

Small Hazard-free Transducers

Johannes Bund¹, Christoph Lenzen¹, and Moti Medina^{*2}

¹Max Planck Institute for Informatics, Saarland Informatics Campus

²Department of Electrical & Computer Engineering, Ben Gurion University of the Negev

November 30, 2018

Abstract

Recently, an unconditional exponential separation between the hazard-free complexity and (standard) circuit complexity of explicit functions has been shown [IKL⁺18]. This raises the question: which classes of functions permit efficient hazard-free circuits?

Our main result is as follows. A *transducer* is a finite state machine that transcribes, symbol by symbol, an input string of length n into an output string of length n . We prove that any function arising from a transducer with s states, that is input symbols which are encoded by ℓ bits, has a hazard-free circuit of size $2^{\mathcal{O}(s+\ell)} \cdot n$ and depth $\mathcal{O}(\ell + s \cdot \log n)$; in particular, if $s, \ell \in \mathcal{O}(1)$, size and depth are asymptotically optimal.

We utilize our main result to derive efficient circuits for *k-recoverable addition*. Informally speaking, a code is *k-recoverable* if it does not increase uncertainty regarding the encoded value, so long as it is guaranteed that it is from $\{x, x + 1, \dots, x + k\}$ for some $x \in \mathbb{N}_0$. We provide an asymptotically optimal *k-recoverable* code. We also realize a transducer with $\mathcal{O}(k)$ states that adds two codewords from this *k-recoverable* code. Combined with our main result, we obtain a hazard-free adder circuit of size $2^{\mathcal{O}(k)}n$ and depth $\mathcal{O}(k \log n)$ with respect to this code, i.e., a *k-recoverable* adder circuit that adds two codewords of n bits each. In other words, *k-recoverable* addition is fixed-parameter tractable with respect to k .

* Author email addresses: jbund@mpi-inf.mpg.de, clenzen@mpi-inf.mpg.de, medinamo@bgu.ac.il

1 Introduction

In this work, we study a classic extension of Boolean logic due to Kleene [Kle52, §64] that allows for the presence of *unstable* (i.e., transitioning, oscillating, unspecified, etc.) signals, which are represented by a third logical value \mathbf{u} . Referring to the Boolean values $\mathbb{B} := \{0, 1\}$ as *stable*, in Kleene logic the basic gates¹ output a stable value if and only if the stable inputs already determine this output:

or	0	1	\mathbf{u}	and	0	1	\mathbf{u}	not	0	1
0	0	0	0	0	0	1	\mathbf{u}	0	1	1
1	0	1	\mathbf{u}	1	1	1	1	1	0	0
\mathbf{u}	0	\mathbf{u}	\mathbf{u}	\mathbf{u}	\mathbf{u}	1	\mathbf{u}	\mathbf{u}	\mathbf{u}	\mathbf{u}

By induction over the circuit structure, this defines for any circuit C consisting of such gates the function $C: \mathbb{T}^n \rightarrow \mathbb{T}^m$ it implements, where $\mathbb{T} := \{0, 1, \mathbf{u}\}$. A main feature differentiating Kleene logic from Boolean logic is that the law of the excluded middle does not hold: $\mathbf{or}(x, \bar{x}) \neq 1$ and $\mathbf{and}(x, \bar{x}) \neq 0$. This critically distinguishes \mathbf{u} from an unknown Boolean value, and it explains why we allow for constant-0 (and constant-1) gates in addition to **or**, **and**, and **not**.

A circuit has a hazard, if, unlike basic gates, it deviates from the best possible behavior in face of unstable inputs.

Definition 1.1 (Resolution and Hazards). *For $x \in \mathbb{T}^n$, its resolutions are*

$$\text{res}(x) := \{y \in \mathbb{B}^n \mid \forall i \in \{1, \dots, n\}: x_i \neq \mathbf{u} \Rightarrow y_i = x_i\},$$

i.e., the set of binary strings that can be obtained from x by replacing each \mathbf{u} by either 0 or 1.

