower bound showing that the upper bound on the global sensitivity is essentially tight up to polylogarithmic factors.

However, prior analyses measure sensitivity with respect to *single-character edits* [AFI23, BLG25]. While this aligns with the standard notion of differential privacy, it does not accurately capture realistic leakage scenarios. In particular, reconsider the motivating example above, where an attacker leverages the length of the compressed file to eliminate incorrect password guesses. A careful reader will observe that the two strings w and w' in this example are not neighboring in the usual sense: they differ in an entire substring (the password), rather than at a single character. Thus, we cannot directly invoke an (ϵ, δ)-DP guarantee to argue that an attacker cannot distinguish between w and w' based on the (noisy) compression length. More broadly, in many applications sensitive information appears as a *contiguous block* of symbols. Examples include passwords and passphrases, authentication tokens and cookies, and database records where entire rows (corresponding to individual users) must be protected. In such settings, it is natural to consider changes that affect a block of length $g \gg 1$, rather than a single symbol.

One possible approach is to extend single-character guarantees via group privacy. However, for (ε, δ)-differential privacy this approach often leads to poor parameter degradation: group privacy guarantees scale roughly as $(g\varepsilon,\ ge^{g\varepsilon}\delta)$ [Vad17]. Approximate DP is necessary for any mechanism that adds input-independent randomized padding, including that of Blocki et al. [BLG25]. Since padding must be nonnegative, neighboring inputs with different compressed lengths induce different supports for the padded output length, ruling out pure DP. Thus, the δ term cannot simply

be eliminated to avoid its degradation under group privacy. Similarly, a naive extension of global sensitivity bounds would scale linearly with g, yielding upper bounds of $O(g \cdot W^{2/3} \log n)$ for LZ77 and thus $O(n \log n)$ when $g = O(n^{1/3})$ and $W = \Omega(n)$. In this regime, the amount of padding required would entirely negate the benefits of compression. Even when g is small (e.g., $g = O(1)$), this naive upper bound still incurs a multiplicative increase in the required noise, inflating the padding length by a constant factor that can be significant in practice.

Our Contributions. To address these limitations, we introduce a refined notion of neighboring inputs tailored to compression leakage. We say that two strings w and w' are *g-neighbors* if they differ only within a contiguous interval of length g. This notion directly captures the setting where a short secret substring is modified while the rest of the message remains unchanged, and avoids the pessimistic degradation inherent in general group privacy arguments.

Our main technical contribution is a nearly tight characterization of the g-consecutive sensitivity of the LZ77 compression algorithm (see Section 2.1 for the formal definition of g-consecutive sensitivity). We show that the g-consecutive sensitivity of LZ77 is at most

$$O\left(W^{2/3} \log n \;+\; g \log n \;+\; \sqrt{Wg} \log n\right),$$

which grows significantly more slowly than the upper bound $gW^{2/3}$ one would get by applying naïve composition to the bound of [BLG25]. In particular, when $g \le W^{1/3}$, our bound simplifies to $O(W^{2/3} \log n)$, matching the known bound for single-character edits ($g = 1$). Combining our sensitivity bound with the `DPCompress` transformation of Blocki et al. [BLG25] yields (ε, δ)-differential privacy for g-consecutive neighbors without group-privacy degradation (see Remark 1). When $g \le W^{1/3}$, our bound simplifies to $O(W^{2/3} \log n)$, so protecting a contiguous substring of length g requires the same asymptotic padding scale as protecting a single character. Moreover, an instantiation of `DPCompress` calibrated to their single-character sensitivity bound already provides (ε', δ')-differential privacy for g-consecutive neighbors, with $\varepsilon' = \Theta(\varepsilon)$ and negligible δ' whenever δ is negligible, without modifying the mechanism or adding further padding. We complement our upper bound with lower bounds showing that the g and $\sqrt{Wg}$ terms are necessary. In particular, when $W = n = \Theta(g^2)$, the binary-alphabet g-consecutive sensitivity is at least $\Omega(g^{3/2}) = \Omega(\sqrt{Wg})$, matching the corresponding term in our upper bound up to the logarithmic factor — the lower bound of [BLG25] implies that a $\tilde{\Omega}(W^{2/3})$ term is already necessary even if $g = 1$.

2 Preliminaries and Definitions

2.1 g-Consecutive Sensitivity

For strings $w, w' \in \Sigma^n$ of length n, we say that w, w' are *g-neighboring* (denoted by $w \overset{g}{\sim} w'$) if there exists some $j \in [n-g+1]$ such that for all $i \in [n] \setminus [j, j+g-1]$ we have $w[i] = w'[i]$. Since our usual notion of the local/global sensitivity cannot properly address the edits of g consecutive characters, we extend the notion of local/global sensitivity to define local/global g-consecutive sensitivity as follows.

Definition 1 (Local/Global g-consecutive Sensitivity). *Let Σ, Σ' be sets of alphabets. The* local g-consecutive sensitivity *of the compression algorithm* $\texttt{Compress} : \Sigma^* \to (\Sigma')^*$ *at $w \in \Sigma^n$ is defined as*

$$\mathrm{LS}^{g}_{\texttt{Compress}}(w) := \max_{w' \in \Sigma^n : w \overset{g}{\sim} w'} \left| |\texttt{Compress}(w)| - |\texttt{Compress}(w')| \right|,$$

and the global g-consecutive sensitivity *of the compression algorithm for strings of length n is defined as* $\mathtt{GS}^{g}_{\mathtt{Compress}}(n) := \max_{w\in\Sigma^n} \mathtt{LS}^{g}_{\mathtt{Compress}}(w)$. *If it is clear from context, then one can omit the parameter n (length of a string) and simply write* $\mathtt{GS}^{g}_{\mathtt{Compress}}$ *to denote the global g-consecutive sensitivity.*

Unless otherwise specified, we simply write *g-consecutive sensitivity* to refer to the global g-consecutive sensitivity.

Remark 1. We can apply the global g-consecutive sensitivity directly to the DPCompress framework of Blocki et al. [BLG25]. In particular, if $\Delta_g := \mathtt{GS}^{g}_{\mathtt{Compress}}$ denotes the global g-consecutive sensitivity, then we can instantiate the $\mathtt{DPCompress}$ transformation with $Z \leftarrow \mathrm{Lap}(\Delta_g/\varepsilon)$ and $k \leftarrow \frac{\Delta_g}{\varepsilon}\ln\left(\frac{1}{2\delta}\right) + \Delta_g + 1$. The proof of [BLG25, Theorem 2] then gives, for all g-consecutive neighbors $w \overset{g}{\sim} w'$ and all $S \subseteq \mathbb{N}$,

$$\Pr[|\mathtt{DPCompress}(w)| \in S] \le e^{\varepsilon} \Pr[|\mathtt{DPCompress}(w')| \in S] + \delta.$$

Thus, calibrating padding to Δ_g gives (ε, δ)-DP directly, whereas group privacy yields only $\left(g\varepsilon, \delta \sum_{i=0}^{g-1} e^{i\varepsilon}\right)$-DP from a single-character guarantee [Vad17].

By post-processing, the same inequality holds for any fixed event defined by an adversary's output after observing the padded compressed length. For example, let B be the event that an attacker successfully guesses a password, PIN, SSN, token, or cookie after observing the padded compressed length. Let x contain a sensitive string s (e.g., a password, PIN, SSN, token, or cookie), and let x' replace s with an independent random string of the same length. If the replacement lies within a contiguous block of length g, then we have the following security guarantee $\Pr[B \mid x] \le e^{\varepsilon} \Pr[B \mid x'] + \delta$ where B is the event that the attacker correctly guesses the *original* secret s in either experiment. Intuitively, we expect that the baseline probability $\Pr[B \mid x']$ is small because the file x' does not even contain the sensitive string (password, PIN, SSN, cookie). In this case, the post-processing guarantee ensures that $\Pr[B \mid x]$ remains small, up to a multiplicative factor of e^{ε} and an additive term of δ.

Scope, Limitations, and Extensions. Our notion of contiguous replacements directly captures fixed-length sensitive fields such as cookies, tokens, SSNs, and PINs, as well as passwords padded to a fixed maximum length. For unpadded passwords, our guarantee protects contents conditional on length. Extending the neighboring relation to allow contiguous insertions and deletions would provide a more natural notion of privacy for variable-length secrets, protecting both contents and length. For example, a file containing a password of length at most g could be compared with a file in which the password is deleted entirely.

Establishing tight sensitivity bounds for such edits presents additional technical challenges. For unbounded-window LZ77, the suffixes can be realigned after the edited region, suggesting that our proof techniques may extend with additional care for phrases and source intervals crossing the edit boundary. For bounded windows, an insertion or deletion can also change the distance between a source occurrence and its target when they lie on opposite sides of the edit, potentially moving the source across the window threshold. Determining tight bounds for contiguous insertion and deletion edits remains an interesting direction for future work.

2.2 LZ77 Compression Scheme

We begin by providing a brief overview of the LZ77 compression algorithm $\mathtt{LZ77} : \Sigma^* \to (\Sigma')^*$. The algorithm takes as input a string $x \in \Sigma^n$ and outputs a sequence of blocks $(B_1^x, \ldots, B_{t_x}^x) \leftarrow \mathtt{LZ77}(x)$

where the block $B_i^x = [q_i^x, \ell_i^x, c_i^x] \in \mathbb{Z}_n \times \mathbb{Z}_n \times \Sigma$ encodes the next $\ell_i^x + 1$ characters of x. In particular, if the decoding algorithm has already recovered the substring $x[1,p]$ then the decoding algorithm can recover the substring $x[p+1, p+\ell_i^x+1] = x[q_i^x, q_i^x+\ell_i^x-1] \circ c_i$, i.e., for $r = 0$ to $\ell_i^x - 1$ copy $x[p+r+1] \leftarrow x[q_i^x+r]$ and then set $x[p+\ell_i^x+1] \leftarrow c_i^x$. For the base case, the compression algorithm will set $B_1^x = [0, 0, x[1]]$ to encode the first character of x. Supposing that blocks $B_1^x, \ldots, B_{i-1}^x$ encode the first $p = (i-1)+\ell_1^x+\ldots+\ell_{i-1}^x$ characters $x[1,p]$ the encoding algorithm will generate the ith block B_i^x by finding the *longest* prefix of $x[p+1,n]$ that is also a substring of $x[\max\{1, p-W+1\}, p]$ where $W \leq n$ is the length of the sliding window — many implementations of the LZ77 compression algorithm save space by only storing the most recent W characters of x during the encoding process. There is also a self-referencing variant of LZ77 where the block B_i^x is generated by finding the longest prefix of $x[p+1,n]$ that is also a substring of $x[\max\{1, p-W+1\}, n]$, i.e., we allow that $q_i^x + \ell_i^x > p$ since it will still be possible to decode the block. To distinguish between the two versions of LZ77, we use "non-overlapping" LZ77 to refer to the regular version, where we require that $q_i^x + \ell_i^x - 1 < s_i^x$, and we use "self-referencing LZ77" to refer to the version where this requirement is dropped.

Useful LZ77 Notation. For each block B_i^x, let s_i^x and f_i^x denote the starting and ending positions of the substring of x encoded by B_i^x, respectively; thus, $s_i^x := 1 + \sum_{r=1}^{i-1}(\ell_r^x + 1)$ and $f_i^x := \sum_{r=1}^{i}(\ell_r^x + 1)$. By definition, $x[s_i^x, f_i^x] = x[q_i^x, q_i^x + \ell_i^x - 1] \circ c_i^x$. Now fix strings $w, w' \in \Sigma^n$. We say that a block $B_k^{w'}$ *starts inside* a block B_i^w if $s_i^w \leq s_k^{w'} \leq f_i^w$. For $i \in [t_w]$, let $\mathcal{M}_i^{w,w'} := \{k \in [t_{w'}] : s_i^w \leq s_k^{w'} \leq f_i^w\}$ denote the set of block indices in $\texttt{LZ77}(w')$ whose blocks start inside B_i^w. We say that B_i^w is a *type-r block* if $|\mathcal{M}_i^{w,w'}| = r$. If we let $y_r^w := \left|\left\{i : |\mathcal{M}_i^{w,w'}| = r\right\}\right|$ denote the total number of type-r blocks, then we can observe that we have

$$t_w = \sum_r y_r^w \quad \text{and} \quad t_{w'} = \sum_{r \geq 1} r \cdot y_r^w \quad \text{and} \quad t_{w'} - t_w \leq \sum_{r \geq 2}(r-1)y_r^w \,. \tag{1}$$

Throughout the paper, for strings $w, w' \in \Sigma^n$ with $w \overset{g}{\sim} w'$, we write $J := [j, j+g-1]$ for an interval such that $w[i] = w'[i]$ for all $i \notin J$. That is, w and w' agree outside J, while they may differ arbitrarily on J.

3 Technical Overview

In the exposition below, we will first focus on the regular (non-overlapping) version of LZ77 without any restriction on the window size $W = n$. Afterwards, we will briefly discuss how the analysis generalizes to the settings where $W < n$ and/or we allow for a self-referencing version. We will show that the g-consecutive sensitivity for both versions of LZ77 (non-overlapping and self-referencing) is upper bounded by $O\left((W^{2/3} + g + \sqrt{Wg})(\log n + \log|\Sigma|)\right)$. Finally, we will construct pairs of strings $w \overset{g}{\sim} w' \in \Sigma^n$ to tightly lower bound the g-consecutive sensitivity.

3.1 Review of Prior Analysis [BLG25] for special case $g = 1$

We first review the analysis of Blocki et al. [BLG25] who bounded the g-consecutive sensitivity of LZ77 in the special case where $g = 1$. We will then highlight which parts of the analysis do not extend to the more general case $g \geq 1$.

First, we can observe that each LZ77 block can be encoded as a binary string of length $2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil$. Thus, when we analyze the global sensitivity of LZ77, it suffices to focus on the total number of blocks and upper bound $\max_{w \sim w'} |t_w - t_{w'}|$ where the maximum is taken over all pairs of neighboring strings $w, w' \in \Sigma^n$ with Hamming distance at most one.

Let $w \sim w'$ be given and without loss of generality, assume that $t_{w'} \geq t_w$. To upper bound $t_{w'} - t_w$, Blocki et al. [BLG25] relied on an observation of Akagi et al. [AFI23] who proved that if $w, w' \in \Sigma^n$ have Hamming distance 1, then $|\mathcal{M}_i^{w,w'}| \leq 2$ for all $i \in [t_w]$. In particular, each block B_i^w of w can have *at most* two w' blocks that start inside B_i^w, i.e., block B_i^w is either type 0, type 1, or type 2. Because there are no type $r > 2$ blocks, we have $y_r^w = 0$ for all $r > 2$. Thus, from Equation (1) we have $t_{w'} - t_w \leq y_2^w$. In particular, to upper bound $t_{w'} - t_w$, it suffices to upper bound y_2^w, the number of type-2 blocks.

To upper bound y_2^w, Blocki et al. [BLG25] rely on several observations about type-2 blocks. Since the Hamming distance between w and w' is 1, there is a unique index $j \in [n]$ such that $w[j] \neq w'[j]$. The first observation is that if B_i^w is a type-2 block, then $f_i^w > j$, i.e., no type-2 blocks can finish before j. The second observation is that if $B_i^w = [q_i^w, \ell_i^w, c_i^w]$ is a type-2 block and $s_i^w > j$ (i.e., the block starts after j) then we must have $j \in [q_i^w, q_i^w + \ell_i^w)$, i.e., the copied region must include $w[j]$. The third observation is that if $B_i^w = [q_i^w, \ell_i^w, c_i^w]$ and $B_k^w = [q_k^w, \ell_k^w, c_k^w]$ are both type-2 blocks then $(q_i^w, \ell_i^w) \neq (q_k^w, \ell_k^w)$. Combining the second and third observation, we can conclude that there are at most ℓ type-2 blocks B_i^w with length $\ell_i^w = \ell$ and $s_i^w > j$. It follows that there are at most $\sum_{\ell=1}^{n^{1/3}} \ell = O(n^{2/3})$ type-2 blocks with length at most $n^{1/3}$ and $s_i^w > j$. Clearly there are at most $\frac{n}{n^{1/3}} = n^{2/3}$ blocks B_i^w with length $\ell_i^w \geq n^{1/3}$ since $\sum_{i=1}^{t_w} \ell_i^w \leq n$. Thus, it follows that there are at most $y_2^w = O(n^{2/3})$ type-2 blocks.

3.2 Challenges in Extending Prior Analysis [BLG25] for $g > 1$

We are interested in upper-bounding the g-consecutive sensitivity of LZ77. For strings $w, w' \in \Sigma^n$ of length n, we say that w, w' are *g-neighboring* (denoted by $w \overset{g}{\sim} w'$) if there exists some $j \in [n-g+1]$ such that for all $i \in [n] \setminus [j, j+g-1]$ we have $w[i] = w'[i]$. We will use $J = [j, j+g-1]$ to denote the interval where w and w' differ. As before, it suffices to upper bound $\max_{w \overset{g}{\sim} w'} |t_{w'} - t_w|$ since each LZ77 block can be encoded as a binary string of length $2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil$.

Without loss of generality, let $w \overset{g}{\sim} w'$ be given and assume that $t_{w'} > t_w$. Let j be given such that $w[i] = w'[i]$ for all $i \in [n] \setminus [j, j+g-1]$. Previously, when $g = 1$, we were able to argue that $t_{w'} - t_w \leq y_2^w$ because $y_r^w = 0$ for all $r > 2$ [AFI23, BLG25]. Unfortunately, because the Hamming distance between w and w' might be larger than 1, we can no longer conclude that $y_r^w = 0$.

However, several of the other observations from [BLG25] can still be generalized to our setting. First, if B_i^w is a type $r \geq 2$ block, then $f_i^w > j$, i.e., no type $r \geq 2$ blocks can finish before j — see Lemma 1. Second, if $B_i^w = [q_i^w, \ell_i^w, c_i^w]$ is a type $r \geq 2$ block and $s_i^w \geq j + g$ (i.e., the block starts after the differing region $w[j, j+g-1]$) then we must have $\left|[j, j+g-1] \cap [q_i^w, q_i^w + \ell_i^w)\right| \geq 1$, i.e., the region copied must *intersect* the region $w[j, j+g-1]$ where w and w' differ — see Claim 1. The third observation is that if $B_i^w = [q_i^w, \ell_i^w, c_i^w]$ is a type-$r_i \geq 2$ block and $B_k^w = [q_k^w, \ell_k^w, c_k^w]$ is a type-$r_k \geq 2$ block, then we must have $(q_i^w, \ell_i^w) \neq (q_k^w, \ell_k^w)$ — see Claim 6.

Combining the second and third observations, we can similarly conclude that there are at most $\ell + g$ total blocks B_i^w with $r_i \geq 2$, length $\ell_i^w = \ell$ and $s_i^w \geq j + g$ — see Claim 12. Thus, there are at most $\sum_{\ell=1}^{n^{1/3}} (\ell + g) = O(n^{2/3} + n^{1/3} g)$ total blocks B_i^w with type $r_i \geq 2$ and $s_i^w \geq j + g$ and, as

before, there can be at most $n^{2/3}$ total blocks B_i^w with length $\ell_i^w \geq n^{1/3}$. Thus, we can conclude that $\sum_{r\geq 2} y_r^w = O(n^{2/3} + n^{1/3}g)$. The extra $n^{1/3}g$ term arises because the upper bound on the number of type $r \geq 2$ blocks is now $\ell + g$ instead of ℓ. Unfortunately, the $n^{1/3}g$ term is already too loose for a tight analysis.

Even if we ignore the loose $n^{1/3}g$ term, which is already larger than we want, the previous observation that $\sum_{r\geq 2} y_r^w = O(n^{2/3} + n^{1/3}g)$ still does not yield a meaningful upper bound on $t_{w'} - t_w$ because we cannot assume that $y_r^w = 0$ for $r > 2$. In particular, to upper bound the g-consecutive sensitivity using Equation (1), we want to upper bound the sum $\sum_{r\geq 2}(r-1)y_r^w$ instead of $\sum_{r\geq 2} y_r^w$, and it is possible that $\sum_{r\geq 2}(r-1)y_r^w \gg \sum_{r\geq 2} y_r^w$. For example, a block B_i^w with length ℓ_i^w could have type $r_i = \Omega(\ell_i^w)$. Supposing that $g = O(n^{1/3})$ one could imagine a worst case where $y_r^w = \Omega(r)$ for all $r = O(n^{1/3})$, e.g., suppose that for all $\ell \leq n^{1/3}/10$ we have $\ell/10$ blocks of length ℓ and that each of these blocks B_i^w has type $r_i = \ell/10$. If such a worst case were possible, we would have $\sum_{r\geq 2}(r-1)y_r^w = \sum_{\ell \leq n^{1/3}/10}(\ell/10)(\ell/10 - 1) = \Omega(n)$. This falls far short of the upper bound $t_{w'} - t_w = O(n^{2/3} + g + \sqrt{ng})$ that we will prove, e.g., if $g = O(n^{1/3})$ then we will have $t_{w'} - t_w = O(n^{2/3})$ which tightly matches the upper bound of [BLG25] for the special case $g = 1$.

3.3 Potential Function

We introduce a potential function to help upper bound the summation $\sum_{r\geq 2}(r-1)y_r^w$. In particular, given our two strings $w \overset{g}{\sim} w'$ along with the interval $J = [j, j+g-1]$ such that $w[i] = w'[i]$ for all $i \notin [j, j+g-1]$, we define a potential function $\phi_J^w(t)$ with the following properties:

(1) **[Bounded]** $\phi_J^w(0) \leq 2\sqrt{ng}$.
(2) **[Non-Increasing]** $\phi_J^w(t) \leq \phi_J^w(t-1)$ for all $1 \leq t \leq t_w$.
(3) **[Positive]** $\phi_J^w(t_w) \geq 0$.
(4) **[Type-$r > 2$ Blocks Decrease Potential]** If block B_i^w has type r_i, and $s_i^w \geq j + g$ then $\phi_J^w(i) \leq \phi_J^w(i-1) - (r_i - 2)$.
(5) **[Internal-Copy Blocks of Type $r \geq 2$ Decrease Potential]** We call a block $B_i^w = [q_i^w, \ell_i^w, c_i^w]$ an *internal-copy* block if $[q_i^w, q_i^w + \ell_i^w] \subseteq J$. If B_i^w is an internal-copy block with type $r_i \geq 2$ and $s_i^w \geq j + g$, then $\phi_J^w(i) \leq \phi_J^w(i-1) - (r_i - 1)$.

Let $y_{r,\text{after}}^w$ denote the total number of blocks B_i^w with type $r_i = r$ and $s_i^w \geq j + g$. Similarly, we use $y_{r,\text{after}}^{w,\text{in}} \leq y_{r,\text{after}}^w$ to denote the total number *internal-copy blocks* B_i^w with type $r_i = r$ and $s_i^w \geq j + g$. Suppose that we can find a potential function that satisfies the above properties. The potential drop directly controls the blocks that start after J, i.e., $s_i^w \geq j + g$. To account for blocks that intersect J, for each block B_i^w, we define $\eta_i := |\{k : s_k^{w'} \in [s_i^w, f_i^w] \cap J\}|$. Since the target intervals $[s_i^w, f_i^w]$ partition $[n]$, we observe $\sum_i \eta_i \leq |\{k : s_k^{w'} \in J\}| \leq g$ — see Lemma 2. The only additional boundary case is the unique block crossing the right boundary of J; in Lemma 4, its after-suffix is charged to the same potential accounting. Taking these together, we have

$$\sum_{r\geq 2}(r-1)y_{r,\text{after}}^{w,\text{in}} + \sum_{r\geq 2}(r-2)\left(y_r^w - y_{r,\text{after}}^{w,\text{in}}\right) \leq \phi_J^w(0) + \sum_i \eta_i \leq 2\sqrt{ng} + g.$$

See Lemma 4 for the details.

As we observed previously (see Lemma 1), there are no type $r \geq 2$ blocks B_i^w with $f_i^w < j$. Now, it follows that

$$\sum_{r\geq 2}(r-2)y_r^w = \sum_{r\geq 2}(r-2)\left(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) + \sum_{r\geq 2}(r-2)y_{r,\mathsf{after}}^{w,\mathsf{in}} \leq g + 2\sqrt{ng} - \sum_{r\geq 2} y_{r,\mathsf{after}}^{w,\mathsf{in}}.$$

Thus, it follows that

$$\begin{aligned}\sum_{r\geq 2}(r-1)y_r^w &= \sum_{r\geq 2} y_r^w + \sum_{r\geq 2}(r-2)y_r^w \leq \sum_{r\geq 2}(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}) + g + 2\sqrt{ng} \\ &= \sum_{r\geq 2}(y_r^w - y_{r,\mathsf{after}}^w) + \sum_{r\geq 2}(y_{r,\mathsf{after}}^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}) + g + 2\sqrt{ng}.\end{aligned}$$

We already know, from Lemma 1 and Lemma 2, that $\sum_{r\geq 2}(y_r^w - y_{r,\mathsf{after}}^w) \leq g$. Hence, it remains to upper bound $\sum_{r\geq 2}(y_{r,\mathsf{after}}^w - y_{r,\mathsf{after}}^{w,\mathsf{in}})$. Previously, we had upper bounded $\sum_{r\geq 2} y_r^w$ directly as $\sum_{r\geq 2} y_r^w = O(n^{2/3} + n^{1/3}g)$, but this bound is not tight when $g \gg n^{1/3}$. Now we are able to tightly upper bound $\sum_{r\geq 2}(y_{r,\mathsf{after}}^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}) = O(n^{2/3})$ — see Claim 8. While there may be up to $\ell + g$ total length ℓ blocks of type $r \geq 2$ blocks, we are now allowed to exclude *internal-copy* blocks. Once we exclude these blocks, it is not too difficult to verify that there are *at most* 2ℓ remaining type $r \geq 2$ blocks of length exactly ℓ (see Claim 12 in Appendix C.1). This insight is sufficient to upper bound $\sum_{r\geq 2}(y_{r,\mathsf{after}}^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}) = O(n^{2/3})$ following similar analysis as [BLG25, Lemma 2].

It follows that the g-consecutive sensitivity of LZ77 can be upper bounded as $O((n^{2/3}+g+\sqrt{ng})\cdot(\log n + \log|\Sigma|))$ — see Theorem 2. If $g \leq n^{1/3}$ then the $n^{2/3}$ term dominates and the g-consecutive sensitivity of LZ77 is just $O(n^{2/3}(\log n + \log|\Sigma|))$ which matches the upper bound of [BLG25] for the special case $g = 1$.

It remains to find a potential function ϕ_J^w that satisfies all five properties specified above. Let $J = [j, j+g-1]$ be the interval where w and w' may differ, and consider a subinterval $I := [k, k'] \subseteq J$. We say that I is *w-covered by time t* if there is some block B_i^w starting after the interval J (i.e., with $s_i^w \geq j + g$) such that $q_i^w \leq k \leq k' < q_i^w + \ell_i^w$. Our potential function $\phi_J^w(t)$ simply counts the number of subintervals I of length $|I| \leq \sqrt{n/g}$ which are not covered by time t and then adds the distance $(n - f_t^w)$ to the end of the strings, normalized by the length parameter $K := \sqrt{n/g}$. More formally, we define

$$\phi_J^w(t) := |\{I \in \mathcal{I}_K(J) : \mathsf{Uncovered}(w, I, t) = 1\}| + \frac{n - f_t^w}{K},$$

where $\mathcal{I}_K(J) := \{I \subseteq J : |I| \leq K\}$ denotes the set of subintervals of J whose length is at most K, and the predicate $\mathsf{Uncovered}(w, I, t) = 1$ if and only if I is not w-covered by time t — see Definition 2 for a formal definition of the predicate $\mathsf{Uncovered}$.

It is easy to verify that the potential function is non-negative (i.e., $\phi_J^w(t) \geq 0$ since $f_t^w \leq n$ for all blocks B_t^w) and non-increasing since $f_{t+1}^w > f_t^w$ and $\mathsf{Uncovered}(w, I, t) = 0 \Rightarrow \mathsf{Uncovered}(w, I, t+1) = 0$, i.e., once an interval I is covered, it will remain covered. Finally, it is easy to see that $\phi_J^w(0) \leq 2\sqrt{ng}$ since $\frac{n-f_t^w}{K} \leq \frac{n}{K} = \sqrt{ng}$, and there are *at most* $gK = \sqrt{ng}$ subintervals of J with length at most K. The main challenge is to prove properties (4) and (5). In particular, we need to show that any type-$r_i > 2$ block B_i^w starting after J causes the potential to drop by at least $(r_i - 2)$,

i.e., $\phi_J^w(i) \le \phi_J^w(i-1) - (r_i - 2)$. Similarly, we need to show that for *internal-copy* blocks B_i^w of type $r_i \ge 2$ starting after J, the potential drop is at least $(r_i - 1)$, i.e., $\phi_J^w(i) \le \phi_J^w(i-1) - (r_i - 1)$. We prove this claim formally in Lemma 3. Below, we give a proof sketch.

Consider first an internal-copy block $B_i^w = [q_i^w, \ell_i^w, c_i^w]$ of type $r_i \ge 2$ starting after the interval J, i.e., with $s_i^w \ge j + g$. Let $B_k^{w'}, B_{k+1}^{w'}, \ldots, B_{k+r_i-1}^{w'}$ denote the r_i blocks of $\texttt{LZ77}(w')$ that start inside B_i^w. Now consider block $B_{k+r}^{w'}$ for any $0 \le r < r_i - 1$. We first observe that $w'[s_{k+r}^{w'}, f_{k+r}^{w'}] = w[s_{k+r}^{w'}, f_{k+r}^{w'}]$ since $s_{k+r}^{w'} \ge j + g$ and $w[v] = w'[v]$ for all $v \ge j + g$. Second, we observe that $w[s_{k+r}^{w'}, f_{k+r}^{w'}] = w[I]$ for some interval $I \subseteq J$ of length $\ell_{k+r}^{w'} + 1$ since B_i^w is an internal-copy block and therefore $s_i^w \le s_{k+r}^{w'} < f_{k+r}^{w'} < s_{k+r+1}^{w'} \le f_i^w$ (see Claim 2. This second observation may not hold when $r = r_i - 1$ because we may have $f_{k+r_i-1}^{w'} > f_i^w$ and $q_i^w + ((f_{k+r_i-1}^{w'} - 1) - s_i^w) \ge j + g$.). Third, we observe that the interval I above was previously uncovered (see Claim 4). Otherwise, we would have $w'[p, p + \ell_{k+r}^{w'}] = w[I]$ for some $p < s_{k+r}^{w'} - \ell_{k+r}^{w'}$ meaning that $\texttt{LZ77}(w')$ would have picked a longer block $B_{k+r}^{w'}$ to have length *at least* $|I| + 1$ by setting $q_{k+r}^{w'} = p$ and $\ell_{k+r}^{w'} = |I|$. This contradicts the choice of interval of length $|I| = \ell_{k+r}^{w'} + 1$. Thus, for each $0 \le r < r_i - 1$ the block $B_{k+r}^{w'}$ either has length $\ell_{k+r}^{w'} + 1 \ge K$ or it covers a previously uncovered interval I of length $|I| < K$. Either option causes the potential function to drop by *at least* 1. It follows that $\phi_J^w(i) \le \phi_J^w(i-1) - (r_i - 1)$.

The potential argument is very similar for the case when a type-$r_i \ge 3$ $B_i^w = [q_i^w, \ell_i^w, c_i^w]$ with $s_i^w \ge j + g$ is not an internal-copy block. Let $B_k^{w'}, B_{k+1}^{w'}, \ldots, B_{k+r_i-1}^{w'}$ denote the r_i blocks of $\texttt{LZ77}(w')$ that start inside B_i^w. For each $0 \le r \le r_i - 1$, we still have $w[s_{k+r}^{w'}, f_{k+r}^{w'}] = w[s_{k+r}^{w'}, f_{k+r}^{w'}]$ since $s_{k+r}^{w'} \ge j + g$ and $w[v] = w'[v]$ for all $v \ge j + g$. The second observation that $w[s_{k+r}^{w'}, f_{k+r}^{w'}] = w[I]$ for some interval $I \subseteq J$ *does not* necessarily hold when $r = 0$, because we may have $q_i^w + (s_k^{w'} - s_i^w) < j$. However, we can prove that $q_i^w + (s_k^{w'} - s_i^w) + \ell_k^{w'} \ge j$ which implies that $q_i^w + (s_{k+r}^{w'} - s_i^w) \ge j$ for each $r \ge 1$. Similarly, we can show that $q_i^w + (f_{k+r}^{w'} - s_i^w) < j + g$ for each $r < r_i - 1$. Thus, for all $1 \le r < r_i - 1$, we still have $w[s_{k+r}^{w'}, f_{k+r}^{w'}] = w[I]$ for some interval $I \subseteq J$. As before, the interval I must be previously uncovered meaning that the potential function must drop by *at least* 1 for each middle block $1 \le r < r_i - 1$ — either we cover a new interval of length $|I| \le K$ causing the potential to drop by at least one or the block $B_{k+r}^{w'}$ has length at least $|I| > K$ causing the potential to drop by at least $|I|/K \ge 1$. It follows that $\phi_J^w(i) \le \phi_J^w(i-1) - (r_i - 2)$ — see Lemma 3.

3.4 Bounded Sliding Window W

We can prove even tighter bounds when LZ77 is implemented with a sliding window of size $W \ll n$. In particular, the g-consecutive sensitivity of $\texttt{LZ77}_W$ is at most $O\left(\left(W^{2/3} + g + \sqrt{Wg}\right)\log n\right)$ — see Theorem 1. In the special case $g = 1$, Blocki et al. [BLG25] observed that any type-2 block B_i^w must have $s_i^w \le j + W$ where j is the unique index such that $w[j] \ne w'[j]$. A similar observation extends to our setting — see Claim 10. In particular, we must have $s_i^w < j + g + W$ for any type $r \ge 2$ block B_i^w, i.e., if block B_i^w starts after $s_i^w \ge j + g + W$ the block *must* have type $r_i = 0$ or type $r_i = 1$. Using this observation, we can obtain a tighter bound on the sum $\sum_{r \ge 2} y_r^w = O(W^{2/3} + g)$. To obtain a tighter upper bound for the sum $\sum_{r \ge 3}(r-2)y_r^w$, we only consider the window-reachable portion of J, defined as $J_W := \left[\max\{j, j + g - W\}, j + g - 1\right]$ and tweak the potential function as follows:

$$\phi_{J_W}^w(t) := \left|\left\{I \in \mathcal{I}_{K_W}(J_W) : \mathsf{Uncovered}(w, I, t) = 1\right\}\right| + \frac{\max\{0, H_W - f_t^w\}}{K_W},$$

where $K_W \coloneqq \max\{1, \sqrt{W/g}\}$, $H_W \coloneqq 2W + j + g - 1$, and recall that $\mathcal{I}_{K_W}(J_W) = \{I \subseteq J_W : |I| \le K_W\}$. Now we only count uncovered subintervals of J_W of length at most $\lfloor K_W \rfloor$ instead of $\lfloor \sqrt{n/g} \rfloor$ and we replace the old $\frac{n - f_t^w}{\sqrt{n/g}}$ term with $\frac{\max\{0, H_W - f_t^w\}}{K_W}$ because there will not be any type-$r \ge 2$ blocks which start after the index $j + g + W - 2$ and any block has length at most $W + 1$, implying that such a block will finish before the index $2W + j + g$. Thus, during this step, the second term decreases by exactly $\frac{f_i^w - f_{i-1}^w}{K_W}$. Taking $\max\{0, H_W - f_t^w\}$ ensures that the potential remains nonnegative, since the term $\frac{H_W - f_{t_w}^w}{K_W} = \frac{H_W - n}{K_W}$ could be negative at the end of the parsing if $H_W < n$. As before, we can argue that a type-$r_i \ge 3$ block B_i^w starting after J decreases potential by at least $(r_i - 2)$ and by at least $(r_i - 1)$ if B_i^w is an internal-copy block of type $r_i \ge 2$ starting after J. If i^* is the first index where the target interval of $B_{i^*}^w$ intersects J_W, then we can show that $\phi_{J_W}^x(i^* - 1) \le 5 \min\{W, \sqrt{Wg}\}$. Along with the same η_i arguments from the unbounded case, this allows us to upper bound $\sum_{r \ge 2}(r-1) y_{r,\mathsf{after}}^{w,\mathsf{in}} + \sum_{r \ge 2}(r-2)\left(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) \le O(g + \min\{W, \sqrt{Wg}\})$ and $\sum_{r \ge 2}\left(y_{r,\mathsf{after}}^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) \le O(W^{2/3})$ — see Lemma 8 and Lemma 10. Together with Claim 7, we obtain our primary result in Theorem 1 below.

Theorem 1. *Let* $\mathtt{LZ77}_W : \Sigma^* \to (\Sigma')^*$ *be the LZ77 compression function with sliding window size* $W \le n$. *Then for* $g \ge 1$,

$$\begin{aligned}\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) &\le \left(5 \min\{W, \sqrt{Wg}\} + (3W)^{2/3} + (3W)^{1/3} + 2g\right)\left(2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil\right) \\ &= O\left(\left(W^{2/3} + \sqrt{Wg} + g\right) \log n\right).\end{aligned}$$

Self-Referencing LZ77. The self-referencing variant of LZ77 (denoted by $\mathtt{LZ77}^{\mathtt{SR}}_W$) requires one additional consideration in the bounded-window case. Recall that in this variant, a block may copy from a source interval that overlaps the target interval currently being generated, and hence the copied length need not be bounded by the window size W. Nevertheless, the same internal-copy/non-internal-copy decomposition still applies. For an internal-copy block $B_i^w \in Y_{w,\mathsf{after}}$ of type $r \ge 2$, the potential-drop argument is unchanged: the translated source-side span is contained in J, so the periodic tail created by self-referencing cannot create an additional exception. Thus, $\phi_{J_W}^w(i-1) - \phi_{J_W}^w(i) \ge r - 1$. For a non-internal-copy block of type $r \ge 3$, the self-referencing behavior may create one additional exceptional interval: besides the first and last translated intervals, one intermediate translated interval may enter the periodic tail after J. Therefore, the guaranteed drop becomes $r - 3$, rather than $r - 2$; see Lemma 12. Telescoping the bounded-window potential gives $\sum_{r \ge 2}(r-1) y_{r,\mathsf{after}}^{w,\mathsf{in}} + \sum_{r \ge 3}(r-3)\left(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) \le g + 5 \min\{W, \sqrt{Wg}\}$ (see Lemma 13). The remaining contribution comes from counting the non-internal-copy type-≥ 2 blocks that start after J. Since self-referencing can make the last such block arbitrarily long, we remove the terminal non-internal-copy type-≥ 2 after-block before applying the bounded-window length budget. This gives $\sum_{r \ge 2}\left(y_{r,\mathsf{after}}^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) \le (3W)^{2/3} + (3W)^{1/3} + 1$ (see Lemma 14). Combining these bounds yields $\mathtt{GS}^g_{\mathtt{LZ77}^{\mathtt{SR}}_W}(n) \le O\left(\left(W^{2/3} + \sqrt{Wg} + g\right) \log n\right)$, as stated in Theorem 3.

3.5 Lower Bound

In the special case when $g = 1$ and $W = n$ Blocki et al. [BLG25] proved that the additive sensitivity of LZ77 was at least $\Omega(n^{2/3}\log^{1/3} n)$. Fujie et al. [FSNI25] claim a tighter $\Omega(n^{2/3}\log n)$ lower bound though their construction requires a very large alphabet Σ, while the alphabet size in [BLG25] is constant $|\Sigma| = O(1)$. These results demonstrate that any upper bound on the g-consecutive sensitivity will need to include a $W^{2/3}\log n$ term since increasing g can only increase the additive sensitivity. It is natural to wonder whether the remaining $g(\log n + \log|\Sigma|)$ and $\sqrt{Wg}(\log n + \log|\Sigma|)$ terms from our upper bounds are necessary.

To ease exposition, we first present lower bound examples which use a very large alphabet $|\Sigma| = \Theta(g^2)$. In particular, we will set $\Sigma = \{c_i\}_{i=1}^{2g+1} \cup \{b_i\}_{i=1}^{g} \cup \{d_{i,j}\}_{i,j\le g}$. We later explain how to transfer these examples to the binary alphabet using a random encoding.

We first argue that the $\Omega(g)$ term is necessary. Let $w = (c_1 \circ c_2 \circ \ldots c_g) \circ (c_1 \circ c_2 \circ \ldots c_g) \circ c_{g+1}$ and let $w' = (b_1 \circ b_2 \circ \ldots b_g) \circ (c_1 \circ c_2 \circ \ldots c_g) \circ c_{g+1}$. Clearly, we have $w \overset{g}{\sim} w'$. Setting $W = n = 2g+1$, it is easy to verify that LZ77 will use exactly $g+1$ blocks to compress w and exactly $n = 2g+1$ blocks to compress w', i.e., w' has g more blocks than w. Thus, the compression of w' will be $g(2\lceil\log n\rceil + \lceil\log|\Sigma|\rceil)$ bits longer than the compression of w — see Theorem 4.

We now illustrate the necessity of the $O(\sqrt{Wg}(\log n + \log|\Sigma|))$ term. Let $C := c_1 \circ \ldots \circ c_g$. We define $x_t = C \circ L_1 \circ L_2 \circ \ldots \circ L_t$ and $x'_t = b_1 \circ \ldots \circ b_g \circ L_1 \circ L_2 \circ \ldots \circ L_t$ where the strings L_r will be defined below. Clearly, $x_t \overset{g}{\sim} x'_t$ since the strings are identical after the first g characters. We define the string $L_r = C[1,r] \circ d_{r,1} \circ C[2,r+1] \circ d_{r,2} \circ \ldots \circ C[g-r+1,g] \circ d_{r,g-r+1}$ for each $r < g$. Note that substring L_r has length $(g-r+1)(r+1)$. The total length of the string x_t is $g + \sum_{r=1}^{t}(g-r+1)(r+1) = O(gt^2)$.