A circuit C on n inputs with m outputs has a hazard at input $x \in \mathbb{T}^n$ (and output bit $i \in \{1, \dots, m\}$), iff $C_i(x) = \mathbf{u}$, yet there is a bit $b \in \mathbb{B}$ such that $C_i(y) = b$ for all $y \in \text{res}(x)$. It is a k -bit hazard, iff the number of \mathbf{u} s in x is at most $k \in \mathbb{N}$.

In contrast, a circuit is (k -bit) *hazard-free*² iff it has no (k -bit) hazards. An already mentioned trivial example for a circuit with a hazard is $\mathbf{and}(x, \bar{x})$, as it implements the constant Boolean function $f(x) = 0$, but $\mathbf{and}(\mathbf{u}, \bar{\mathbf{u}}) = \mathbf{u}$; a hazard-free implementation would use a constant-0 gate, ignoring the input to the circuit. The smallest non-trivial example is a multiplexer, which has the Boolean specification $\text{MUX}(a, b, 0) = a$ and $\text{MUX}(a, b, 1) = b$, see, e.g., [IKL⁺18].

The importance of this fundamental concept is illustrated by its surfacing in areas as diverse as digital circuit design [Huf57, Cal58, FFL18], logic [Kle52, Kör66], and cybersecurity [TWM⁺09, HOI⁺12]. Accordingly, the design of hazard-free circuits has received significant attention over the years, cf. Section 2. Early on, Huffman established that all Boolean functions admit a hazard-free circuit implementing them [Huf57]. However, recently unconditional exponential lower bounds for explicit functions have been shown [IKL⁺18], including functions that admit hazardous circuits of polynomial size.

This naturally raises the question for which classes of functions small hazard-free circuits exist. But which classes are of particular interest? In [FFL18], it is proposed to use hazard-free circuits in the design of digital controllers for “analog” (i.e., continuously-valued) actuating variables. A topological argument [Mar81] proves that in this setting metastability (i.e., instability of circuit outputs that may last arbitrarily long) cannot be avoided deterministically. Hazard-free circuits open up the possibility to avoid the delay incurred by waiting for metastability to decay with

¹The specific choice of basic gates does not matter, see [IKL⁺18]; accordingly, we stick with **and**, **or**, and **not** here.

²The term hazard-free has varying meanings in the literature; we follow the definition from [DDT78].

sufficiently high probability, enabling the control loop to counteract accumulating errors quicker; ultimately, this results in better performance, despite going from a probabilistic to a deterministic correctness guarantee.

A major open question posed in [FFL18] is whether efficient circuits implementing the proposed approach exist. As a promising example, a positive answer has been provided for sorting [BLM18]. However, arguably more arithmetic operations are required to make the concept practical: we need hazard-free circuits for addition, multiplication, division, etc; in fact, besides sorting, division by 2 and addition are the operations missing to fully implement the computations of the clock synchronization module proposed in [FFL18] by a digital circuit.

In a seminal article, Ladner and Fischer [LF80] showed that any function that can be realized by a constant-size *finite-state transducer* can be implemented by a circuit of linear size and logarithmic depth. Addition and, by extension, multiplication and division by constants, fall into this class of functions. Interestingly, the sorting circuit presented in [BLM18] relies on a special case of this framework, suggesting the possibility of a generalization to hazard-free circuits. Given these considerations and the central role addition plays in arithmetics, this work sets out to study hazard-free circuits for the class of functions arising from such transducers.

Transducers

A transducer is a finite state machine transcribing an input to an output string. We phrase our results for *Moore machines* [Moo56], but our techniques are not specific to this type of transducer.

Definition 1.2 (Moore Machine). *A Moore machine is a 6-tuple $T = (S, s_0, \Sigma, \Lambda, t, o)$, where*

- S is the finite set of states,
- $s_0 \in S$ is the starting state,
- Σ is the input alphabet,
- Λ is the output alphabet,
- $t: S \times \Sigma \rightarrow S$ is the state transition function, and
- $o: S \times \Sigma \rightarrow \Lambda$ is the output function.