Intuitively, for each $r \le t$ and each $1 \le z \le g-r$, the segment $C[z, z+r-1] \circ d_{r,z}$ from L_r will correspond to a single LZ77 block when we compress x_t, but will require multiple LZ77 blocks when we compress x'_t. In particular, if we set $W = |x_t|$ then the total number of blocks for the compression of x_t is exactly $g + \sum_{i=1}^{t}(g-i)$. By contrast, when we compress x'_t we will use g blocks to encode the first g characters $c_1 \circ \ldots \circ c_g$, followed by $2g$ blocks to encode L_1, $2g-2$ blocks to encode L_2, etc. The total number of blocks for the compression of w' is exactly $g + 2\sum_{i=1}^{t}(g-i)$. The difference is exactly $\sum_{i=1}^{t}(g-i) = \Omega(gt)$ so the lengths of the corresponding LZ77 compressions will differ by $\Omega(gt(\log n + \log|\Sigma|))$ bits. Since, $W = \Theta(t^2 g)$ we have $gt = \Theta(\sqrt{gW})$. It follows that the g-consecutive sensitivity is at least $\Omega(\sqrt{gW}(\log n + \log|\Sigma|))$ — see Theorem 5.

As an example, if we set $t = \sqrt{g}$ then $|x_t| = W = O(g^2)$ and $gt = \Theta(\sqrt{Wg})$. Thus, the lengths of the corresponding LZ77 compressions will differ by $\Omega(\sqrt{Wg}(\log n + \log|\Sigma|))$ bits. Observe that, in this example, the term $\sqrt{Wg} = \Theta(g^{1.5})$ dominates the other terms g and $W^{2/3} = \Theta(g^{4/3})$ from our upper bound. Thus, the $O(\sqrt{Wg}(\log n + \log|\Sigma|))$ term is truly necessary.

Finally, Section 5.2 transfers these large-alphabet constructions to the binary alphabet via a random encoding. We show that the encoding introduces no unintended long bit-level matches and preserves the block-count gaps up to constant factors — see Lemma 6 and Lemma 7. This gives binary lower bounds $\Omega(g)$ and $\Omega(\sqrt{Wg})$ as stated in Theorem 6, so the g and $\sqrt{Wg}$ terms remain necessary over constant-size alphabets up to the logarithmic factor in the upper bound.

4 Upper Bound for the g-Consecutive Sensitivity of LZ77

Recall that each LZ77 block is encoded using at most $2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil$ bits. Hence, the g-consecutive sensitivity of LZ77 is bounded by $(2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil) \cdot \max_{w \overset{g}{\sim} w'} |t_w - t_{w'}|$, where t_x denotes the number of blocks in $\mathtt{LZ77}(x)$. Thus, it suffices to upper-bound the difference between the numbers of blocks in $\mathtt{LZ77}(w)$ and $\mathtt{LZ77}(w')$.

4.1 g-Consecutive Sensitivity of LZ77 with Unbounded Window

We first analyze the non-self-referencing version of LZ77 with an unbounded sliding window, i.e., $W = n$. This setting contains the main combinatorial ideas of the proof while avoiding two technical complications: the bounded-window restriction and the self-referencing variant. We later extend the analysis to bounded windows in Section 4.2 and to self-referencing LZ77 in Section 4.3.

Let $w \overset{g}{\sim} w' \in \Sigma^n$ with edited interval $J = [j, j+g-1]$, i.e., $w[i] = w'[i]$ for all $i \in [n] \setminus J$. Recall that y_r^w denotes the total number of type-r blocks of $\mathtt{LZ77}(w)$. Let $y^w_{r,\mathsf{before}}$ denote the total number of type-r blocks of $\mathtt{LZ77}(w)$ with $f_i^w < j$, $y^w_{r,\mathsf{intersect}}$ denote the total number of type-r blocks of $\mathtt{LZ77}(w)$ that intersect J, i.e., $[s_i^w, f_i^w] \cap J \neq \emptyset$, and $y^w_{r,\mathsf{after}}$ denote the total number of type-r blocks of $\mathtt{LZ77}(w)$ that starts after J, i.e., $s_i^w \geq j+g$. We first observe that there are no type-$r \geq 2$ blocks B_i^w with $f_i^w < j$.

Lemma 1. *For every $r \geq 2$, $y^w_{r,\mathsf{before}} = 0$.*

Proof. Since $w[1, j-1] = w'[1, j-1]$, the LZ77 parsings of w and w' are identical up to the first block that intersects J. Therefore, all the blocks of $\mathtt{LZ77}(w)$ with $f_i < j$ must be of type 1. This concludes the proof. □

We also observe that there can be at most g blocks of $x \in \{w, w'\}$ which start inside the interval J.

Lemma 2. *For $x \in \{w, w'\}$, $|\{k : s_k^x \in J\}| \leq g$ and $\sum_{r \geq 0} y^x_{r,\mathsf{intersect}} \leq g$.*

Proof. The starting positions s_k^x are strictly increasing. Since J has length g, at most g distinct starting positions can lie in J. Hence, $|\{k : s_k^x \in J\}| \leq g$. By definition, $\sum_{r \geq 0} y^x_{r,\mathsf{intersect}} = |\{k : [s_k^x, f_k^x] \cap J \neq \emptyset\}|$. Since the target intervals of distinct blocks are pairwise disjoint, and each block counted on the right-hand side contains at least one position of J, we have $\sum_{r \geq 0} y^x_{r,\mathsf{intersect}} \leq |J| = g$. □

Now for $x \in \{w, w'\}$, let $Y_{x,\mathsf{after}} := \{B_i^x : s_i^x \geq j+g\}$ denote the set of blocks of $\mathtt{LZ77}(x)$ that starts after J. We first show that if a block in $Y_{w,\mathsf{after}}$ is of type r for $r \geq 2$, i.e., more than one block of $\mathtt{LZ77}(w')$ starts inside the block B_i^w, then B_i^w must copy a substring that *intersects* J.

Claim 1. *If $B_i^w = [q_i^w, \ell_i^w, c_i^w] \in Y_{w,\mathsf{after}}$ is type-r block for $r \geq 2$, then $[q_i^w, q_i^w + \ell_i^w - 1] \cap J \neq \emptyset$.*

The proof of Claim 1 follows largely the same strategy as [BLG25, Claim 16] and therefore we defer the proof to Appendix C.1.

Definition 2 (Predicate $\mathsf{Uncovered}$). *Let $x \in \Sigma^n$ be a string and $(B_1^x, \ldots, B_{t_x}^x) \leftarrow \mathtt{LZ77}(x)$. We say that an interval $I := [k, k']$ is x-covered by time t if there exists some block $B_i^x \in Y_{x,\mathsf{after}}$ with $i \leq t$ such that $q_i^x \leq k \leq k' < q_i^x + \ell_i^x$.*

We define the predicate $\mathsf{Uncovered}(x, I, t)$ as follows:

$$\mathsf{Uncovered}(x, I, t) := \begin{cases} 1 \text{ if } I \text{ is not } x\text{-covered by time } t, \text{ and} \\ 0 \text{ otherwise.} \end{cases}$$

Definition 3 (Potential Function). *Let $K := \sqrt{n/g}$. Given an interval $J := [j, j+g-1]$, let $\mathcal{I}_K(J) := \{[k,k'] \subseteq J : |[k,k']| \le K\}$ denote the set of subintervals of J whose length is at most K. Now, we define a* potential function $\phi_J^x : \mathbb{Z}_{\ge 0} \to \mathbb{R}_{\ge 0}$ *as follows:*

$$\phi_J^x(t) := |\{I \in \mathcal{I}_K(J) : \mathsf{Uncovered}(x, I, t) = 1\}| + \frac{n - f_t^x}{K},$$

where $f_0^x := 0$.

The first term in $\phi_J^x(t)$ counts the number of subintervals of J of length at most K that are not x-covered by time t. The second term is the remaining unparsed length normalized by K; when the parsed prefix length increases from f_t^x to f_{t+1}^x, this term decreases by $\left(f_{t+1}^x - f_t^x\right)/K$.

At time $t = 0$, no interval in $\mathcal{I}_K(J)$ is x-covered, and $f_0^x = 0$. Thus, $\phi_J^x(0) = |\mathcal{I}_K(J)| + \frac{n}{K}$. Since J has length g, the number of subintervals of J of length ℓ is $g - \ell + 1$. Hence, for $L := \min\{g, \lfloor K \rfloor\}$, we have $|\mathcal{I}_K(J)| = \sum_{\ell=1}^{L}(g - \ell + 1) \le Kg$ and $\phi_J^x(0) \le Kg + \frac{n}{K} = 2\sqrt{ng}$.

Definition 4 (Internal-Copy Block). *We say that a block $B_i^x = [q_i^x, \ell_i^x, c_i^x] \in Y_{x,\mathsf{after}}$ is an* internal-copy *block if its source-side span satisfies $[q_i^x, q_i^x + \ell_i^x] \subseteq J$.*

Lemma 3. *Let $(B_1^w, \dots, B_{t_w}^w) \leftarrow \mathtt{LZ77}(w)$ and $(B_1^{w'}, \dots, B_{t_{w'}}^{w'}) \leftarrow \mathtt{LZ77}(w')$ for strings $w \overset{g}{\approx} w' \in \Sigma^n$ that agree outside J. Let $B_i^w \in Y_{w,\mathsf{after}}$ be a type-r block.*

(1) If $r \ge 3$, then $\phi_J^w(i-1) - \phi_J^w(i) \ge r - 2$.
(2) If B_i^w is an internal-copy block and $r \ge 2$, then $\phi_J^w(i-1) - \phi_J^w(i) \ge r - 1$.

Before proving Lemma 3, we explain the intuition behind the two bounds. Suppose that $B_i^w \in Y_{w,\mathsf{after}}$ is a type-r block, and let $B_{k_1}^{w'}, \dots, B_{k_r}^{w'}$ be the blocks of $\mathtt{LZ77}(w')$ that start inside B_i^w, ordered from left to right. After translating these blocks back to the source side of B_i^w, the translated intervals corresponding to the interior blocks $B_{k_2}^{w'}, \dots, B_{k_{r-1}}^{w'}$ lie inside J, while the first and last translated intervals may intersect the boundary of J. Each interior translated interval is newly w-covered at time i; otherwise, the corresponding block of $\mathtt{LZ77}(w')$ could be extended past the next block boundary. This gives the loss $r - 2$ in the general case. If B_i^w is an internal-copy block, then the source-side span $[q_i^w, q_i^w + \ell_i^w]$ is contained in J. Therefore, the translated intervals corresponding to $B_{k_1}^{w'}, \dots, B_{k_{r-1}}^{w'}$ are all contained in J; only the translated interval corresponding to $B_{k_r}^{w'}$ may fail to be contained in J. The same extension argument then gives $r - 1$ newly w-covered intervals.

If a newly covered interval I has length at most K, then it directly decreases the first term of the potential. If $|I| > K$, then we instead charge it to the second term $(n - f_t^w)/K$. The charged intervals are pairwise disjoint and lie inside the portion encoded by B_i^w. Hence, if there are h such long intervals, then $hK \le f_i^w - f_{i-1}^w$. The second term decreases by $(f_i^w - f_{i-1}^w)/K$, which is at least h. Thus, each newly covered interval contributes one unit of potential drop, either through the first term or through the second term.

Example 1. Figure 1 illustrates the mechanism behind the potential drop in the type-3 case. Suppose that two strings w, w' differ in an interval $J = [3, 7]$, highlighted in red. In Figure 1(a), the middle block $B_{k_2}^{w'}$ corresponds, after translation to the source side of B_i^w, to the interval $[4, 6] \subseteq J$. Hence, if $w[4, 6]$ is uncovered by time $i - 1$, then it becomes covered at time i, so the potential decreases

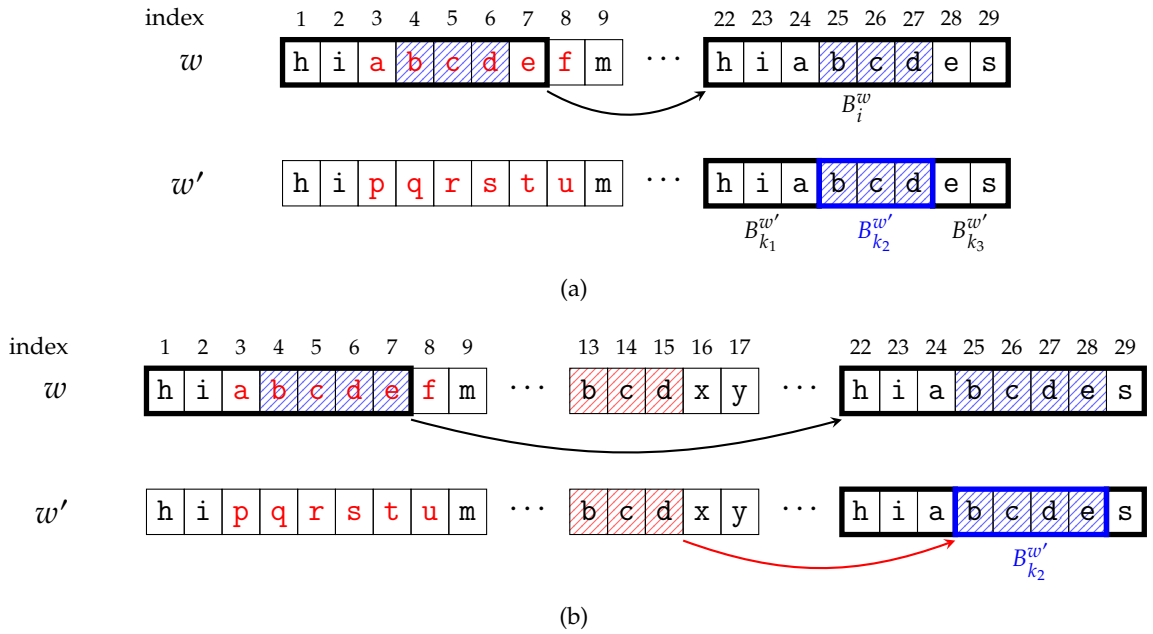

Figure 1: An example of the type-3 case in Lemma 3. See Example 1 for details.

by at least one. Figure 1(b) shows why $w[4,6]$ must be uncovered by time $i-1$. If it were already covered earlier, then the substring bcd would already occur in an earlier copied target, allowing $B_{k_2}^{w'}$ to extend to the longer substring bcde. Therefore, $w[4,6]$ is newly covered at time i.

Proof of Lemma 3. Since $B_i^w \in Y_{w,\mathsf{after}}$ is of type r, there exist distinct indices $k_1 < \ldots < k_r$ such that the blocks $B_{k_1}^{w'}, \ldots, B_{k_r}^{w'}$ start inside B_i^w. Now define $\delta_i := s_i^w - q_i^w$ and for each $m \in [r]$, let $I_m := \left[s_{k_m}^{w'} - \delta_i, f_{k_m}^{w'} - \delta_i\right]$ be the corresponding translated intervals on the source side of B_i^w.

For the first argument, assume $r \geq 3$. Let $J' := [q_i^w, q_i^w + \ell_i^w - 1] \cap J \subseteq J$. We observe that by Claim 1, $J' \neq \emptyset$. Now we establish the following claims.

Claim 2. *$I_m \subset J'$ for $m = 2, \ldots, r-1$.*

Proof of Claim 2. It suffices to prove the claim for I_2 and I_{r-1}, since the intervals $I_1, \ldots, I_r$ are pairwise disjoint and occur in this order along the source interval of B_i^w. The first translated interval I_1 may contain the portion of the source interval lying *before* J, i.e., $[q_i^w, q_i^w + \ell_i^w - 1] \setminus J$, and the last translated interval I_r may contain the portion lying *after* J. However, no intermediate translated interval can meet either side of $[q_i^w, q_i^w + \ell_i^w - 1] \setminus J$; otherwise, the corresponding block in $\mathtt{LZ77}(w')$ could be merged with its predecessor or successor, contradicting the maximality of the LZ77 parsing. Therefore, $I_m \subset J'$ for every $m = 2, \ldots, r-1$. ■

Claim 3. *For every $m = 2, \ldots, r-1$, $\mathsf{Uncovered}(w, I_m, i) = 0$, i.e., I_m is w-covered by time i.*

Proof of Claim 3. By Claim 2, $I_m \subset J' \subseteq [q_i^w, q_i^w + \ell_i^w - 1]$ for each $m = 2, \ldots, r-1$. Therefore, by definition, I_m is w-covered by time i. ■

Claim 4. *For every $m = 2, \ldots, r-1$, $\mathsf{Uncovered}(w, I_m, i-1) = 1$, i.e., I_m is not w-covered by time $i-1$.*

The proof of Claim 4 is a maximality argument: if I_m had already been w-covered by an earlier after-block, then the corresponding block $B^{w'}_{k_m}$ would have an earlier occurrence in w' and could be extended past its endpoint. We defer the details to Appendix C.1.

We will also need the following observation for the internal copy case with $r \geq 2$. Suppose that B^w_i is an internal-copy block. Then the same conclusions as in Claim 2, Claim 3, and Claim 4 also hold for $m = 1$. Indeed, since B^w_i is internal, we have $[q^w_i, q^w_i + \ell^w_i] \subseteq J$. Moreover, because $B^{w'}_{k_1}$ is not the last block starting inside B^w_i, we have $s^w_i \leq s^{w'}_{k_1} \leq f^{w'}_{k_1} < s^{w'}_{k_2} \leq f^w_i$. After translating by $-\delta_i$, we get $I_1 \subseteq [q^w_i, q^w_i + \ell^w_i - 1] \subseteq J$. Thus, I_1 is w-covered by time i. The same maximality argument as in Claim 4 shows that I_1 was not w-covered by time $i-1$; otherwise $B^{w'}_{k_1}$ could be extended past its endpoint, contradicting the maximality of the LZ77 block.

We now complete the proof of the two assertions separately. For the first assertion, assume $r \geq 3$ and define $M = \{2, \ldots, r-1\}$. For the second assertion, assume that B^w_i is an internal-copy block and $r \geq 2$, and define $M = \{1, \ldots, r-1\}$. By the claims above, for every $m \in M$, the interval I_m is not w-covered by time $i-1$ and becomes w-covered by time i. We split M into $M_{\leq K} := \{m \in M : |I_m| \leq K\}$ and $M_{>K} := \{m \in M : |I_m| > K\}$. For each $m \in M_{\leq K}$, the interval I_m is counted in the first term of $\phi^w_J(i-1)$ but not in the first term of $\phi^w_J(i)$. Since the intervals I_m are pairwise disjoint, the first term of the potential decreases by at least $|M_{\leq K}|$.

It remains to charge the intervals in $M_{>K}$. These intervals are pairwise disjoint; after translating them forward by δ_i, their corresponding target intervals lie inside the target interval $[s^w_i, f^w_i]$ encoded by B^w_i. Hence, $\sum_{m \in M_{>K}} |I_m| \leq f^w_i - f^w_{i-1}$. Since each interval in $M_{>K}$ has length greater than K, we have $\sum_{m \in M_{>K}} |I_m| \geq K|M_{>K}|$ and thus $|M_{>K}| \leq \frac{f^w_i - f^w_{i-1}}{K}$. On the other hand, the second term of the potential decreases by $\frac{n - f^w_{i-1}}{K} - \frac{n - f^w_i}{K} = \frac{f^w_i - f^w_{i-1}}{K} \geq |M_{>K}|$. Therefore,

$$\phi^w_J(i-1) - \phi^w_J(i) \geq |M_{\leq K}| + |M_{>K}| = |M|.$$

For the first assertion, $|M| = r - 2$, and for the second assertion, $|M| = r - 1$. This concludes the proof. □

Let $y^{w,\mathsf{in}}_r$ be the number of type-r internal copy blocks, and let $y^{w,\mathsf{in}}_{r,\mathsf{after}}$ denote the number of type-r internal-copy blocks in $Y_{w,\mathsf{after}}$. Then we observe the following lemma.

Lemma 4. *Let* $(B^w_1, \ldots, B^w_{t_w}) \leftarrow \mathtt{LZ77}(w)$ *and* $(B^{w'}_1, \ldots, B^{w'}_{t_{w'}}) \leftarrow \mathtt{LZ77}(w')$ *for strings* $w \overset{g}{\sim} w' \in \Sigma^n$ *that agree outside* J*. Then*

$$\sum_{r \geq 2} (r-1) y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r \geq 2} (r-2)\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) \leq g + 2\sqrt{ng}.$$

Proof. For each block B^w_i, let r_i be its type, and define

$$\eta_i := |\{k : s^{w'}_k \in [s^w_i, f^w_i] \cap J\}|.$$

Since the target intervals $[s^w_i, f^w_i]$ partition $[n]$, Lemma 2 gives

$$\sum_i \eta_i = |\{k : s^{w'}_k \in J\}| \leq g. \tag{2}$$

Now define the following sets:

– $A^{\text{in}}_{\text{after}} := \{i : B^w_i \in Y_{w,\text{after}} \text{ is an internal-copy block of type } r_i \geq 2\}$,
– $A^{\text{non}} := \{i : B^w_i \text{ has type } r_i \geq 2\} \setminus A^{\text{in}}_{\text{after}}$, and
– $A^{\text{non}}_{\text{after}} := A^{\text{non}} \cap \{i : B^w_i \in Y_{w,\text{after}}\}$.

If there exists a *crossing* block B^w_e satisfying $s^w_e < j + g \leq f^w_e$, then this block is unique, since target intervals are disjoint. In this case, we call the interval $T_e := [j + g, f^w_e]$ the *after-suffix* of B^w_e, and let $\rho_e := |\{k : s^{w'}_k \in T_e\}|$ be the number of blocks of $\texttt{LZ77}(w')$ that start inside this after-suffix. If no such block exists, set $\rho_e := 0$. Then we first prove the following auxiliary claim.

Claim 5. $\displaystyle\sum_{i \in A^{\text{in}}_{\text{after}}} (r_i - 1) + \sum_{i \in A^{\text{non}}_{\text{after}}} (r_i - 2) + \max\{\rho_e - 1, 0\} \leq 2\sqrt{ng}.$

Proof of Claim 5. By Lemma 3, if $i \in A^{\text{in}}_{\text{after}}$, then $\phi^w_J(i-1) - \phi^w_J(i) \geq r_i - 1$, and if $i \in A^{\text{non}}_{\text{after}}$, then $\phi^w_J(i-1) - \phi^w_J(i) \geq r_i - 2$. It remains to account for the after-suffix of the crossing block. If $\rho_e \leq 1$, there is nothing to charge. Suppose $\rho_e \geq 2$, and let $B^{w'}_{k_1}, \ldots, B^{w'}_{k_{\rho_e}}$ be the blocks of $\texttt{LZ77}(w')$ whose starting positions lie in T_e, ordered by their starting positions. Translate their target intervals back to the source side with the offset $s^w_e - q^w_e$, and write $I_a := [s^{w'}_{k_a} - (s^w_e - q^w_e), f^{w'}_{k_a} - (s^w_e - q^w_e)]$ for $a = 1, \ldots, \rho_e$.

Since B^w_e is a crossing block, i.e., $s^w_e < j + g$, I_a's can intersect the left boundary of J, but they cannot cross the right boundary of J. Hence, $I_2, \ldots, I_{\rho_e}$ are all contained in J (I_1 can cross the left boundary of J). Moreover, each such interval is newly charged in the same way as in Lemma 3: if some I_a with $a \geq 2$ had already been covered earlier, then the corresponding block $B^{w'}_{k_a}$ could be extended, contradicting maximality. Thus, the after-suffix contributes at least $\rho_e - 1$ additional charged intervals.

Combining the charges from the blocks starting after J and from this after-suffix block, every charged interval either has length at most K and is charged to a distinct initially uncovered interval counted by the first term of $\phi^w_J(0)$, or has length greater than K and is charged to the length budget in the second term. The same disjointness argument in Lemma 3 shows that the total number of charges is at most $\phi^w_J(0)$. Hence, $\sum_{i \in A^{\text{in}}_{\text{after}}} (r_i - 1) + \sum_{i \in A^{\text{non}}_{\text{after}}} (r_i - 2) + \max\{\rho_e - 1, 0\} \leq \phi^w_J(0) \leq 2\sqrt{ng}$. ∎

We now return to the proof of Lemma 4. By the definitions of $A^{\text{in}}_{\text{after}}$ and A^{non},

$$\sum_{r \geq 2} (r-1) y^{w,\text{in}}_{r,\text{after}} + \sum_{r \geq 2} (r-2)\left(y^w_r - y^{w,\text{in}}_{r,\text{after}}\right) = \sum_{i \in A^{\text{in}}_{\text{after}}} (r_i - 1) + \sum_{i \in A^{\text{non}}} (r_i - 2). \tag{3}$$

Now consider any $i \in A^{\text{non}} \setminus Y_{w,\text{after}}$. By Lemma 1, every such block must intersect J. If such a block does not cross the right endpoint of J, i.e., if $f^w_i < j + g$, then all blocks of w' starting inside B^w_i — except possibly the unique one starting before J — start inside J. Hence, $r_i - 2 \leq \eta_i$.

The only remaining non-after case is the crossing block B^w_e, if it exists. At most one block of w' starting inside B^w_e can start before J. The remaining starts are either inside J or inside the after-suffix T_e. This implies $r_e \leq 1 + \eta_e + \rho_e$. Hence, $r_e - 2 \leq \eta_e + \max\{\rho_e - 1, 0\}$. Therefore,

$$\sum_{i \in A^{\text{non}}} (r_i - 2) \leq \sum_{i \in A^{\text{non}}_{\text{after}}} (r_i - 2) + \max\{\rho_e - 1, 0\} + \sum_i \eta_i. \tag{4}$$

Combining Equation (3), Equation (4) with Claim 5, we have

$$\begin{aligned}\sum_{r\geq 2}(r-1)y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r\geq 2}(r-2)\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) &= \sum_{i\in A^{\mathsf{in}}_{\mathsf{after}}}(r_i-1) + \sum_{i\in A^{\mathsf{non}}}(r_i-2)\\ &\leq \sum_{i\in A^{\mathsf{in}}_{\mathsf{after}}}(r_i-1) + \sum_{i\in A^{\mathsf{non}}_{\mathsf{after}}}(r_i-2) + \max\{\rho_e-1,0\} + \sum_i \eta_i\\ &\leq 2\sqrt{ng}+g,\end{aligned}$$

where the last inequality follows from the fact that $\sum_i \eta_i \leq g$ as in Equation (2). This completes the proof. □

Claim 6. *Let $i_1 < i_2$. If $B^w_{i_1} = [q^w_{i_1}, \ell^w_{i_1}, c^w_{i_1}] \in Y_{w,\mathsf{after}}$ is a type-r_1 block and $B^w_{i_2} = [q^w_{i_2}, \ell^w_{i_2}, c^w_{i_2}] \in Y_{w,\mathsf{after}}$ is a type-r_2 block for both $r_1, r_2 \geq 2$. Then $(q^w_{i_1}, \ell^w_{i_1}) \neq (q^w_{i_2}, \ell^w_{i_2})$.*

The proof of Claim 6 is deferred to Appendix C.1 as it is similar to the proof of [BLG25, Claim 17].

Lemma 5. *Let $(B^w_1, \ldots, B^w_{t_w}) \leftarrow \mathtt{LZ77}(w)$ and $(B^{w'}_1, \ldots, B^{w'}_{t_{w'}}) \leftarrow \mathtt{LZ77}(w')$ for $w \overset{g}{\sim} w'$. Then $|t_w - t_{w'}| \leq 2g + 2\sqrt{ng} + \left(\frac{3n}{2}\right)^{2/3} + \left(\frac{3n}{2}\right)^{1/3}$.*

Proof. By symmetry, it suffices to upper bound $t_{w'} - t_w$. Recall that $t_w = \sum_{r\geq 0} y^w_r$ and $t_{w'} = \sum_{r\geq 1} r\cdot y^w_r$. Therefore,

$$\begin{aligned}t_{w'} - t_w &\leq \sum_{r\geq 1}(r-1)y^w_r\\ &= \sum_{r\geq 1}(r-1)y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r\geq 2}(r-1)\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right)\\ &= \sum_{r\geq 2}(r-1)y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r\geq 2}(r-2)\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) + \sum_{r\geq 2}\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right)\\ &\leq g + 2\sqrt{ng} + \sum_{r\geq 2}\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right). \qquad \blacktriangleleft \text{ by Lemma 4}\end{aligned} \tag{5}$$

Thus, it remains to upper bound $\sum_{r\geq 2}\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right)$. We observe that

$$\sum_{r\geq 2}\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) = \sum_{r\geq 2}\left(y^w_r - y^w_{r,\mathsf{after}}\right) + \sum_{r\geq 2}\left(y^w_{r,\mathsf{after}} - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right). \tag{6}$$

We upper bound each part in Claim 7 and Claim 8.

Claim 7. $\sum_{r\geq 2}\left(y^w_r - y^w_{r,\mathsf{after}}\right) \leq g$.

Proof. Recall that $y^w_r - y^w_{r,\mathsf{after}} = y^w_{r,\mathsf{before}} + y^w_{r,\mathsf{intersect}}$. By Lemma 1, for every $r \geq 2$, we have $y^w_{\mathsf{before}} = 0$. Thus, $\sum_{r\geq 2}\left(y^w_r - y^w_{r,\mathsf{after}}\right) = \sum_{r\geq 2}\left(y^w_{r,\mathsf{before}} + y^w_{r,\mathsf{intersect}}\right) = \sum_{r\geq 2} y^w_{\mathsf{intersect}} \leq g$, where the last inequality comes from Lemma 2. ■

Claim 8. $\sum_{r\geq 2}\left(y_r^{w,\mathsf{after}} - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) \leq \left(\frac{3n}{2}\right)^{2/3} + \left(\frac{3n}{2}\right)^{1/3}$.

Proof Sketch. The proof follows largely the same strategy as prior work [BLG25, Lemma 2], and we give a proof sketch here. The full proof can be found in Appendix C.1. For each $\ell \geq 1$, let

$$S_{\ell,\mathsf{after}}^{w,\mathsf{non}} := \{i : B_i^w = [q_i^w, \ell_i^w = \ell, c_i^w] \in Y_{w,\mathsf{after}}, \\ B_i^w \text{ is of type } r \geq 2 \text{ and } B_i^w \text{ is not an internal-copy block}\},$$

and $x_{\ell,\mathsf{after}}^{w,\mathsf{non}} := |S_{\ell,\mathsf{after}}^{w,\mathsf{non}}|$. Then the total number of non-internal-copy type-$r(\geq 2)$ blocks in $Y_{w,\mathsf{after}}$ is given by the sum $\sum_{\ell\geq 1} x_{\ell,\mathsf{after}}^{w,\mathsf{non}}$. Rearranging the sum by summing over $r \geq 2$, we have

$$\sum_{r\geq 2}\left(y_{r,\mathsf{after}}^{w} - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) = \sum_{\ell\geq 1} x_{\ell,\mathsf{after}}^{w,\mathsf{non}}.$$

Thus, it remains to upper bound $\sum_{\ell\geq 1} x_{\ell,\mathsf{after}}^{w,\mathsf{non}}$. At this point, the argument is analogous to the final optimization step in the proof of [BLG25, Lemma 2]. There, one maximizes $\sum_{\ell\geq 0} x_{\ell,\mathsf{after}}^{w}$ subject to the constraints (1) $x_{\ell,\mathsf{after}}^{w} \leq \ell$, and (2) $\sum_{\ell\geq 0} x_{\ell,\mathsf{after}}^{w}(\ell+1) \leq n$. In our setting, constraint (1) is replaced by $x_{\ell,\mathsf{after}}^{w,\mathsf{non}} \leq 2\ell$ (see Claim 12 in Appendix C.1) and constraint (2) is replaced by $\sum_{\ell\geq 1} x_{\ell,\mathsf{after}}^{w,\mathsf{non}}(\ell+1) \leq n$. Now, a standard swapping argument implies that the optimal solution is obtained by filling the smallest lengths first, i.e., there is a threshold z such that $x_{\ell,\mathsf{after}}^{w,\mathsf{non}} = 2\ell$ for all $\ell < z$, $0 < x_{z,\mathsf{after}}^{w,\mathsf{non}} \leq 2z$, and $x_{\ell,\mathsf{after}}^{w,\mathsf{non}} = 0$ for all $\ell > z$. The budget constraint (updated constraint (2) above) then implies $z \leq \left(\frac{3n}{2}\right)^{1/3}$. Setting this value of z gives the desired result. ∎

To finish proving Lemma 5, combining Equation (5) and Equation (6), along with Claim 7 and Claim 8, we conclude that

$$\begin{aligned} t_{w'} - t_w &\leq 2g + 2\sqrt{ng} + \left(\frac{3n}{2}\right)^{2/3} + \left(\frac{3n}{2}\right)^{1/3} \\ &= O\left(n^{2/3} + \sqrt{ng} + g\right). \end{aligned}$$ □

Combining Lemma 1, Lemma 2, and Lemma 5 altogether, we finally get the g-consecutive sensitivity of the LZ77 compression as follows.

Theorem 2. *Let* $\mathtt{LZ77} : \Sigma^* \to (\Sigma')^*$ *be the LZ77 compression function with unbounded sliding window. Then for* $g \geq 1$,

$$\begin{aligned} \mathtt{GS}_{\mathtt{LZ77}}^{g}(n) &\leq \left(2\sqrt{ng} + \left(\frac{3n}{2}\right)^{2/3} + \left(\frac{3n}{2}\right)^{1/3} + 2g\right)\left(2\lceil \log n\rceil + \lceil \log |\Sigma|\rceil\right) \\ &= O\left((n^{2/3} + \sqrt{ng} + g)\log n\right). \end{aligned}$$

Proof. Recall that each LZ77 block is encoded using at most $2\lceil \log n\rceil + \lceil \log |\Sigma|\rceil$ bits. Hence,

$$\mathtt{GS}_{\mathtt{LZ77}}^{g}(n) \leq \left(2\lceil \log n\rceil + \lceil \log |\Sigma|\rceil\right) \cdot \max_{w\in\Sigma^n} \max_{w'\in\Sigma^n : w \overset{g}{\sim} w'} |t_w - t_{w'}|. \tag{7}$$

Combining Equation (7) with Lemma 5, we get the desired result. □

4.2 g-Consecutive Sensitivity of LZ77 with a Bounded Window

Our analysis extends to the setting with a bounded-window size $W < n$, where every copied substring must lie entirely within the previous W symbols. Since the main technical points were already addressed in Section 3.4, we defer the details of the discussion and the proof of Theorem 1 to Appendix A.

Reminder of Theorem 1. *Let* $\mathtt{LZ77}_W : \Sigma^* \to (\Sigma')^*$ *be the LZ77 compression function with sliding window size* $W \le n$. *Then for* $g \ge 1$,

$$\begin{aligned}\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) &\le \Big(5\min\{W, \sqrt{Wg}\} + (3W)^{2/3} + (3W)^{1/3} + 2g\Big)\big(2\lceil \log n\rceil + \lceil \log|\Sigma|\rceil\big)\\ &= O\Big(\Big(W^{2/3} + \sqrt{Wg} + g\Big)\log n\Big).\end{aligned}$$

Theorem 1 exhibits a natural transition at $g = W^{1/3}$. If $g \le W^{1/3}$, then $\sqrt{Wg} \le W^{2/3}$ and therefore Theorem 1 simplifies to $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = O\big(W^{2/3}\log n\big)$. Thus, for contiguous edits of length at most $W^{1/3}$, the asymptotic sensitivity bound matches the known single-character sensitivity bound for bounded-window LZ77. In particular, calibrating randomized padding to the single-character sensitivity already suffices, up to asymptotic order, to protect contiguous substrings of length $g \le W^{1/3}$.

On the other hand, if $W^{1/3} < g \le W$, then the term $\sqrt{Wg}$ dominates $W^{2/3}$ and g and Theorem 1 simplifies to $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = O\Big(\sqrt{Wg}\log n\Big)$. If $g > W$, then the g term dominates, and the bound becomes $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = O(g\log n)$. Hence, beyond the $W^{1/3}$ threshold, protecting a length-g contiguous substring requires additional padding beyond what is suggested by the single-character sensitivity alone.

This additional dependence on $\sqrt{Wg}\log n + g\log n$ is not merely an artifact of our proof: in the unbounded-window case $W = n$, our lower bound, proved later in Theorem 5, shows that these terms are necessary when the alphabet size $|\Sigma| = \Theta(n)$ is large. Even if the alphabet is binary, the g-consecutive sensitivity can be lower bounded as $\tilde{\Omega}(g)$ and $\tilde{\Omega}(\sqrt{Wg})$ so, up to polylogarithmic factors, the $\sqrt{Wg}\log n + g\log n$ terms in Theorem 1 are necessary.

4.3 g-Consecutive Sensitivity of LZ77 with Self-Referencing

Our analysis also extends to the self-referencing variant of LZ77. In this variant, a copied substring may overlap with the target interval currently being generated. In the unbounded-window setting, this does not change the after-part analysis. The potential-drop argument still relies on maximality of the chosen match, and the source-pair uniqueness/counting argument for non-internal-copy type-$r \ge 2$ after-blocks remains valid. The only additional contribution comes from blocks whose target intervals intersect the edited interval J. Since these target intervals are pairwise disjoint and each such block contributes at least one position of J, there are at most g such blocks. This contribution is already included in the g-term of the unbounded-window bound.

In the bounded-window setting, however, one cannot apply the non-self-referencing proof verbatim. The window size only bounds the offset of a copy, but not necessarily the *copied length*; a block may encode a long periodic substring by referring to symbols that are themselves being generated.

This requires two modifications to the proof. First, for non-internal-copy type-$r \geq 3$ blocks that start after J, the potential-drop argument is weakened from an $(r-2)$ drop to an $(r-3)$ drop: one additional intermediate translated interval may cross from J into the periodic tail generated by self-referencing. Internal-copy blocks are unaffected and still give a drop of $r-1$, because their relevant translated source-side span is contained in J. Second, when counting non-internal-copy type-$r \geq 2$ blocks that start after J, the bounded-window length budget is applied *after removing* the last type-$r(\geq 2)$ "after"-block, which can be arbitrarily long; this removed block contributes only one to the block count. These modifications affect only constant factors. Consequently, the same asymptotic bounds hold for the g-consecutive sensitivity of LZ77 with self-referencing. We defer the details to Appendix B.

Theorem 3. *Let* $\mathtt{LZ77}_W^{\mathtt{SR}} : \Sigma^* \to (\Sigma')^*$ *denote the self-referencing LZ77 compression function with sliding window size* $W \leq n$*. Then, for every* $g \geq 1$*, the following bounds hold.*

If $W = n$*,*

$$\mathtt{GS}^g_{\mathtt{LZ77}_W^{\mathtt{SR}}}(n) \leq \left(2\sqrt{ng} + \left(\frac{3n}{2}\right)^{2/3} + \left(\frac{3n}{2}\right)^{1/3} + 2g\right)\left(2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil\right).$$

If $W < n$*,*

$$\mathtt{GS}^g_{\mathtt{LZ77}_W^{\mathtt{SR}}}(n) \leq \left(3g + 5\min\{W, \sqrt{Wg}\} + 2(3W)^{2/3} + 2(3W)^{1/3} + 2\right)\left(2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil\right).$$

In particular, for every $W \leq n$, $\mathtt{GS}^g_{\mathtt{LZ77}_W^{\mathtt{SR}}}(n) \leq O\left(\left(W^{2/3} + \sqrt{Wg} + g\right)\log n\right)$.

5 Lower Bound for the g-consecutive Sensitivity of LZ77

In this section, we present our lower bounds on the g-consecutive sensitivity of LZ77. We show that $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = \Omega(g \log n)$ and $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = \Omega(\sqrt{Wg}\log n)$. This implies that our upper bounds are asymptotically tight when the alphabet size is large $|\Sigma| = \Theta(n)$. We can reduce the size of the alphabet in our lower bounds to 2 by replacing each symbol in Σ with a random binary string of length $O(\log n)$ — this substitution will impact the lower bounds but only by a $O(\log n)$ factor. Thus, when the alphabet size is constant $O(1)$, our upper bounds are still tight up to a $O(\log n)$ factor.

5.1 String Constructions for Lower Bounds (Large Alphabet)

5.1.1 Special Case: Large $g > W^{2/3}$. We first present our lower bounds using a large alphabet $|\Sigma| = \Theta(n)$. We now demonstrate that when the block edit size g is sufficiently large, the global sensitivity of LZ77 can grow linearly with g. In particular, we exhibit an example showing that when $g, W = \Theta(n)$ the sensitivity satisfies $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = \Omega(g \log n)$.

Theorem 4. *Suppose that we have* $g, W = \Theta(n)$ *with* $W \geq g$ *and alphabet size* $|\Sigma| = \Omega(n)$ *then*

$$\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) \geq g(2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil) = \Omega(g \log n)\,.$$

Proof. Let $y = c_1 \circ c_2 \circ \cdots \circ c_g \in \Sigma^g$ and define the strings $w = y \circ y \circ c_{2g+1} \in \Sigma^{2g+1}$, and $w' = y \circ c_{g+1} \circ \cdots \circ c_{2g} \circ c_{2g+1} \in \Sigma^{2g+1}$, over the alphabet $\Sigma = \{c_1, \ldots, c_{2g+1}\}$. Observe that: (1) Both strings have length $n = 2g+1$, (2) the strings differ only on the interval $J = [g+1, 2g]$, so $w \overset{g}{\sim} w'$, (3) the alphabet size is $|\Sigma| = 2g+1$.