For $n \in \mathbb{N}$, Moore machine T gives rise to the transcription function $\tau_{T,n}: \Sigma^n \rightarrow \Lambda^n$ in the following way. Define for $i \in \{1, \dots, n\}$ and $x \in \Sigma^n$ the state s_i after i steps inductively via $s_i := t(s_{i-1}, x_i)$. Then $\tau_{T,n}(x)_i := o(s_{i-1}, x_i)$. In the following, we will omit T and n from the notation whenever clear from context, i.e., we simply write τ instead of $\tau_{T,n}$.

Unless we impose restrictions, any function $f: \mathbb{B}^n \rightarrow \mathbb{B}$ can (essentially) be realized by a transcription function. For instance, even when restricting to $\Sigma = \Lambda = \mathbb{B}$, the state machine can simply memorize the entire input string x and output $f(x)$ on reception of the last input bit. However, this requires $|S|$ to be exponential in n , i.e., the state machine is very large.

Accordingly, it is of interest to consider *small* transducers. In particular, essential basic operations, like addition, max, and min, can be implemented by constant-size transducers. However, there is more to the task than providing hazard-free circuits for such operations.

Encoding Matters!

In the Boolean world, the choice of encoding may influence the complexity of an operation, but not its precision. In face of hazards, however, the encoding may cause unnecessary loss of information.

To illustrate this point, we describe the desired behavior first. Suppose that we are given two k -preserving inputs in n -bit unary encoding, which means that input $x \in \mathbb{T}^n$ satisfies that $\text{res}(x)$

contains at most $k + 1$ codewords, all of which are consecutive. Such inputs are of the form $1^*u^*0^*$, where the number of u 's is at most k and the length of the string is n . The following definition is helpful in expressing how such inputs are constructed.

Definition 1.3 (The $*$ -operator). *For $x, y \in \mathbb{T}$, $x * y = x$ if $x = y$ and $x * y = u$ otherwise. We extend this associative and commutative operator to strings $x, y \in \mathbb{T}^n$ by applying it bit-wise.*

Denoting by γ_u the unary encoding function, we get that inputs are of the form $*_{\ell=i}^j \gamma_u(\ell) = *_{\ell=i}^j 1^\ell 0^{n-\ell} = 1^i u^{j-i} 0^{n-j}$, where $i, j \in [n + 1]$ and $j - i \leq k$. We can obtain a $(2n)$ -bit unary encoding of their sum that treats the u s in the “right” way by sorting the $2n$ input bits with respect to the order $0 \prec u \prec 1$. This is straightforward, as `and` and `or` implement `max` and `min` with respect to this order, meaning that they together realize a \prec -comparator. Feeding inputs $1^i u^{j-i} 0^{n-j}$ and $1^{i'} u^{j'-i'} 0^{n-j'}$ into a sorting network for $2n$ inputs using such comparators outputs $1^{i+i'} u^{j+j'-i-i'} 0^{2n-j-j'} = *_{\ell=i+i'}^{j+j'} \gamma_u(\ell)$.

Note that this addition maintains information to the best possible degree: The codewords that are resolutions of these strings represent the numbers $\{i, \dots, j\}$, $\{i', \dots, j'\}$, and $\{i + i', \dots, j + j'\}$, respectively. In other words, the uncertainty in the output regarding which value is represented matches the uncertainty of the inputs. The reader may think of this as a faithful representation of interval addition.

The above unary encoding has another crucial property. We can consistently map any *resolution* of the output back to a codeword. We extend γ_u^{-1} (the decoding function that is defined on $\gamma_u([n+1])$) to \mathbb{B}^n as follows: For $x \in \mathbb{B}^n$, let $i \in [n+1]$ be maximal such that $x_j = 1$ for all $j \in [i+1]$. Then $\gamma_u^{-1}(x) := 1^i 0^{n-i}$. Observe that this guarantees that $\gamma_u^{-1}(\text{res}(*_{\ell=i}^j \gamma(\ell))) = \{i, \dots, j\}$, i.e., after stabilization we can recover a value from the range that was represented by the string $1^i u^{j-i} 0^{n-j}$ using γ_u^{-1} . This property is very similar to the ability of error correction codes to recover the correct codeword in face of bit flips; here, we recover a “correct” value in the sense that it is consistent with the information that was available. We refer to this as the code being *recoverable*. We will see that this is both highly useful and the best we can expect.