We will set $W = n$ and consider running the LZ77 compression scheme on both strings w and w'. For the string w' the LZ77 blocks are $B_i^{w'} = [0, 0, c_i]$ for all $i \le 2g+1$ since each character in w' is distinct. For the string w, the first g characters are distinct so we have $B_i^w = [0, 0, c_i]$ for all $i \le g$. However, LZ77 only needs a single block $B_i^w = [1, g, c_{2g+1}]$ to encode the final $g+1$ characters of w. Thus, we use $2g+1$ blocks to encode w' and only $g+1$ blocks to encode w. We have $|\mathtt{LZ77}_W(w')| = (2g+1)(2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil)$, but $|\mathtt{LZ77}_W(w)| = (g+1)(2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil)$. It follows that

$$\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) \ge \big||\mathtt{LZ77}_W(w')| - |\mathtt{LZ77}_W(w)|\big| \ge g(2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil)\,. \qquad \square$$

5.1.2 Main Lower Bound: The $\sqrt{Wg}$ Term. We now illustrate the necessity of the $O(\sqrt{Wg}(\log n + \log|\Sigma|))$ term from our upper bound. Theorem 5 shows that, in the unbounded-window regime $W = n$ with $g > W^{1/3}$, we have $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = \Omega(\sqrt{Wg}(\log n + \log|\Sigma|))$ over large alphabets $|\Sigma| = \Omega(n)$.

Theorem 5. *Suppose that $|\Sigma| = \Omega(n)$ and $W = n$. Then for any parameter $g \ge W^{1/3}$ we have*

$$\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = \Omega(\sqrt{Wg}(\log n + \log|\Sigma|))\,.$$

Proof. We will set $W = n$ and define the alphabet $\Sigma = \{c_i\}_{i=1}^g \cup \{b_i\}_{i=1}^g \cup \{d_{i,j}\}_{i,j \le g}$. If necessary, we add unused dummy symbols to the alphabet so that $|\Sigma| = \Omega(n)$.

Let $C := c_1 \circ \ldots \circ c_g$. For a parameter $t \le g$, we define $x_t := C \circ L_1 \circ L_2 \circ \ldots \circ L_t$ and $x'_t := b_1 \circ \ldots \circ b_g \circ L_1 \circ L_2 \circ \ldots \circ L_t$ where, for each $r \le t$, $L_r := C[1, r] \circ d_{r,1} \circ C[2, r+1] \circ d_{r,2} \circ \ldots C[i, i+r-1] \circ d_{r,i} \circ \ldots \circ C[g-r+1, g] \circ d_{r,g-r+1}$. We make a couple of observations: (1) We have $x_t[i] = x'_t[i]$ for all $i > g$. Thus, $x_t \overset{g}{\sim} x'_t$. (2) The string L_r has length $|L_r| = (g-r+1)(r+1)$. The total length of the string x_t is $g + \sum_{r=1}^t (g-r+1)(r+1) = \Theta(gt^2)$. Thus, we set $t = \Theta(\sqrt{n/g})$ so that $|x_t| = n$.

We now consider the LZ77 blocks of x_t. The first g blocks are $B_i^x = [0, 0, c_i]$ for $i \le g$ — encoding $x_t[1, g]$. After this the next g blocks are $B_{g+i}^x = [i, 1, d_{1,i}]$ for $i \le g$. These g blocks encode L_1 and the individual block B_{g+i}^x encodes $C[i] \circ d_{1,i}$. After this the next $g-1$ blocks are $B_{2g+i}^x = [i, 2, d_{2,i}]$ for $i \le g-1$. These blocks encode L_2 and the individual block B_{2g+i}^x encodes $C[i, i+1] \circ d_{2,i}$. Define $p_r := g + \big(g + (g-1) + \ldots + (g-(r-2))\big)$. In general after L_{r-1} has been encoded the next $g-r+1$ blocks will be $B_{p_r+i}^x = [i, r, d_{r,i}]$ for $i \le g-r+1$. Thus, the total number of blocks to encode x_t is $p_{t+1} = g + \big(g + (g-1) + \ldots + (g-t+1)\big)$.

We now consider the LZ77 blocks of x'_t. The first g blocks are $B_i^{x'} = [0, 0, b_i]$ for $i \le g$ — encoding $x'_t[1, g]$. After this the next $2g$ blocks are $B_{g+2i-1}^{x'} = [0, 0, c_i]$ and $B_{g+2i}^{x'} = [0, 0, d_{1,i}]$ for $i \le g$. These $2g$ blocks encode L_1. The key difference is that we now use two blocks $B_{g+2i-1}^{x'} = [0, 0, c_i]$ and $B_{g+2i}^{x'} = [0, 0, d_{1,i}]$ to encode $C[i] \circ d_{1,i}$. Similarly, to encode L_2 we will need $2g-2$ blocks i.e., we need two blocks $B_{3g+2i-1}^{x'}$ and $B_{3g+2i}^{x'}$ to encode the string $C[i, i+1] \circ d_{2,i}$. Because the string $C[i, i+1]$ has not been seen before, but the string $C[i, i] = c_i$ has, so the block $B_{g+2i-1}^{x'}$ will encode

$C[i, i+1]$ and the block $B^{x'}_{g+2i} = [0, 0, d_{2,i}]$ will only encode the character $d_{2,i}$ since this is the first and only occurrence of that character. In general, we will need $2(g-r+1)$ blocks to encode L_r once L_{r-1} has been encoded. In particular, we need two blocks $B^{x'}_{p'_r+2i-1}$ and $B^{x'}_{p'_r+2i}$ to encode the string $C[i, i+r-1] \circ d_{r,i}$ where $p'_r := g + 2\big(g + (g-1) + \ldots + (g-(r-2))\big)$. The block $B^{x'}_{p'_r+2i-1}$ will encode $C[i, i+r-1]$ since the string $C[i, i+r-1]$ has not been seen before but the substring $C[i, i+r-2]$ has been seen before. The block $B^{x'}_{p'_r+2i} = [0, 0, d_{r,i}]$ will only encode the next character $d_{r,i}$ since this is the first and only occurrence of that character. Thus, the total number of blocks to encode x'_t is $p'_{t+1} = g + 2\big(g + (g-1) + \ldots + (g-t+1)\big)$.

The difference between the number of blocks for x and x' is $p'_{t+1} - p_{t+1} = \sum_{i=0}^{t-1}(g-i) = \Theta(gt)$. Since $W = \Theta(t^2 g)$ we have $p'_{t+1} - p_{t+1} = \Theta(\sqrt{Wg})$. Thus, it follows that

$$||\mathtt{LZ77}_W(x_t)| - |\mathtt{LZ77}_W(x'_t)|| = \Omega\left(\sqrt{Wg}\left(\log n + \log|\Sigma|\right)\right). \qquad \square$$

Discussion. If we set $t = \sqrt{g}$ then $|x_t| = W = O(g^2)$ and $gt = \Theta(\sqrt{Wg})$. Thus, the lengths of the corresponding LZ77 compressions will differ by $\Omega(\sqrt{Wg}(\log n + \log|\Sigma|))$ bits. Observe that, in this example, the term $\sqrt{Wg} = \Theta(g^{1.5})$ dominates the other terms g and $W^{2/3} = \Theta(g^{4/3})$ from our upper bound. Thus, the $O(\sqrt{Wg}(\log n + \log|\Sigma|))$ term is truly necessary.

5.2 Binary Alphabet

We now explain how to reduce the size of the alphabet to 2 by replacing each symbol in Σ with a random binary string. To avoid confusing the parameters before and after the encoding, in this subsection, we write N, g_0, W_0 for the length, edit length, and window size of the large-alphabet construction, and reserve n, g, W for the corresponding binary parameters. In this subsection, we transfer the large-alphabet construction from Theorem 5. The binary $\Omega(g)$ lower bound will be obtained by taking $t = 1$ in the same construction.

Let $\mathcal{D}$ be the set of delimiter symbols $d_{r,i}$ that appear in the chosen construction from Theorem 5, and fix an arbitrary ordering $\mathcal{D} = \{d_1, \ldots, d_M\}$. Let $d_\star$ be an auxiliary delimiter symbol that does not appear in either w or w'. Define the dictionary prefix

$$P_{\mathcal{D}} := d_1 \circ \ldots \circ d_M \circ d_\star.$$

For the binary transfer, instead of encoding w and w' directly, we encode the prefixed strings $\widehat{w} := P_{\mathcal{D}} \circ w$ and $\widehat{w}' := P_D \circ w'$. The prefix is identical in the two strings and lies before the edited interval J. Moreover, every delimiter symbol $d_{r,i}$ has already appeared before it is used in the lower-bound gadget.

Random Binary Encoding. Let $w, w' \in \Sigma^N$ be the large-alphabet strings from Theorem 5, and let $\widehat{w}, \widehat{w}' \in (\Sigma \cup \{d_\star\})^{N^*}$ be the prefixed strings defined above, where $N^* := N + M + 1$. Since $M = O(N)$ in the construction, we have $N^* = \Theta(N)$. We choose a random map $\varphi : (\Sigma \cup \{d_\star\}) \to \{0,1\}^h$ by sampling the codeword $\varphi(a)$ independently and uniformly from $\{0,1\}^h$ for each $a \in \Sigma \cup \{d_\star\}$, where $h = C\lceil \log N^* \rceil$ for a sufficiently large constant C. We extend φ to a map $\Phi : (\Sigma \cup \{d_\star\})^{N^*} \to \{0,1\}^{hN^*}$ such that for $x = x_1 \circ \ldots \circ x_{N^*}$,

$$\Phi(x) := \varphi(x_1) \circ \ldots \circ \varphi(x_{N^*}).$$

For each $i = 1, \dots, N^*$, we call $\varphi(x_i)$ a *codeword* of the encoding $\Phi(x)$. If w and w' differ only on an interval of length g_0, then $\widehat{w}$ and $\widehat{w}'$ also differ only on an interval of length g_0, shifted by the common prefix. Hence, $\Phi(\widehat{w})$ and $\Phi(\widehat{w}')$ differ only on the corresponding binary interval of length $g := hg_0$. The resulting binary strings have length $n := hN^*$. For the window size, we use $W := h(W_0 + M + 1)$. Since $M = O(N)$ and $W_0 = \Theta(N)$ in the construction, we have $W = \Theta(hW_0)$. Since $h = \Theta(\log N^*)$, we also have $h = \Theta(\log n)$.

For a position $p \in [n]$, let

$$\mathsf{idx}_h(p) := \left\lceil \frac{p}{h} \right\rceil \qquad \text{and} \qquad \mathsf{off}_h(p) := p - (\mathsf{idx}_h(p) - 1)h.$$

Here, $\mathsf{idx}_h(p)$ denotes the codeword index, i.e., p lies in the $\mathsf{idx}_h(p)^{th}$ codeword and has offset $\mathsf{off}_h(p) \in [h]$. Equivalently, for every $x \in (\Sigma \cup \{d_\star\})^{N^*}$, we have

$$\Phi(x)[p] = \varphi(x_{\mathsf{idx}_h(p)})[\mathsf{off}_h(p)].$$

The main reason why we cannot directly translate the large-alphabet analysis into the binary-alphabet case is that LZ77 over the binary alphabet is not forced to respect codeword boundaries. Thus, after replacing symbols by binary codewords, the parser could, in principle, use a long non-aligned bit-level match that did not exist at the symbol level. Lemma 6 rules this possibility out: for a suitable choice of the random codewords, every sufficiently long binary match inside $\Phi(\widehat{w})$ or $\Phi(\widehat{w}')$ contains a long symbol-level match.

Lemma 6 (Good Encoding Lemma). *Let $\widehat{w}, \widehat{w}' \in (\Sigma \cup \{d_\star\})^{N^*}$ be the prefixed large-alphabet strings defined above, and suppose that $|\Sigma \cup \{d_\star\}| \le (N^*)^A$ for some constant $A > 0$. Then there exists a constant $C = C(A) > 0$ such that, for $h = C\lceil \log N^* \rceil$, there exists a map $\varphi : (\Sigma \cup \{d_\star\}) \to \{0,1\}^h$ with the following property: for every $x \in \{\widehat{w}, \widehat{w}'\}$, every pair of positions p, q, and every $L \ge 2h$, if*

$$\Phi(x)[p, p + L - 1] = \Phi(x)[q, q + L - 1],$$

then there exist intervals I_1, I_2 in x such that $x[I_1] = x[I_2]$ and $|I_1| = |I_2| \ge \left\lfloor \frac{L}{h} \right\rfloor - 1$.

Proof. Sample $\varphi(a) \leftarrow\!\!\$\ \{0,1\}^h$ independently and uniformly for every $a \in \Sigma \cup \{d_\star\}$. For $x \in \{\widehat{w}, \widehat{w}'\}$, let $\mathsf{bad}(x, p, q, L)$ be the event that $\Phi(x)[p, p + L - 1] = \Phi(x)[q, q + L - 1]$ but there are no intervals I_1, I_2 in x such that $x[I_1] = x[I_2]$ and $|I_1| = |I_2| \ge \left\lfloor \frac{L}{h} \right\rfloor - 1$. Now we observe the following claim.

Claim 9. *For every fixed $x \in \{\widehat{w}, \widehat{w}'\}$, positions p, q, and $L \ge 2h$,*

$$\Pr[\mathsf{bad}(x, p, q, L)] \le 2^{-h/2}.$$

Proof of Claim 9. Let $p_0 := \min\{u \ge p : \mathsf{off}_h(u) = 1\}$ and $m := \left\lfloor \frac{L}{h} \right\rfloor - 1$. Then since $p_0 < p + h$, we have $[p_0, p_0 + mh - 1] \subseteq [p, p + L - 1]$. Let $q_0 := q + (p_0 - p)$. On the event $\Phi(x)[p, p + L - 1] = \Phi(x)[q, q + L - 1]$, we have

$$\Phi(x)[p_0, p_0 + mh - 1] = \Phi(x)[q_0, q_0 + mh - 1]. \tag{8}$$

The left-hand side of Equation (8) is the aligned encoding of the symbol interval $x[\mathsf{idx}_h(p_0), \mathsf{idx}_h(p_0) + m - 1]$. If $\mathsf{off}_h(q_0) = 1$ and $x[\mathsf{idx}_h(p_0), \mathsf{idx}_h(p_0) + m - 1] = x[\mathsf{idx}_h(q_0), \mathsf{idx}_h(q_0) + m - 1]$, then the desired symbol-level intervals exist, which contradicts the bad event $\mathsf{bad}(x, p, q, L)$. Hence, on $\mathsf{bad}(x, p, q, L)$, Equation (8) is not induced by the same aligned symbol sequence. Now, there are two cases:

(1) If q_0 is aligned, i.e., $\mathsf{off}_h(q_0) = 1$, then the two aligned symbol sequences differ in some position. Hence, Equation (8) forces $\varphi(a) = \varphi(b)$ for two distinct symbols $a \neq b \in \Sigma \cup \{d_\star\}$. This has a probability 2^{-h}.
(2) If q_0 is not aligned, then an aligned length-h codeword on the left is matched against a length-h substring starting at a nonzero offset on the right. In the worst case, this forces a random codeword to agree with a nontrivial shift of itself. For a nonzero shift d, the probability that a uniform h-bit word has period d is at most $2^{-(h-\gcd(h,d))} \leq 2^{-h/2}$. If the right-hand side crosses two independent codewords, the probability becomes only smaller.

Thus, in either case, $\Pr[\mathsf{bad}(x, p, q, L)] \leq 2^{-h/2}$. ■

Going back to the proof of Lemma 6, it remains to take a union bound. There are 2 choices for x. Each encoded string has length hN^*, so there are at most $(hN^*)^2$ choices for p, q and at most hN^* choices for L. Thus, if we define $\mathsf{bad}(x)$ to be the event that $\mathsf{bad}(x, p, q, L)$ occurs for some p, q, L, then

$$\Pr[\mathsf{bad}(x)] \leq (hN^*)^3 \cdot 2^{-h/2}.$$

Taking a sufficiently large C makes $(hN^*)^3 \cdot 2^{-h/2} < 1$. Therefore, there exists a choice of $\varphi : \Sigma \cup \{d_\star\} \to \{0,1\}^h$ for which no bad event occurs. □

The dictionary prefix handles a separate first-occurrence issue. Many delimiter symbols $d_{r,i}$ appear only once inside the large-alphabet gadget. Without the prefix, the binary codeword $\varphi(d_{r,i})$ might require several LZ77 blocks the first time it is encountered, and this cost could depend on the surrounding context. Since every delimiter codeword already appears in the common prefix, each later occurrence of $\varphi(d_{r,i})$ can be encoded using one block from the prefix. The final dummy delimiter $d_\star$ appears only in the common prefix and prevents symbol-level matches from crossing the prefix/body boundary. Thus, delimiter codewords do not introduce an asymmetric first-occurrence cost.

By Lemma 6, any sufficiently long match used by the binary LZ77 parser inside $\Phi(\widehat{w})$ or $\Phi(\widehat{w}')$ must come from a symbol-level match in $\widehat{w}$ or $\widehat{w}'$. Therefore, after the dictionary prefix removes first-occurrence effects, the binary encoding cannot create the long unintended matches that would merge the forced extra blocks in the large-alphabet construction. Hence, the block-count gap from the large-alphabet construction is preserved up to a constant factor, as shown in Lemma 7 below.

Lemma 7. *Let $w, w' \in \Sigma^N$ be the large-alphabet strings from Theorem 5, and let $\Delta := |\#\mathtt{LZ77}_W(w) - \#\mathtt{LZ77}_W(w')|$ denote the block-count gap, where $\#\mathtt{LZ77}_W(\cdot)$ denotes the number of blocks. Let $\widehat{w} = P_{\mathcal{D}} \circ w$ and $\widehat{w}' = P_{\mathcal{D}} \circ w'$ be the prefixed strings defined above. Then there exist binary strings $z := \Phi(\widehat{w})$ and $z' := \Phi(\widehat{w}')$ of length $n = hN^*$, with edit length $g = hg_0$ and window size $W = h(W_0 + M + 1) = \Theta(hW_0)$, such that $z \overset{g}{\sim} z'$ and*

$$\big||\mathtt{LZ77}_W(z)| - |\mathtt{LZ77}_W(z')|\big| = \Omega\big(\Delta \cdot \log n\big).$$

Proof Sketch. Here, we only give a proof sketch, and the full proof can be found in Appendix C.2. Let φ be a *good encoding* from Lemma 6, and set $z = \Phi(\widehat{w})$ and $z' = \Phi(\widehat{w}')$. Since the prefix $P_{\mathcal{D}}$ is common and lies before the edited interval, we have $z \overset{g}{\sim} z'$, where $g = hg_0$. Also, $n = hN^*$ and $W = h(W_0 + M + 1) = \Theta(hW_0)$. The common prefix contributes identically to both parsings. Moreover, it ensures that every delimiter codeword $\varphi(d_{r,i})$ has already appeared before the gadget is parsed. Hence, later occurrences of delimiter codewords can be encoded using one block from

the prefix, and there is no asymmetric first-occurrence cost for these codewords. The final dummy delimiter $d_\star$ prevents symbol-level matches from crossing the prefix/body boundary.

We claim that the block-count gap from the large-alphabet construction is preserved up to a constant factor. The only way this could fail is if the binary parser used long bit-level matches that do not correspond to symbol-level matches, thereby merging many of the forced extra blocks in the large-alphabet parsing. However, by Lemma 6, every binary match of length at least $L \geq 2h$ inside either z or z' contains a symbol-level match of length $\Omega(L/h)$, after losing only a constant number of boundary codewords. Therefore, such a match cannot bypass the forced block boundaries in the large-alphabet construction, except for a constant number of boundary losses. Discarding these constant-size boundary losses preserves a constant fraction of the Δ forced extra blocks.

Consequently, each constant-size group of forced extra large-alphabet blocks still forces at least one extra binary LZ77 block. Thus, the binary block-count gap is $\Omega(\Delta)$. Finally, each binary LZ77 block is encoded using $2\lceil \log n \rceil + \log 2 = \Theta(\log n)$ bits. This concludes the proof. □

Theorem 6 gives the binary-alphabet version of the lower bounds for the g-consecutive sensitivity of LZ77 from Section 5.1.

Theorem 6. *Let $\Sigma = \{0, 1\}$. For the non-self-referencing LZ77 with sliding window size W, the following binary-alphabet lower bounds hold:*

(1) There are infinitely many choices of (n, g, W) with $W = \Theta(n)$ such that $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = \Omega(g)$.
(2) There are infinitely many choices of (n, g, W) with $W = \Theta(n)$ such that $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = \Omega(\sqrt{Wg})$.

Proof. (1) For the first construction from Theorem 5 with $t = 1$, we have $\Delta = \Omega(g_0)$. Since $g = hg_0$ and $h = \Theta(\log n)$, by Lemma 7, we have $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = \Omega(\Delta \cdot \log n) = \Omega(g_0 \log n) = \Omega(g)$.

(2) For the main construction from Theorem 5, we have $\Delta = \Omega(\sqrt{W_0 g_0})$. Since $W = \Theta(hW_0)$, $g = hg_0$, and $h = \Theta(\log n)$, Lemma 7 gives $\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) = \Omega(\Delta \cdot \log n) = \Omega\left(\sqrt{W_0 g_0} \cdot \log n\right) = \Omega(\sqrt{Wg})$. □

Remark 2. The second lower bound in Theorem 6 can be instantiated in the regime $W = n = \Theta(g^2)$. Recall that g_0 is the large-alphabet edit length, $h = \Theta(\log N) = \Theta(\log n)$, and the final binary edit length is $g = hg_0$. Choose $t = \left\lfloor \sqrt{g_0 \log g_0} \right\rfloor$ in the large-alphabet construction. Then $N = \Theta(g_0 t^2) = \Theta(g_0^2 \log g_0)$ and $h = \Theta(\log N) = \Theta(\log g_0)$. Hence, the final binary edit length is $g = hg_0 = \Theta(g_0 \log g_0)$, and $n = hN^* = \Theta(hN) = \Theta(g_0^2 (\log g_0)^2) = \Theta(g^2)$.

Moreover, by Lemma 7, the binary compressed-length gap is $\Omega(\Delta \cdot \log n) = \Omega(g_0 t \cdot h) = \Omega((g_0 \log g_0)^{3/2}) = \Omega(g^{3/2})$. Since $W = n = \Theta(g^2)$, this is precisely $\Omega(\sqrt{Wg})$.

Acknowledgments. This work was supported in part by NSF CAREER Award CNS-2047272 and by an Amazon Research Award. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or Amazon.

References

AFI23. Tooru Akagi, Mitsuru Funakoshi, and Shunsuke Inenaga. Sensitivity of string compressors and repetitiveness measures. *Inf. Comput.*, 291(C), March 2023.

AS16. Jyrki Alakuijala and Zoltan Szabadka. Brotli Compressed Data Format. RFC 7932, July 2016.

BLG25. Jeremiah Blocki, Seunghoon Lee, and Brayan Sebastián Yepes Garcia. Differentially private compression and the sensitivity of LZ77. In Benny Applebaum and Huijia (Rachel) Lin, editors, *TCC 2025, Part IV*, volume 16271 of *LNCS*, pages 356–386. Springer, Cham, December 2025.

BM08. David L. Black and David McGrew. Using Authenticated Encryption Algorithms with the Encrypted Payload of the Internet Key Exchange version 2 (IKEv2) Protocol. RFC 5282, August 2008.

Deg21. Jean Paul Degabriele. Hiding the lengths of encrypted messages via gaussian padding. In Giovanni Vigna and Elaine Shi, editors, *ACM CCS 2021*, pages 1549–1565. ACM Press, November 2021.

Deu96. L. Peter Deutsch. DEFLATE Compressed Data Format Specification version 1.3. RFC 1951, May 1996.

FSNI25. Yuto Fujie, Hiroki Shibata, Yuto Nakashima, and Shunsuke Inenaga. Tight additive sensitivity on lz-style compressors and string attractors. In *String Processing and Information Retrieval: 32nd International Symposium, SPIRE 2025, London, UK, September 8–11, 2025, Proceedings*, page 109–123, Berlin, Heidelberg, 2025. Springer-Verlag.

GHP13. Yoel Gluck, Neal Harris, and Angelo Prado. BREACH: Reviving the CRIME Attack. https://breachattack.com/, 2013.

PFPSÚGDZ22. Rafael Palacios, Andrea Fariña Fernández-Portillo, Eugenio F Sánchez-Úbeda, and Pablo García-De-Zúñiga. Htb: a very effective method to protect web servers against breach attack to https. *IEEE Access*, 10:40381–40390, 2022.

PSPW20. Brandon Paulsen, Chungha Sung, Peter A. H. Peterson, and Chao Wang. Debreach: mitigating compression side channels via static analysis and transformation. In *Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering*, ASE '19, page 899–911. IEEE Press, 2020.

RD12. J. Rizzo and T. Duong. The CRIME attack. Presentation at Ekoparty 2012, 2012.

Vad17. Salil Vadhan. The complexity of differential privacy. In *Tutorials on the Foundations of Cryptography: Dedicated to Oded Goldreich*, pages 347–450. Springer, 2017.

A g-Consecutive Sensitivity with a Bounded Window

In this section, we consider the bounded-window size $W < n$, where every copied substring must lie entirely within the previous W symbols. Equivalently, if a block $B_i^x = [q_i^x, \ell_i^x, c_i^x]$ is generated at position s_i^x, then its source interval must satisfy $s_i^x - W \le q_i^x \le q_i^x + \ell_i^x - 1 < s_i^x$. Throughout this section, we consider the non-self-referencing version of the LZ77 with a bounded window of size W. We call this compression scheme $\texttt{LZ77}_W$.

We first record two facts that remain unchanged for every window size W. First, before the parsing reaches the edited interval J, the two bounded-window parsings see the same previous W symbols at every step. Hence, as in Lemma 1, there are no type-$r \ge 2$ blocks before J. Second, as in Lemma 2, the number of blocks of w (or w') whose starting position lies inside J is at most g.

The bounded-window restriction affects the after-part analysis in two ways. First, the potential argument only needs to count subintervals of $J \cap [j+g-W, j+g-1]$ instead of J. Second, every type-$r \ge 2$ block in $Y_{w,\mathsf{after}}$ must start within the first W positions after J. We make these two observations precisely below.

Lemma 8. *For* $\texttt{LZ77}_W$ *with* $W < n$,

$$\sum_{r\ge 2}(r-1)y_{r,\mathsf{after}}^{w,\mathsf{in}} + \sum_{r\ge 2}(r-2)\Big(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\Big) \le g + 5\min\{W, \sqrt{Wg}\}.$$

Proof. For the bounded window size $W < n$, define the window-reachable portion of J by

$$J_W := \big[\max\{j, j+g-W\}, j+g-1\big],$$

and let $m := |J_W| = \min\{g, W\}$. Now define a *truncated* potential function $\phi_{J_W}^w(t)$ as follows. For $K_W := \max\Big\{1, \sqrt{W/g}\Big\}$ and $H_W := 2W + j + g - 1$, let

$$\phi_{J_W}^w(t) := \Big|\big\{I \in \mathcal{I}_{K_W}(J_W) : \mathsf{Uncovered}(w, I, t) = 1\big\}\Big| + \frac{\max\{0, H_W - f_t^w\}}{K_W}.$$

Here, we only count uncovered subintervals of J_W of length at most K_W, and we replace the old $\frac{n-f_t^w}{K}$ with $\frac{\max\{0,H_W-f_t^w\}}{K_W}$ because there will not be any type-$r \ge 2$ blocks which start *after* the index $j+g+W-2$ and any block has length at most $W+1$. The reason we define the second term by $\frac{\max\{0,H_W-f_t^w\}}{K_W}$ is to make sure that the telescoping sum is well bounded by $\phi_{J_W}^w(i^*-1)$, since if we define the second term as $\frac{H_W-f_t^w}{K_W}$, then the inequality $\sum_{i=i^*}^{t_w}\Big(\phi_{J_W}^w(i-1) - \phi_{J_W}^w(i)\Big) = \phi_{J_W}^w(i^*-1) - \phi_{J_W}^w(t_w) \le \phi_{J_W}^w(i^*-1)$ is *no longer guaranteed* because $f_{t_w}^w = n$ and $H_W - f_{t_w}^w < 0$ if $H_W < n$.

We now want to repeat the proof of Lemma 4 using the same telescoping argument with this truncated potential. Let i^* be the first index such that the target interval of $B_{i^*}^w$ intersects J_W, i.e., $[s_{i^*}^w, f_{i^*}^w] \cap J_W \ne \varnothing$. If no such index exists, then after discarding the w'-blocks starting inside J, there is no contribution to the weighted sum, and the lemma follows directly from Lemma 2. Hence, assume such i^* exists.

Since $B_{i^*}^w$ has target length at most $W+1$ and intersects J_W, we have $s_{i^*}^w \ge \max\{j, j+g-W\} - W$. Thus, $f_{i^*-1}^w = s_{i^*}^w - 1 \ge \max\{j, j+g-W\} - W - 1$. Therefore, $\phi_{J_W}^w(i^*-1) \le |\mathcal{I}_{K_W}(J_W)| + \frac{H_W - f_{i^*-1}^w}{K_W} \le |\mathcal{I}_{K_W}(J_W)| + \frac{H_W - \max\{j, j+g-W\} + W + 1}{K_W}$.

Let $L' := \min\{m, \lfloor K_W \rfloor\}$. Since J_W has length m, the number of subintervals of J_W of length ℓ is $m - \ell + 1$. Hence, we have $|\mathcal{I}_{K_W}(J_W)| = \sum_{\ell=1}^{L'}(m-\ell+1) = (m+1)L' - \frac{L'(L'+1)}{2} \le mL'$. Hence, $\phi^w_{J_W}(i^*-1) \le mL' + \frac{H_W - \max\{j, j+g-W\}+W+1}{K_W}$. Now we consider two cases:

- If $W \ge g$, then $m = g$ and $L' = \left\lfloor \sqrt{W/g} \right\rfloor$. Hence,

$$\phi^w_{J_W}(i^*-1) \le g\sqrt{\frac{W}{g}} + (3W+g)\sqrt{\frac{g}{W}} \le 5\sqrt{Wg}.$$

  In this case, we see that $5\sqrt{Wg} \le 5W$.
- If $W < g$, then $m = W$ and $K_W = 1$, which implies $L' = 1$. Hence, $\phi^w_{J_W}(i^*-1) \le W + 4W = 5W$. In this case, we see that $5W \le 5\sqrt{Wg}$.

Taken together, we have $\phi^w_{J_W}(i^*-1) \le 5\min\{W, \sqrt{Wg}\}$.

We now use the bounded-window analogue of Claim 5 with the truncated potential $\phi^w_{J_W}$ in place of ϕ^w_J. This gives $\sum_{i\in A^{\mathsf{in}}_{\mathsf{after}}}(r_i-1) + \sum_{i\in A^{\mathsf{non}}_{\mathsf{after}}}(r_i-2) + \max\{\rho_e - 1, 0\} \le \phi^w_{J_W}(i^*-1) \le 5\min\{W, \sqrt{Wg}\}$.

Now as in Lemma 4, we have

$$\begin{aligned}
\sum_{r\ge 2}(r-1)y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r\ge 2}(r-2)\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) &= \sum_{i\in A^{\mathsf{in}}_{\mathsf{after}}}(r_i-1) + \sum_{i\in A^{\mathsf{non}}}(r_i-2) \\
&\le \sum_{i\in A^{\mathsf{in}}_{\mathsf{after}}}(r_i-1) + \sum_{i\in A^{\mathsf{non}}_{\mathsf{after}}}(r_i-2) + \max\{\rho_e-1,0\} + \sum_i \eta_i \\
&\le 5\min\{W, \sqrt{Wg}\} + g. \qquad \square
\end{aligned}$$

Recall that

$$\begin{aligned}
S^{w,\mathsf{non}}_{\ell,\mathsf{after}} := \{i : B^w_i = [q^w_i, \ell^w_i = \ell, c^w_i] \in Y_{w,\mathsf{after}},& \\
B^w_i \text{ is of type } r \ge 2 \text{ and } B^w_i \text{ is not an internal-copy block}\}&
\end{aligned}$$

denotes the index of a non-internal-copy block of $\mathtt{LZ77}_W(w)$ with length ℓ (as we consider the bounded-window setting) that starts after the interval J and of type $r \ge 2$. Also, recall that $x^{w,\mathsf{non}}_{\ell,\mathsf{after}} := |S^{w,\mathsf{non}}_{\ell,\mathsf{after}}|$. Now, we observe that any block that starts after $j+g+W-2$ cannot be a type-$r \ge 2$ block, as shown in Claim 10 below.

Claim 10. *For every $\ell \ge 1$ and every $i \in S^{w,\mathsf{non}}_{\ell,\mathsf{after}}$, we have*

$$j + g \le s^w_i \le j + g + W - 1.$$

Proof. Fix $\ell \ge 1$ and $i \in S^{w,\mathsf{non}}_{\ell,\mathsf{after}}$. By definition, $B^w_i \in Y_{w,\mathsf{after}}$ is a type-r block for some $r \ge 2$. Hence, by Claim 1, its source interval intersects J, i.e., $[q^w_i, q^w_i + \ell - 1] \cap J \ne \emptyset$. In the bounded-window setting, the copied substring lies entirely within the previous W symbols, so $[q^w_i, q^w_i + \ell - 1] \subseteq [s^w_i - W, s^w_i - 1]$. Since $B^w_i \in Y_{w,\mathsf{after}}$, we also have $s^w_i \ge j+g$. Combining those together, we conclude that $j + g \le s^w_i \le j + g + W - 1$. $\square$

Now, we have an updated budget constraint.

Lemma 9. *For* $\mathtt{LZ77}_W$ *with* $W < n$,

$$\sum_{\ell \geq 1} x^{w,\mathsf{non}}_{\ell,\mathsf{after}}(\ell + 1) \leq 2W.$$

Proof. Since $i \in S^{w,\mathsf{non}}_{\ell,\mathsf{after}}$ implies $\ell \leq W$, the target interval of B^w_i is $[s^w_i, f^w_i] = [s^w_i, s^w_i + \ell]$, and hence is contained in $[j + g, j + g + 2W - 1]$. This interval has length $2W$. Since target intervals of distinct LZ77 blocks are pairwise disjoint, the total target length of all blocks counted by $x^{w,\mathsf{non}}_{\ell,\mathsf{after}}$ is at most $2W$. Thus, we have $\sum_{\ell \geq 1} x^{w,\mathsf{non}}_{\ell,\mathsf{after}}(\ell + 1) \leq 2W$. □

Lemma 10. *For* $\mathtt{LZ77}_W$ *with* $W < n$,

$$\sum_{r \geq 2} \left(y^w_{r,\mathsf{after}} - y^{w,\mathsf{in}}_{r,\mathsf{after}} \right) \leq (3W)^{2/3} + (3W)^{1/3}.$$

Proof. Recall that $\sum_{r \geq 2} \left(y^w_{r,\mathsf{after}} - y^{w,\mathsf{in}}_{r,\mathsf{after}} \right) = \sum_{\ell \geq 1} x^{w,\mathsf{non}}_{\ell,\mathsf{after}}$. By Claim 12, for every $\ell \geq 1$, we have $x^{w,\mathsf{non}}_{\ell,\mathsf{after}} \leq 2\ell$. By Lemma 9, we also have $\sum_{\ell \geq 1} x^{w,\mathsf{non}}_{\ell,\mathsf{after}}(\ell + 1) \leq 2W$. Repeating the same swapping argument as in the proof of Claim 8 with budget $2W$ in place of n, gives $\sum_{r \geq 2} \left(y^w_{r,\mathsf{after}} - y^{w,\mathsf{in}}_{r,\mathsf{after}} \right) \leq (3W)^{2/3} + (3W)^{1/3}$. □

Now we are ready to prove the g-consecutive sensitivity of $\mathtt{LZ77}_W$ with $W < n$, as shown in Theorem 1.

Reminder of Theorem 1. *Let* $\mathtt{LZ77}_W : \Sigma^* \to (\Sigma')^*$ *be the LZ77 compression function with sliding window size* $W \leq n$. *Then for* $g \geq 1$,

$$\begin{aligned} \mathtt{GS}^g_{\mathtt{LZ77}_W}(n) &\leq \left(5 \min\{W, \sqrt{Wg}\} + (3W)^{2/3} + (3W)^{1/3} + 2g \right) \left(2\lceil \log n \rceil + \lceil \log |\Sigma| \rceil \right) \\ &= O\left(\left(W^{2/3} + \sqrt{Wg} + g \right) \log n \right). \end{aligned}$$

Proof of Theorem 1. By symmetry, it suffices to upper bound $t_{w'} - t_w$; the case $t_w - t_{w'}$ is obtained by applying the same argument with w and w' interchanged. Recall that $t_w = \sum_{r \geq 0} y^w_r$ and $t_{w'} = \sum_{r \geq 1} r \cdot y^w_r$. Therefore,

$$\begin{aligned} t_{w'} - t_w &\leq \sum_{r \geq 1} (r-1) y^w_r \\ &= \sum_{r \geq 1} (r-1) y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r \geq 2} (r-1) \left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}} \right) \\ &= \sum_{r \geq 2} (r-1) y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r \geq 2} (r-2) \left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}} \right) + \sum_{r \geq 2} \left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}} \right) \\ &\leq g + 5 \min\{W, \sqrt{Wg}\} + \sum_{r \geq 2} \left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}} \right). \qquad \blacktriangleleft \text{ by Lemma 8} \end{aligned}$$

Thus, it remains to upper bound $\sum_{r\geq 2}\left(y_r^w - y_{r,\text{after}}^{w,\text{in}}\right)$. We observe that

$$\sum_{r\geq 2}\left(y_r^w - y_{r,\text{after}}^{w,\text{in}}\right) = \sum_{r\geq 2}\left(y_r^w - y_{r,\text{after}}^{w}\right) + \sum_{r\geq 2}\left(y_{r,\text{after}}^{w} - y_{r,\text{after}}^{w,\text{in}}\right).$$

By Claim 7, we have $\sum_{r\geq 2}\left(y_r^w - y_{r,\text{after}}^{w}\right) \leq g$. And by Lemma 10, we have

$$\sum_{r\geq 2}\left(y_{r,\text{after}}^{w} - y_{r,\text{after}}^{w,\text{in}}\right) \leq (3W)^{2/3} + (3W)^{1/3}.$$

Combining everything together, we have

$$t_{w'} - t_w \leq 2g + 5\min\{W, \sqrt{Wg}\} + (3W)^{2/3} + (3W)^{1/3}.$$

Finally, each LZ77 block can be encoded using at most $2\lceil \log n\rceil + \lceil \log |\Sigma|\rceil$ bits. Hence,

$$\mathtt{GS}^g_{\mathtt{LZ77}_W}(n) \leq \left(2g + 5\min\{W, \sqrt{Wg}\} + (3W)^{2/3} + (3W)^{1/3}\right)\left(2\lceil \log n\rceil + \lceil \log |\Sigma|\rceil\right).$$

The asymptotic bound follows immediately. □

B g-Consecutive Sensitivity of LZ77 with Self-Referencing

Recall that in the self-referencing variant of LZ77, a block $B_i^x = [q_i^x, \ell_i^x, c_i^x]$ is allowed to copy from a source interval that *overlaps the target interval* currently being generated. That is, we may have $q_i^x + \ell_i^x - 1 \geq s_i^x$. This makes it possible to encode long periodic substrings using a single block.

B.1 Unbounded Window Case

We first consider the unbounded window size case. In this case, we remark that the self-referencing variant introduces only a minor additional complication. In the Hamming-distance-one setting, it is known that self-referencing may allow *one additional type-3 block*, namely the block whose target interval contains the edited position [BLG25]. In our setting, the corresponding exceptional blocks are precisely the blocks whose target intervals intersect the edited interval J. Since the target intervals of LZ77 blocks are pairwise disjoint, there are at most g such blocks, which is exactly the same as Lemma 2.

We further observe that the block-count analysis from the non-self-referencing case also applies to the self-referencing variant with an unbounded window. For each block B_i^w, after discarding η_i blocks of $\mathtt{LZ77}^{\mathtt{SR}}(w')$ whose starting positions lie inside J, the same translation and maximality argument applies.

Thus, the same upper bound applies to the self-referencing LZ77 with unbounded window size $W = n$. This leads to Lemma 11.

Lemma 11. *Let $\mathtt{LZ77}^{\mathtt{SR}} : \Sigma^* \to (\Sigma')^*$ denote the self-referencing LZ77 compression function with sliding window size $W = n$. Then, for every $g \geq 1$,*

$$\begin{aligned}\mathtt{GS}^g_{\mathtt{LZ77}^{\mathtt{SR}}}(n) &\leq \left(2\sqrt{ng} + \left(\frac{3n}{2}\right)^{2/3} + \left(\frac{3n}{2}\right)^{1/3} + 2g\right)\left(2\lceil \log n\rceil + \lceil \log |\Sigma|\rceil\right)\\ &= O\left((n^{2/3} + \sqrt{ng} + g)\log n\right).\end{aligned}$$

B.2 Bounded Window Case

Let $\mathtt{LZ77}_W^{\mathtt{SR}}$ be the LZ77 with self-referencing with bounded window size $W < n$. Unfortunately, in the bounded-window setting, one cannot apply the non-self-referencing proof verbatim. For block $B_i^x = [q_i^x, \ell_i^x, c_i^x]$, let $p_i^x := s_i^x - q_i^x$ denote the offset of the copy. Then the copied substring may have length larger than p_i^x, and the block may generate a long periodic substring. That is, the window size bounds the offset, i.e.,

$$1 \le p_i^x = s_i^x - q_i^x \le W,$$

but it does not necessarily bound the copied length ℓ_i^x.