Naturally, the above positive example is of limited interest, as it is extremely inefficient due to the exponential overhead of unary encoding. So, what happens if we use a non-redundant encoding? It is not hard to see that standard binary encoding results in disaster. For instance, denoting by $B_n: [2^n] \rightarrow \mathbb{B}^n$ an n -bit code, we have that $B_4(7) = 0111$ and $B_4(0) * B_4(1) = 000u$; for this pair of strings, it is unclear whether their sum should be 7 or 8. In contrast, $B_4(7 + 0) * B_4(7 + 1) = 0111 * 1000 = uuuu$, i.e., the resulting string holds no information whatsoever. This is an extreme example, in which (even given all other bits) a single input bit affects all output bits, which in Kleene logic entails that setting this bit to u implies that all outputs are u .³ In other words, the code provides no non-trivial guarantees in face of uncertainty at its input.

The above examples show that we should aim for codes that behave similarly to unary encoding in face of unstable inputs, have a high rate,⁴ and yet permit efficient circuits for arithmetic and other basic computations.

Our Contribution

Our contribution is twofold. First, on the conceptual level, we propose a meaningful addition in Kleene logic, providing and motivating the necessary key definitions, and demonstrating that they are realizable. We formalize the properties of unary encoding discussed above in the next definition.

³This can be seen by induction on the circuit size, where the base case is given by the behavior of the basic gates.

⁴For encoding function $\gamma: [2^m] \rightarrow \mathbb{B}^\ell$, the rate is the ratio $m/\ell \leq 1$.

Definition 4.1 (*k*-preserving & recoverable Codes). A code is an injective function $\gamma: [M] \rightarrow \mathbb{B}^n$ (where $[M] := \{0, \dots, M-1\}$), whose image is the set of codewords. By γ^{-1} we denote the decoding function, i.e., the inverse of γ on its image. For range $r = \langle i, i+k \rangle_M$, where $\langle i, i+k \rangle_M$ denotes $k+1$ consecutive integers modulo M starting at i , define $x_r := * \gamma(r)$ as the corresponding extended codeword and define its range as $r_{x_r} := r$.

- (i) For extended codeword x , $p_x := |r_x| - 1$ is its imprecision (i.e., codewords are 0-imprecise).
- (ii) Let $k \in \mathbb{N}_0$. Code γ is *k*-preserving, iff for all $p \leq k$ and *p*-imprecise extended codewords x it holds that $\gamma([M]) \cap \text{res}(x) = \gamma(r_x)$. That is, $\text{res}(x)$ does not contain “new” codewords which are not in $\gamma(r_x)$.
- (iii) Let $k \in \mathbb{N}_0$. Code γ is *k*-recoverable, iff there is an extension of γ^{-1} to \mathbb{B}^n (i.e., the entire codomain of γ) such that for all extended codewords x of imprecision $p_x \leq k$ and all $y \in \text{res}(x)$ it holds that $\gamma^{-1}(y) \in r_x$.

We prove that any *k*-recoverable code requires at least *k* redundant bits.

Theorem 4.2. For $k, n \in \mathbb{N}$, no code $\gamma: [(k+1) \cdot 2^n] \rightarrow \mathbb{B}^{n+k}$ is *p*-recoverable for $p > k$.

We give a simple and natural code that requires *k* redundant bits in order to be *k*-preserving.

Lemma 4.13. $\gamma_{n,k}$ is *k*-preserving.

The same code γ can be made *k*-recoverable by using $2k - 1$ redundant bits, this means that the obtained code is asymptotically optimal, that is, its rate is $1 - O\left(\frac{1}{n}\right)$.

Theorem 4.15. $\gamma_{n,k}: [(k+1) \cdot 2^{n-k}] \rightarrow \mathbb{B}^n$ is $\lceil k/2 \rceil$ -recoverable.

We show that using this code comes at no loss in computational power compared to (inefficient) unary encoding, so long as the limits of the uncertainty the code can tolerate are not exceeded. This is formalized in the following lemma.

Lemma 4.3. Let $f: [M] \rightarrow \mathbb{B}^m$. For any *k*-recoverable code $\gamma: [M] \rightarrow \mathbb{B}^n$, there is a combinational circuit C satisfying for all $p \leq k$ and *p*-imprecise codewords x that

$$C(x) = \underset{i \in r_x}{*} f(i).$$

We also establish that efficient operations can be performed on this code, most prominently addition. In short, we propose a foundation for arithmetic computations in Kleene logic. We begin by defining addition in this context.

Definition 5.1 (*k*-recoverable Addition). For code $\gamma: [M] \rightarrow \mathbb{B}^n$, we define γ -addition in the natural way: For codewords $x, y \in \gamma[M]$, $x +_\gamma y = \gamma(\gamma^{-1}(x) + \gamma^{-1}(y) \bmod M)$. If γ is *k*-recoverable for $k \in \mathbb{N}$, we extend this definition to all $x, y \in \mathbb{B}^n$ by the (not necessarily unique) extension of γ^{-1} to \mathbb{B}^n *k*-recoverability guarantees. A circuit implements *k*-recoverable addition w.r.t. *k*-recoverable code γ , iff it implements (the extended) operator $+_\gamma$ on extended codewords x and y satisfying that $p_x + p_y \leq k$ without hazards.

We also obtain the following hazard free adder.

Theorem 5.11. There is a circuit of size $2^{\mathcal{O}(k)}n$ and depth $\mathcal{O}(k \log n)$ that implements $\lceil k/2 \rceil$ -recoverable addition on $\gamma_{n,k}$.

Most of the insights mentioned above can be shown using straightforward arguments; once the pieces are in place, a general picture emerges naturally. We view our contribution in this regard primarily in posing the right questions and distilling the appropriate definitions. An exception is a key ingredient to our machinery for addition, which we consider to be the main technical challenge and second main contribution of this work.

Theorem 3.13. *For any $\ell, n \in \mathbb{N}$ and Moore machine $T = (S, s_0, \Sigma \subseteq \mathbb{B}^\ell, \Lambda, t, o)$, there is a hazard-free circuit of size $\mathcal{O}((2^{2\ell+|S|} + 2^{3|S|})n)$ and depth $\mathcal{O}(\ell + |S| \log n)$ implementing $\tau_{T,n}$.*

We remark that this result stands out against the lower bound from [IKL⁺18], which proves an exponential dependence of the circuit size on n for any general construction of hazard-free circuits. While the above theorem incurs exponential overheads in terms of the size of the transducer, the dependence on n is asymptotically optimal. Thus, for constant size transducers, we obtain asymptotically optimal hazard-free implementations of their transcription functions, both with respect to size and depth.

Organization of this Article

We discuss related work in Sec. 2. In Sec. 3, we present our generic construction of hazard-free circuits for transcription functions, i.e., we prove Thm. 3.13. Subsequently, in Sec. 4, we define k -preserving and k -recoverable codes. We provide a k -recoverable code and prove that its rate is asymptotically optimal.⁵ Moreover, we formalize the notion of “an uncertainty of k in an input” and show that a k -recoverable code allows for implementing any function such that this uncertainty is preserved; this serves as a strong indicator that requiring k -recoverability (for a suitable k) is a good choice. This result is flanked by some additional observations further supporting the view that the proposed concepts may serve as a solid basis for arithmetic computations in Kleene logic. Next, Sec. 5 utilizes the tools developed in Sec. 3 and 4 to construct asymptotically small circuits for k -recoverable addition.

2 Related Work

There are a variety of 3-valued logics in use, but relevant to this work is only Kleene’s “strong logic of indeterminacy” [Kle52, §64] which he published as early as 1938 [Kle38]. Over time, it has been utilized in a large variety of contexts.

Metastability in Electronic Circuits. Early works studied relay and switching networks, but due to the underlying abstraction provided by Kleene logic, there is no qualitative difference to modern electronic circuits. To our knowledge, Goto [Got49] was the first to publish a hazard-free multiplexer. Huffman [Huf57] provided the first general construction of hazard-free circuits.