We use the same internal-copy/non-internal-copy decomposition as in the bounded-window non-overlapping proof. The self-referencing variant affects the proof in two places.

(1) First, in the potential-drop argument (Lemma 3), for non-internal-copy type-$r \ge 3$ after-blocks, the potential drop loses one additional unit, i.e., a drop of at least $r-3$ instead of $r-2$. Besides the first and last translated intervals, one intermediate translated interval may enter the periodic tail generated after J. For internal-copy blocks, the original $(r-1)$ drop remains valid, since the relevant translated source-side span is contained in J and cannot enter the periodic tail.
(2) Second, the budget $\sum_{\ell \ge 0} x_{\ell,\mathsf{after}}^{w,\mathsf{non}}(\ell+1) \le 2W$ does not hold verbatim, because ℓ can be much larger than W. Instead, we remove the last non-internal-copy type-$r \ge 2$ after-block and apply a $2W$-budget to the remaining blocks.

Both losses affect only constants, and the final asymptotic bound remains the same as in the non-self-referencing bounded-window case.

B.2.1 The Potential Drop in the Presence of Periodic Tails. Let $B_i^w = [q_i^w, \ell_i^w, c_i^w] \in Y_{w,\mathsf{after}}$ be a self-referencing block that starts after the interval J. Let $p_i^w := s_i^w - q_i^w$. The only additional obstruction created by self-referencing arises when on the copied part of B_i^w, the string w is generated with period p_i^w. Now we establish the following claim.

Claim 11. *Let $B_i^w \in Y_{w,\mathsf{after}}$ be a self-referencing block, and let $p_i^w := s_i^w - q_i^w$. Suppose that a block $B_k^{w'}$ starts inside B_i^w at starting position $u := s_k^{w'}$. If*

$$u - p_i^w \ge j + g,$$

then $B_k^{w'}$ is the last block of $\mathtt{LZ77}_W^{\mathtt{SR}}(w')$ that starts inside B_i^w.

Proof. Suppose that $u - p_i^w \ge j + g$. Then for every $v \in [u, f_i^w - 1]$, both v and $v - p_i^w$ lie outside J. Since w and w' agree outside J, and since the copied part of B_i^w has period p_i^w, we have

$$w'[v] = w[v] = w[v - p_i^w] = w'[v - p_i^w] \qquad \text{for every } v \in [u, f_i^w - 1].$$

Thus, when constructing the block of w' starting at u, the self-referencing parser can copy from offset p_i^w at least until position $f_i^w - 1$, and then append the next symbol. Hence, this block must end at or after f_i^w by maximality of the chosen match. Therefore, no later block of $\mathtt{LZ77}_W^{\mathtt{SR}}(w')$ can start inside B_i^w. □

Claim 11 implies that, among the translated intervals corresponding to the blocks of $\mathtt{LZ77}_W^{\mathtt{SR}}(w')$ that start inside B_i^w, at most one non-last interval can enter the periodic tail after J. This is a new exception compared with the non-self-referencing proof.

Recall the truncated potential function we defined in Appendix A, i.e., for the window-reachable portion $J_W := \big[\max\{j, j+g-W\}, j+g-1\big]$ and for $K_W = \max\Big\{1, \sqrt{W/g}\Big\}$ and $H_W := 2W+j+g-1$,

$$\phi^w_{J_W}(t) := \Big|\big\{I \in \mathcal{I}_{K_W}(J_W) : \mathsf{Uncovered}(w, I, t) = 1\big\}\Big| + \frac{\max\{0, H_W - f^w_t\}}{K_W}.$$

As in the non-self-referencing bounded-window proof, if i^* is the first index such that the target interval of $B^w_{i^*}$ intersects J_W, i.e., $[s^w_{i^*}, f^w_{i^*}] \cap J_W \neq \varnothing$, then $\phi^w_{J_W}(i^*-1) \le 5\min\{W, \sqrt{Wg}\}$. Now, as we discussed in (1) above, we establish the following lemma.

Lemma 12. *For $\mathtt{LZ77}^{\mathtt{SR}}_W$ with bounded window $W < n$, let $B^w_i \in Y_{w,\mathsf{after}}$ be a type-r block.*

(1) If B^w_i is an internal-copy block and $r \ge 2$, then $\phi^w_{J_W}(i-1) - \phi^w_{J_W}(i) \ge r-1$.
(2) If B^w_i is not an internal-copy block and $r \ge 3$, then $\phi^w_{J_W}(i-1) - \phi^w_{J_W}(i) \ge r-3$.

Proof. (1) *(Internal-Copy Case.)* The internal-copy case is the same as in the non-self-referencing bounded-window analysis. Since B^w_i is an internal-copy block, we have $[q^w_i, q^w_i + \ell^w_i] \subseteq J$. Moreover, since $B^w_i \in Y_{w,\mathsf{after}}$, we have $s^w_i \ge j+g$, and since the offset is at most W, we have $q^w_i = s^w_i - p^w_i \ge j+g-W$. Thus, the translated intervals corresponding to the first $r-1$ blocks of $\mathtt{LZ77}^{\mathtt{SR}}_W(w')$ that start inside B^w_i are contained in J_W. Hence, they cannot enter the periodic tail after J, which is the only additional obstruction introduced by self-referencing. Hence, the same maximality argument (as in the proof of Lemma 3) holds: each of these $r-1$ translated intervals is not w-covered by time $i-1$ and becomes w-covered at time i. Otherwise, the corresponding block of $\mathtt{LZ77}^{\mathtt{SR}}_W(w')$ could be extended, contradicting maximality. If such an interval has length at most K_W, it decreases the first term of $\phi^w_{J_W}$ by one. If it has length larger than K_W, we charge it to the second term of the potential, exactly as in the non-self-referencing bounded-window proof. Hence, $\phi^w_{J_W}(i-1) - \phi^w_{J_W}(i) \ge r-1$.

(2) *(Non-Internal-Copy Case.)* Let $B^{w'}_{k_1}, \dots, B^{w'}_{k_r}$ be the blocks of $\mathtt{LZ77}^{\mathtt{SR}}_W(w')$ that start inside B^w_i, ordered by their starting positions. Let $p^w_i = s^w_i - q^w_i$. For each $a \in [r]$, shift the target interval of $B^{w'}_{k_a}$ by $-p^w_i$ and write $I_a := \Big[s^{w'}_{k_a} - p^w_i, f^{w'}_{k_a} - p^w_i\Big]$. As in the non-self-referencing potential-drop argument (Lemma 3), the first and last shifted intervals may meet the boundary of the relevant region and are not charged to the potential. In the self-referencing case, there is one additional possible exceptional interval: a middle shifted interval may cross the right boundary of J and enter the periodic tail after J.
We claim that there is *at most one* such additional exceptional interval. By Claim 11, if a block $B^{w'}_{k_a}$ with $a < r$ satisfies $s^{w'}_{k_a} - p^w_i \ge j+g$, then it must already be the last block starting inside B^w_i, a contradiction. Thus, every non-last shifted interval starts before $j+g$. Since the shifted intervals are ordered and pairwise disjoint, at most one of the middle shifted intervals can cross the right boundary of J.
After discarding the first and last shifted interval, and this possible crossing interval, at least $r-3$ shifted intervals remain, where each remaining interval is contained in J_W. Moreover, each such interval is uncovered by time $i-1$ and becomes covered at time i, following the same maximality argument in Lemma 3. As in the bounded-window proof, we split the remaining shifted intervals according to their length. If a remaining interval I has length at most K_W, then it decreases the first term of $\phi^w_{J_W}$ by one. If $|I| > K_W$, we charge it to the second term. The long

charged intervals are pairwise disjoint, and after translating them back to the target side, they lie inside the portion encoded by B_i^w before the horizon $2W + j + g - 1$. Hence, the decrease of the second term pays for all such long intervals. Therefore, each of the remaining $r-3$ intervals contributes one unit of potential drop. Hence, we can conclude that $\phi^w_{J_W}(i-1) - \phi^w_{J_W}(i) \geq r-3$. □

Lemma 13. *For* $\mathtt{LZ77}^{\mathsf{SR}}_W$ *with bounded window* $W < n$,

$$\sum_{r\geq 2}(r-1)y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r\geq 3}(r-3)\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) \leq g + 5\min\{W, \sqrt{Wg}\}.$$

Proof. Let i^* be the first index such that the target interval of $B^w_{i^*}$ intersects J_W, i.e., $[s^w_{i^*}, f^w_{i^*}] \cap J_W \neq \emptyset$. If no such index exists, then the claim follows from the same η_i-accounting as in Lemma 4. Hence, assume that such an index exists. By the bounded-window potential analysis from Lemma 8, we have $\phi^w_{J_W}(i^* - 1) \leq 5\min\{W, \sqrt{Wg}\}$. Now define the following sets:

- $A^{\mathsf{in}}_{\mathsf{after}} := \{i : B^w_i \in Y_{w,\mathsf{after}} \text{ is an internal-copy block of type } r_i \geq 2\}$,
- $A^{\mathsf{non}} := \{i : B^w_i \text{ has type } r_i \geq 3\} \setminus A^{\mathsf{in}}_{\mathsf{after}}$, and
- $A^{\mathsf{non}}_{\mathsf{after}} := A^{\mathsf{non}} \cap \{i : B^w_i \in Y_{w,\mathsf{after}}\}$.

We follow the proof of Lemma 8, with the following two modifications. First, for non-internal-copy blocks starting after J, we use Lemma 12 instead of Lemma 3, and therefore charge only $r_i - 3$ units rather than $r_i - 2$. Second, for the unique crossing block B^w_e, if it exists, we use the same after-suffix accounting as in Claim 5, but the after-suffix contributes $\max\{\rho_e - 2, 0\}$ charged intervals instead of $\max\{\rho_e - 1, 0\}$ since one additional intermediate interval may enter the periodic tail in the self-referencing case. Thus, the self-referencing bounded-window analogue of Claim 5 gives

$$\sum_{i\in A^{\mathsf{in}}_{\mathsf{after}}}(r_i - 1) + \sum_{i\in A^{\mathsf{non}}_{\mathsf{after}}}(r_i - 3) + \max\{\rho_e - 2, 0\} \leq \phi^w_{J_W}(i^* - 1) \leq 5\min\{W, \sqrt{Wg}\}. \tag{9}$$

Following the same strategy as Lemma 4 but with the update charge above, Equation (3) becomes

$$\sum_{r\geq 2}(r-1)y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r\geq 3}(r-3)\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) = \sum_{i\in A^{\mathsf{in}}_{\mathsf{after}}}(r_i - 1) + \sum_{i\in A^{\mathsf{non}}}(r_i - 3), \tag{10}$$

and Equation (4) becomes

$$\sum_{i\in A^{\mathsf{non}}}(r_i - 3) \leq \sum_{i\in A^{\mathsf{non}}_{\mathsf{after}}}(r_i - 3) + \max\{\rho_e - 2, 0\} + \sum_i \eta_i. \tag{11}$$

Combining Equation (9), Equation (10), and Equation (11),

$$\begin{aligned}
\sum_{r\geq 2}(r-1)y^{w,\mathsf{in}}_{r,\mathsf{after}} + \sum_{r\geq 3}(r-3)\left(y^w_r - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) &= \sum_{i\in A^{\mathsf{in}}_{\mathsf{after}}}(r_i - 1) + \sum_{i\in A^{\mathsf{non}}}(r_i - 3) \\
&\leq \sum_{i\in A^{\mathsf{in}}_{\mathsf{after}}}(r_i - 1) + \sum_{i\in A^{\mathsf{non}}_{\mathsf{after}}}(r_i - 3) + \max\{\rho_e - 2, 0\} + \sum_i \eta_i \\
&\leq 5\min\{W, \sqrt{Wg}\} + g.
\end{aligned}$$

□

B.2.2 Counting Non-Internal-Copy Type-≥ 2 After-Blocks. We next bound the number of non-internal-copy type-≥ 2 blocks in $Y_{w,\mathsf{after}}$. The only difference between the non-self-referencing bounded-window proof is that the copied length ℓ_i^w need not be at most W. Thus, we remove the last non-internal-copy type-≥ 2 block in $Y_{w,\mathsf{after}}$ before applying the length budget.

Lemma 14. *For* $\mathtt{LZ77}_W^{\mathtt{SR}}$ *with self-referencing with bounded window* $W < n$*,*

$$\sum_{r\geq 2}\left(y^w_{r,\mathsf{after}} - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) \leq (3W)^{2/3} + (3W)^{1/3} + 1.$$

Proof. If there is no non-internal-copy type-$r \geq 2$ block in $Y_{w,\mathsf{after}}$, there is nothing to prove. Otherwise, remove the last non-internal-copy type-$r \geq 2$ block in $Y_{w,\mathsf{after}}$. For each $\ell \geq 1$, let $x^{w,*,\mathsf{non}}_{\ell,\mathsf{after}}$ denote the number of remaining non-internal-copy type-$r \geq 2$ blocks in $Y_{w,\mathsf{after}}$ with copied length ℓ.

First, for every $\ell \geq 1$, $x^{w,*,\mathsf{non}}_{\ell,\mathsf{after}} \leq 2\ell$ as before. In particular, a counted non-internal-copy type-≥ 2 after-block has a shifted source-side interval that intersects J_W but is not contained in J_W. Hence, if its copied length is ℓ, its source position must lie within distance ℓ of one of the two boundaries of J_W. Thus, for fixed ℓ, its source position can lie in at most 2ℓ possible positions. By the source-pair uniqueness argument, no two such blocks with copied length ℓ have the same source position.

Second, after removing the last non-internal-copy type-$r \geq 2$ block in $Y_{w,\mathsf{after}}$, the same bounded-window budget as in Appendix A holds. Thus, we have $\sum_{\ell\geq 1} x^{w,*,\mathsf{non}}_{\ell,\mathsf{after}}(\ell+1) \leq 2W$. Applying the same threshold/swapping argument as in Lemma 5, with budget $2W$ in place of n, gives

$$\sum_{\ell\geq 1} x^{w,*,\mathsf{non}}_{\ell,\mathsf{after}} \leq (3W)^{2/3} + (3W)^{1/3}.$$

Adding back the removed last non-internal-copy type-$r \geq 2$ block increases the count by at most 1. Hence,

$$\sum_{r\geq 2}\left(y^w_{r,\mathsf{after}} - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) = 1 + \sum_{\ell\geq 1} x^{w,*,\mathsf{non}}_{\ell,\mathsf{after}} \leq (3W)^{2/3} + (3W)^{1/3} + 1. \qquad \square$$

B.2.3 Final Self-Referencing Bound. Taken the results of Section B.2.1 and Section B.2.2 altogether, we establish the following theorem.

Theorem 7. *Let* $\mathtt{LZ77}_W^{\mathtt{SR}} : \Sigma^* \to (\Sigma')^*$ *denote the self-referencing LZ77 compression function with sliding window size* $W < n$*. Then, for every* $g \geq 1$*,*

$$\begin{aligned}\mathtt{GS}^g_{\mathtt{LZ77}_W^{\mathtt{SR}}}(n) &\leq \left(3g + 5\min\{W, \sqrt{Wg}\} + 2(3W)^{2/3} + 2(3W)^{1/3} + 2\right)\left(2\lceil \log n\rceil + \lceil \log|\Sigma|\rceil\right)\\ &= O\left(\left(W^{2/3} + \sqrt{Wg} + g\right)\log n\right).\end{aligned}$$

Proof. By symmetry, it suffices to upper bound $t_{w'} - t_w$. Recall that $t_w = \sum_{r\geq 0} y^w_r$ and $t_{w'} = \sum_{r\geq 1} r\cdot y^w_r$. Therefore,

$$\begin{aligned}
t_{w'} - t_w &\le \sum_{r\ge 1}(r-1)y_r^w \\
&= \sum_{r\ge 1}(r-1)y_{r,\mathsf{after}}^{w,\mathsf{in}} + \sum_{r\ge 2}(r-1)\left(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) \\
&= \sum_{r\ge 2}(r-1)y_{r,\mathsf{after}}^{w,\mathsf{in}} + \sum_{r\ge 2}(r-3)\left(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) + 2\sum_{r\ge 2}\left(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) \\
&\le g + 5\min\{W, \sqrt{Wg}\} + 2\sum_{r\ge 2}\left(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right). \qquad \blacktriangleleft \text{ by Lemma 13}
\end{aligned}$$

Thus, it remains to upper bound $\sum_{r\ge 2}\left(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right)$. We observe that

$$\sum_{r\ge 2}\left(y_r^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) = \sum_{r\ge 2}\left(y_r^w - y_{r,\mathsf{after}}^{w}\right) + \sum_{r\ge 2}\left(y_{r,\mathsf{after}}^{w} - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right).$$

By Claim 7, we have $\sum_{r\ge 2}\left(y_r^w - y_{r,\mathsf{after}}^{w}\right) \le g$. And by Lemma 14, we have

$$\sum_{r\ge 2}\left(y_{r,\mathsf{after}}^{w} - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) \le (3W)^{2/3} + (3W)^{1/3} + 1.$$

Combining everything together, we have

$$t_{w'} - t_w \le 3g + 5\min\{W, \sqrt{Wg}\} + 2(3W)^{2/3} + 2(3W)^{1/3} + 2.$$

Finally, each LZ77 block can be encoded using at most $2\lceil \log n\rceil + \lceil \log|\Sigma|\rceil$ bits. Hence,

$$\begin{aligned}
\mathtt{GS}^g_{\mathtt{LZ77}^{\mathtt{SR}}_W}(n) &\le \left(3g + 5\min\{W, \sqrt{Wg}\} + 2(3W)^{2/3} + 2(3W)^{1/3} + 2\right)\left(2\lceil \log n\rceil + \lceil \log|\Sigma|\rceil\right) \\
&= O\left(\left(W^{2/3} + \sqrt{Wg} + g\right)\log n\right). \qquad \square
\end{aligned}$$

Reminder of Theorem 3. *Let* $\mathtt{LZ77}^{\mathtt{SR}}_W : \Sigma^* \to (\Sigma')^*$ *denote the self-referencing LZ77 compression function with sliding window size* $W \le n$*. Then, for every* $g \ge 1$*, the following bounds hold.*

If $W = n$*,*

$$\mathtt{GS}^g_{\mathtt{LZ77}^{\mathtt{SR}}_W}(n) \le \left(2\sqrt{ng} + \left(\frac{3n}{2}\right)^{2/3} + \left(\frac{3n}{2}\right)^{1/3} + 2g\right)\left(2\lceil \log n\rceil + \lceil \log|\Sigma|\rceil\right).$$

If $W < n$*,*

$$\mathtt{GS}^g_{\mathtt{LZ77}^{\mathtt{SR}}_W}(n) \le \left(3g + 5\min\{W, \sqrt{Wg}\} + 2(3W)^{2/3} + 2(3W)^{1/3} + 2\right)\left(2\lceil \log n\rceil + \lceil \log|\Sigma|\rceil\right).$$

In particular, for every $W \le n$, $\mathtt{GS}^g_{\mathtt{LZ77}^{\mathtt{SR}}_W}(n) \le O\left(\left(W^{2/3} + \sqrt{Wg} + g\right)\log n\right)$.

Proof. This follows immediately from Lemma 11 and Theorem 7. $\square$

C Missing Proofs

C.1 Missing Proofs from Section 4

Reminder of Claim 1. *If $B_i^w = [q_i^w, \ell_i^w, c_i^w] \in Y_{w,\mathsf{after}}$ is type-r block for $r \geq 2$, then $[q_i^w, q_i^w + \ell_i^w - 1] \cap J \neq \emptyset$.*

Proof. Suppose for contradiction that $[q_i^w, q_i^w + \ell_i^w - 1] \cap J = \emptyset$. Then, since two strings are identical outside J,

$$\begin{aligned} w[s_i^w, f_i^w - 1] &= w[q_i^w, q_i^w + \ell_i^w - 1] \\ &= w'[q_i^w, q_i^w + \ell_i^w - 1] = w'[s_i^w, f_i^w - 1]. \end{aligned} \tag{12}$$

Let k be the smallest element in $\mathcal{M}_i^{w,w'}$, i.e., $s_i^w \leq s_k^{w'} \leq f_i^w$ but $s_{k-1}^{w'} < s_i^w$. If $s_k^{w'} = f_i^w$, then it is clear that $\mathcal{M}_i^{w,w'} = \{k\}$ and B_i^w becomes a type-1 block, which is a contradiction. Hence, we have $s_k^{w'} < f_i^w$. Then, from Equation (12) and the fact that $f_i^w = s_i^w + \ell_i^w$, we observe that the substring $w'[s_i^w, f_i^w - 1]$ can be obtained by shifting the substring $w'[q_i^w, q_i^w + \ell_i^w - 1]$ by $s_i^w - q_i^w$. Hence,

$$\begin{aligned} w'[s_k^{w'}, f_i^w - 1] &= w'[s_k^{w'}, s_i^w + \ell_i^w - 1] \\ &= w'[s_k^{w'} - (s_i^w - q_i^w), s_i^w + \ell_i^w - 1 - (s_i^w - q_i^w)] \\ &= w'[q_i^w + (s_k^{w'} - s_i^w), q_i^w + \ell_i^w - 1]. \end{aligned}$$

This implies that $f_k^{w'} \geq f_i^w$. If there is another $k' \in \mathcal{M}_i^{w,w'}$ with $k' \neq k$, then by choice of k, we have $k < k'$ and $f_i^w \leq f_k^{w'} < s_{k'}^{w'}$. Hence, $|\mathcal{M}_i^{w,w'}| = 1$ and B_i^w becomes a type-1 block, which is a contradiction. Hence, $[q_i^w, q_i^w + \ell_i^w - 1] \cap J \neq \emptyset$. □

Reminder of Claim 4. *For every $m = 2, \ldots, r-1$, $\mathsf{Uncovered}(w, I_m, i-1) = 1$, i.e., I_m is not w-covered by time $i-1$.*

Proof. Suppose for contradiction that there exists some $m' \in \{2, \ldots, r-1\}$ such that $\mathsf{Uncovered}(w, I_{m'}, i-1) = 0$. Then there exists a block $B_{i'}^w = \left[q_{i'}^w, \ell_{i'}^w, c_{i'}^w\right] \in Y_{w,\mathsf{after}}$ for some $i' < i$ such that $I_{m'} \subseteq [q_{i'}^w, q_{i'}^w + \ell_{i'}^w - 1]$. Recall that $I_{m'} = \left[s_{k_{m'}}^{w'} - \delta_i, f_{k_{m'}}^{w'} - \delta_i\right]$. Shifting this back to the target interval of $B_{i'}^w$, we have

$$\begin{aligned} w\left[s_{k_{m'}}^{w'} - \delta_i, f_{k_{m'}}^{w'} - \delta_i\right] &= w\left[s_{k_{m'}}^{w'} - \delta_i + \delta_{i'}, f_{k_{m'}}^{w'} - \delta_i + \delta_{i'}\right] \\ &= w'\left[s_{k_{m'}}^{w'} - \delta_i + \delta_{i'}, f_{k_{m'}}^{w'} - \delta_i + \delta_{i'}\right], \end{aligned} \tag{13}$$

where Equation (13) holds since $B_{i'}^w \in Y_{w,\mathsf{after}}$.

On the other hand, shifting the same interval back to the target interval of B_i^w, we have

$$\begin{aligned} w\left[s_{k_{m'}}^{w'} - \delta_i, f_{k_{m'}}^{w'} - \delta_i\right] &= w\left[s_{k_{m'}}^{w'}, f_{k_{m'}}^{w'}\right] \\ &= w'\left[s_{k_{m'}}^{w'}, f_{k_{m'}}^{w'}\right], \end{aligned} \tag{14}$$

and similarly, Equation (14) holds since $B_i^w \in Y_{w,\mathsf{after}}$. However, combining Equation (13) and Equation (14) yields

$$w'\left[s_{k_{m'}}^{w'} - \delta_i + \delta_{i'}, f_{k_{m'}}^{w'} - \delta_i + \delta_{i'}\right] = w'\left[s_{k_{m'}}^{w'}, f_{k_{m'}}^{w'}\right],$$

which implies that $B_{k_{m'}}^{w'}$ indeed would have copied

$$w'\left[s_{k_{m'}}^{w'} - \delta_i + \delta_{i'}, f_{k_{m'}}^{w'} - \delta_i + \delta_{i'}\right]$$

to $w'\left[s_{k_{m'}}^{w'}, f_{k_{m'}}^{w'}\right]$ and append a new character to generate the block. Hence, $w'\left[s_{k_{m'}}^{w'}, f_{k_{m'}}^{w'}\right]$ is a substring of $w'\left[1, s_{k_{m'}}^{w'} - 1\right]$. This implies that the copied substring used to construct $B_{k_{m'}}^{w'}$ is *not maximal*: the LZ77 algorithm could copy one more symbol to the right, so $B_{k_{m'}}^{w'}$ would have to end strictly after $f_{k_{m'}}^{w'}$. This contradiction shows that I_m cannot be w-covered by time $i-1$ for each $m = 2, \ldots, r-1$. □

Reminder of Claim 6. *Let $i_1 < i_2$. If $B_{i_1}^w = [q_{i_1}^w, \ell_{i_1}^w, c_{i_1}^w] \in Y_{w,\mathsf{after}}$ is a type-r_1 block and $B_{i_2}^w = [q_{i_2}^w, \ell_{i_2}^w, c_{i_2}^w] \in Y_{w,\mathsf{after}}$ is a type-r_2 block for both $r_1, r_2 \geq 2$. Then $(q_{i_1}^w, \ell_{i_1}^w) \neq (q_{i_2}^w, \ell_{i_2}^w)$.*

Proof. The proof follows largely the same strategy as [BLG25, Claim 17]. Suppose for contradiction that there exists $i_1 < i_2$ such that $B_{i_1}^w \in Y_{w,\mathsf{after}}$ is type-r_1 block and $B_{i_2}^w \in Y_{w,\mathsf{after}}$ is a type-r_2 block for $r_1, r_2 \geq 2$ but $(q_{i_1}^w, \ell_{i_1}^w) = (q_{i_2}^w, \ell_{i_2}^w)$. Since $B_{i_1}^w, B_{i_2}^w \in Y_{w,\mathsf{after}}$, we have $j + g \leq s_{i_1}^w \leq f_{i_1}^w < s_{i_2}^w$, meaning that two strings are identical after J. Hence,

$$\begin{aligned} w'[s_{i_1}^w, f_{i_1}^w - 1] &= w[s_{i_1}^w, f_{i_1}^w - 1] = w[q_{i_1}^w, q_{i_1}^w + \ell_{i_1}^w - 1] \\ &= w[q_{i_2}^w, q_{i_2}^w + \ell_{i_2}^w - 1] = w[s_{i_2}^w, f_{i_2}^w - 1] = w'[s_{i_2}^w, f_{i_2}^w - 1]. \end{aligned} \tag{15}$$

Let k be the smallest element in $\mathcal{M}_{i_2}^{w,w'}$, i.e., $s_{i_2}^w \leq s_k^{w'} \leq f_{i_2}^w$ but $s_{k-1}^{w'} < s_{i_2}^w$. If $s_k^{w'} = f_{i_2}^w$, then it is clear that $\mathcal{M}_{i_2}^{w,w'} = \{k\}$ and $B_{i_2}^w$ becomes a type-1 block, which is a contradiction. Hence, we have $s_k^{w'} < f_{i_2}^w$. Then, from Equation (15),

$$\begin{aligned} w'[s_k^{w'}, f_{i_2}^w - 1] &= w'[s_{i_2}^w + (s_k^{w'} - s_{i_2}^w), f_{i_2}^w - 1] \\ &= w'[s_{i_1}^w + (s_k^{w'} - s_{i_2}^w), f_{i_1}^w - 1]. \end{aligned} \tag{16}$$

This implies that $f_k^{w'} \geq f_{i_2}^w$. To see this, suppose for contradiction that $f_k^{w'} < f_{i_2}^w$. By definition of LZ77, $w'[s_k^{w'}, f_k^{w'} - 1]$ is the longest substring of $w'[1, s_k^{w'} - 1]$ starting at $s_k^{w'}$. However, Equation (16) implies that $w'[s_k^{w'}, f_{i_1}^w - 1]$ is also a substring of $w'[1, s_k^{w'} - 1]$ which is *longer than* $w'[s_k^{w'}, f_k^{w'} - 1]$, which is a contradiction.

If there is another $k' \in \mathcal{M}_{i_2}^{w,w'}$ such that $k' \neq k$, then by choice of k, we have $k < k'$ and therefore $s_{k'}^{w'} > f_k^{w'} \geq f_{i_2}^w$. This contradicts the definition of $\mathcal{M}_{i_2}^{w,w'}$ since $B_{k'}^{w'}$ does not start inside $B_{i_2}^w$. It follows that $|\mathcal{M}_{i_2}^{w,w'}| = 1$, which is a contradiction. Hence, $(q_{i_1}^w, \ell_{i_1}^w) \neq (q_{i_2}^w, \ell_{i_2}^w)$. □

Reminder of Claim 8. $\sum_{r\geq 2}\left(y_{r,\mathsf{after}}^w - y_{r,\mathsf{after}}^{w,\mathsf{in}}\right) \leq \left(\frac{3n}{2}\right)^{2/3} + \left(\frac{3n}{2}\right)^{1/3}$.

Proof. Since a block with $\ell = 0$ encodes only one symbol, it can contain at most one start position of a block of $\mathtt{LZ77}(w')$. Hence, no block of copied length 0 can be of type $r \geq 2$. Therefore, we may sum over $\ell \geq 1$. For each $\ell \geq 1$, let

$$S^{w,\mathsf{non}}_{\ell,\mathsf{after}} := \{i : B^w_i = [q^w_i, \ell^w_i = \ell, c^w_i] \in Y_{w,\mathsf{after}}, \\ B^w_i \text{ is of type } r \geq 2 \text{ and } B^w_i \text{ is not an internal-copy block}\},$$

and let $x^{w,\mathsf{non}}_{\ell,\mathsf{after}} := |S^{w,\mathsf{non}}_{\ell,\mathsf{after}}|$. Then the total number of non-internal-copy type-$r(\geq 2)$ blocks in $Y_{w,\mathsf{after}}$ is given by the sum $\sum_{\ell\geq 1} x^{w,\mathsf{non}}_{\ell,\mathsf{after}}$. Rearranging the sum by summing over $r \geq 2$, we have

$$\sum_{r\geq 2}\left(y^w_{r,\mathsf{after}} - y^{w,\mathsf{in}}_{r,\mathsf{after}}\right) = \sum_{\ell\geq 1} x^{w,\mathsf{non}}_{\ell,\mathsf{after}}. \tag{17}$$

Thus, it remains to upper bound $\sum_{\ell\geq 1} x^{w,\mathsf{non}}_{\ell,\mathsf{after}}$. At this point, the argument is analogous to the final optimization step in the proof of [BLG25, Lemma 2]. There, one maximizes $\sum_{\ell\geq 0} x^w_{\ell,\mathsf{after}}$ subject to the constraints

(1) $x^w_{\ell,\mathsf{after}} \leq \ell$, and
(2) $\sum_{\ell\geq 0} x^w_{\ell,\mathsf{after}}(\ell+1) \leq n$.

In our setting, constraint (1) is replaced by $x^{w,\mathsf{non}}_{\ell,\mathsf{after}} \leq 2\ell$ as shown in Claim 12 below, and constraint (2) is replaced by $\sum_{\ell\geq 1} x^{w,\mathsf{non}}_{\ell,\mathsf{after}}(\ell+1) \leq n$.

Claim 12. *For each $\ell \geq 1$, $x^{w,\mathsf{non}}_{\ell,\mathsf{after}} \leq 2\ell$.*

Proof of Claim 12. Recall that $x^{w,\mathsf{non}}_{\ell,\mathsf{after}} := |S^{w,\mathsf{non}}_{\ell,\mathsf{after}}|$. If $i \in S^{w,\mathsf{non}}_{\ell,\mathsf{after}}$, then by Claim 1, the source interval $[q^w_i, q^w_i + \ell - 1]$ must intersect $J = [j, j+g-1]$, i.e., $j - \ell + 1 \leq q^w_i \leq j + g - 1$. Furthermore, since such a block is not an internal-copy block, q^w_i must lie in one of the two boundary ranges $[j-\ell+1, j-1]$ or $[j+g-\ell, j+g-1]$. Hence, the starting position q^w_i can take at most 2ℓ distinct values. On the other hand, by Claim 6, if $i_1, i_2 \in S^{w,\mathsf{non}}_{\ell,\mathsf{after}} \subseteq S^w_{\ell,\mathsf{after}}$ are distinct, then $(q^w_{i_1}, \ell) \neq (q^w_{i_2}, \ell)$, i.e., $q^w_{i_1} \neq q^w_{i_2}$. Therefore, the map $i \mapsto q^w_i$ is injective on $S^{w,\mathsf{non}}_{\ell,\mathsf{after}}$, and thus $x^{w,\mathsf{non}}_{\ell,\mathsf{after}} = |S^{w,\mathsf{non}}_{\ell,\mathsf{after}}| \leq 2\ell$. ∎

The same swapping argument therefore implies that the optimal solution is obtained by filling the smallest lengths first, i.e., there exists a threshold z such that $x^{w,\mathsf{non}}_{\ell,\mathsf{after}} = 2\ell$ for all $\ell < z$, $x^{w,\mathsf{non}}_{z,\mathsf{after}} \in (0, 2z]$, and $x^{w,\mathsf{non}}_{\ell,\mathsf{after}} = 0$ for all $\ell > z$. To identify the threshold z, we first observe that

$$\left[\sum_{\ell=1}^{z-1} 2\ell(\ell+1)\right] + x^{w,\mathsf{non}}_{z,\mathsf{after}}(z+1) = \sum_{\ell\geq 1} x^{w,\mathsf{non}}_{\ell,\mathsf{after}}(\ell+1) \leq n.$$

by our second constraint. Simplifying the sum, we have

$$\begin{aligned} n &\geq \left[\sum_{\ell=1}^{z-1} 2\ell(\ell+1)\right] + x^{w,\mathsf{non}}_{z,\mathsf{after}}(z+1) \\ &\geq \left[\sum_{\ell=1}^{z-1} 2\ell(\ell+1)\right] + (z+1) && \blacktriangleleft \text{ since } x^{w,\mathsf{non}}_{z,\mathsf{after}} \geq 1 \end{aligned}$$

$$= \frac{2}{3}(z-1)z(z+1) + (z+1) \geq \frac{2}{3}z^3. \tag{18}$$

From Equation (18), we have $z \leq \left(\frac{3n}{2}\right)^{1/3}$. Setting this value of z, we have

$$\sum_{\ell \geq 1} x^{w,\text{non}}_{\ell,\text{after}} \leq \sum_{\ell=1}^{\left(\frac{3n}{2}\right)^{1/3}} 2\ell = \left(\frac{3n}{2}\right)^{2/3} + \left(\frac{3n}{2}\right)^{1/3}. \tag{19}$$

By Equation (17) and Equation (19), we get the desired result. □

C.2 Missing Proofs from Section 5

Reminder of Lemma 7. *Let $w, w' \in \Sigma^N$ be the large-alphabet strings from Theorem 5, and let $\Delta := |\#\mathtt{LZ77}_W(w) - \#\mathtt{LZ77}_W(w')|$ denote the block-count gap, where $\#\mathtt{LZ77}_W(\cdot)$ denotes the number of blocks. Let $\widehat{w} = P_{\mathcal{D}} \circ w$ and $\widehat{w}' = P_{\mathcal{D}} \circ w'$ be the prefixed strings defined above. Then there exist binary strings $z := \Phi(\widehat{w})$ and $z' := \Phi(\widehat{w}')$ of length $n = hN^*$, with edit length $g = hg_0$ and window size $W = h(W_0 + M + 1) = \Theta(hW_0)$, such that $z \overset{g}{\sim} z'$ and*

$$\left|\left|\mathtt{LZ77}_W(z)\right| - \left|\mathtt{LZ77}_W(z')\right|\right| = \Omega\big(\Delta \cdot \log n\big).$$

Proof. Let $\varphi : \Sigma \cup \{d_\star\} \to \{0,1\}^h$ be a *good encoding* from Lemma 6, and set $z = \Phi(\widehat{w})$ and $z' = \Phi(\widehat{w}')$. By construction, $|z| = |z'| = n = hN^*$, $g = hg_0$, and $W = h(W_0 + M + 1)$. Moreover, since the prefix $P_{\mathcal{D}}$ is common and lies before the edited interval, if w and w' differ only on a symbol interval of length g_0, then z and z' differ only on the corresponding binary interval of length g. Thus, $z \overset{g}{\sim} z'$.

It remains to compare the number of LZ77 blocks between z and z'. In each construction, there are $\Theta(\Delta)$ split points in the parsing of w' at which the parser is forced to begin a new block, while the corresponding portion of w can be parsed without this extra split. In the construction of Theorem 5, these forced splits occur inside the gadgets $w[i, i+r-1] \circ d_{r,i}$, where the parsing of w uses one block for the whole gadget, while the parsing of w' must split before the delimiter $d_{r,i}$.

The dictionary prefix removes first-occurrence effects for the delimiter codewords. Since each $d_{r,i} \in \mathcal{D}$ appears in $P_{\mathcal{D}}$, each later occurrence of $\varphi(d_{r,i})$ can be encoded using one LZ77 block from the prefix. The final dummy delimiter $d_\star$ appears only in the common prefix and prevents a symbol-level match from crossing the prefix/body boundary. The prefix is identical in z and z', so it contributes the same number of blocks to both parsings and does not affect the block-count gap.

Let $\mathcal{S}$ be the set of these forced split points in w'. Then $|\mathcal{S}| = \Theta(\Delta)$. For each $s \in \mathcal{S}$, the following forced-split property holds:

> *There do not exist intervals I_1, I_2 in w' such that $\min I_2 < s \leq \max I_2$, I_1 occurs before I_2, and $w'[I_1] = w'[I_2]$.*

In Theorem 5, it follows from the freshness of the delimiter $d_{r,i}$.

We now pass these forced splits to the binary encoding. Let $B := 4$. First, discard from $\mathcal{S}$ the forced split points whose distance from the endpoints of the corresponding large-alphabet block L_r is less than B symbols. Then, take a maximal subset $\mathcal{S}' \subseteq \mathcal{S}$ such that $|s - s'| \geq 2B + 1$ for all distinct $s, s' \in \mathcal{S}'$. Since any symbol interval of length $2B+1$ contains only $O(B) = O(1)$ split points, this greedy selection loses only a constant factor. Therefore, $|\mathcal{S}'| = \Omega(|\mathcal{S}|) = \Omega(\Delta)$.

Now fix $s \in \mathcal{S}'$. In the prefixed string $\widehat{w}'$, the split point s occurs at symbol position $M+1+s$. Let $s_b := (M+s)h+1$ be the corresponding binary codeword boundary in z'. We claim that the binary parsing of z' must contain a block boundary in the interval $[s_b - Bh, s_b + Bh]$. Suppose not. Then, some block in the parsing of z' contains the entire interval $[s_b - Bh, s_b + Bh]$. Let the target interval copied by this block be $T = [p, p+L-1]$ and let its source interval be $Q = [q, q+L-1]$ so that $z'[p, p+L-1] = z'[q, q+L-1]$. Since $[s_b - Bh, s_b + Bh] \subseteq T$, we have $L \geq 2Bh \geq 2h$. Then by Lemma 6, there exist intervals I_1, I_2 in $\widehat{w}'$ such that $\widehat{w}'[I_1] = \widehat{w}'[I_2]$ and $|I_1| = |I_2| \geq \left\lfloor \frac{L}{h} \right\rfloor - 1$.

Moreover, because the binary block crosses the Bh-neighborhood of s_b and $B = 4 > 1$, the one-codeword loss allowed by Lemma 6 still leaves an induced symbol-level target interval crossing the split point $M+1+s$ in $\widehat{w}'$. Thus, one of the intervals I_1 and I_2 is a target interval crossing $M+1+s$, and the other is an earlier source interval equal to it.

We now remove the common prefix shift. Since the target interval crosses $M+1+s$, its corresponding body interval crosses the split point s in w'. If the source interval also lies in the body part of $\widehat{w}'$, shifting both intervals by $-(M+1)$ gives two equal intervals in w', with the target interval crossing s, contradicting the forced-split property of s.

It remains to rule out source intervals that meet the prefix. If the source interval lies entirely inside the dictionary prefix, then it consists only of delimiter symbols and possibly the dummy delimiter $d_\star$, whereas the target interval crossing s contains a non-delimiter symbol immediately before the delimiter $d_{r,i}$. If the source interval crosses the boundary between the prefix and the body, then it contains $d_\star$, while the target interval lies entirely inside the body and contains no $d_\star$. Both cases are impossible. Therefore, we obtain a contradiction in all cases.

On the other hand, the forced split points in $\mathcal{S}'$ do not create a comparable number of additional binary block boundaries in z. The copied symbol intervals used by the large-alphabet parsing of w lift to aligned binary copies in z, and whenever a delimiter codeword $\varphi(d_{r,i})$ is needed, it can be encoded by one block using its earlier occurrence in the common dictionary prefix. Thus, the binary parsing of z follows the large-alphabet parsing of w with only a constant-factor overhead per gadget, whereas the forced split points above force $\Omega(\Delta)$ additional boundaries in z'. Consequently,

$$|\#\mathtt{LZ77}_W(z) - \#\mathtt{LZ77}_W(z')| = \Omega(\Delta).$$

Finally, each binary LZ77 block is represented using $2\lceil \log n \rceil + \log 2 = \Theta(\log n)$ bits. This completes the proof. □