By the late 70’s, there was an intense debate among hardware developers whether the problem of *metastability* can be dealt with deterministically, by suitable design of circuits [Pec76, Wor77, Mar77, SC79]. Metastability denotes an unstable third equilibrium state of bistable elements, such as latches and flip-flops. Metastability can lead to late output transitions of such elements, conflicting interpretations of the stored value, or even metastability of downstream circuit components, breaking the abstraction of Boolean logic. Marino [Mar81] resolved this question by giving a topological argument that metastability cannot be avoided in general. Two main strategies are in use to handle this problem. The first is based on timing guarantees, which is infeasible when

⁵The rate is actually $1 - \mathcal{O}(\frac{1}{n})$.

communicating across boundaries of different clock domains or when taking and digitally storing measurements of continuously-valued variables. In such cases, so-called *synchronizers* are employed, see, e.g., [Kin08, Chap. 2]. Synchronizers trade time for decreased probability of ongoing metastability (and thus resulting errors).

Another option applicable in some cases connects this discussion to hazards and Kleene logic: *logical masking* (c.f. [SKK⁺02]). By this term, hardware designers refer to the property of basic gates to output a stable value if the stable inputs determine this value. Put differently, hazard-free circuits are “maximally logically masking.” It may hence surprise that the respective literature does not mention Huffman’s work, suggesting that it has been overlooked due to different terminology.

Metastability-Containing Control Loops. Friedrichs et al. [FFL18] proposed the use of hazard-free (or “metastability-containing”) circuits in mixed-signal control loops. Mixed-signal circuits combine both analog and digital elements, which here means that a continuously-valued (i.e., analog) actuating variable is controlled by a digital circuit. This comes with the advantage of standardized, small, and in particular fast digital components, but by Marino’s result mentioned earlier [Mar81], using digital components necessarily incurs the risk of metastability.

While [FFL18] proves the approach to be feasible in principle, it leaves open the question whether *efficient* solutions exist. In light of this, [FFL18] points out that Gray codes offer imprecision-1, too, providing hope that small circuits are a possibility. In [FKLP17], it is shown that imprecision-1 Gray code time measurements can be obtained by practical circuits. Astonishingly, the measurement circuit produces the same k -recoverable codes we give in this work, as an intermediate result before conversion to imprecision-1 Gray code. These codes emerge naturally in this context for a completely different reason, suggesting that they are of general significance.

Complexity of Hazard-free Circuits. In [IKL⁺18] it was shown that for monotone functions, their *hazard-free complexity* (i.e., the size of the smallest hazard-free implementation) equals their monotone complexity (i.e., the size of the smallest implementation without negation gates). This yields a number of unconditional lower bounds as corollaries of results on monotone circuits. In particular, an exponential separation between hazard-free and standard circuit complexity follows from [AB87, Tar88], and the naive monotone circuit of cubic size for Boolean matrix multiplication is optimal [Pat75, MG76]. These lower bounds are complemented by a general construction yielding circuits of size $n^{\mathcal{O}(k)}|C|$ without k -bit hazards, where C is an arbitrary circuit implementing the desired function. Thus, for constant k , the overhead for removing k -bit hazards is polynomial in n . The above separation result implies that an overhead of $2^{k^{\Omega(1)}}$ is necessary, but it remains open whether the task is fixed-parameter tractable w.r.t. k .

In contrast, some functions and specific hazards admit much more efficient solutions. If the possible positions of \mathbf{u} inputs are restricted to index set I , a construction based on hazard-free multiplexers avoids the respective hazards with a circuit of size $\mathcal{O}(2^{|I|}|C|)$ [IKL⁺18, Lemma 5.2], where C is as above. This can be seen as combining speculative computing [TY12, TYM14] with hazard-free multiplexers. Another example is given by sorting of imprecision-1 Gray code inputs, which admit asymptotically optimal sorting networks [BLM18]. This is the special case of $k = 1$ for the k -recoverable code we present in Section 4; we show how to generalize the result to arbitrary k .

It is worth noting that the lower bound can be circumvented using non-combinational logic [FFL18]. Using clocked circuits and so-called masking registers, k -bit hazards can be eliminated with factor $\mathcal{O}(k)$ -overhead. Masking registers also strictly increase the computational power of the system with each clock cycle. However, in this work we consider combinational logic only.

Parallel Prefix Computation. The result from [BLM18] is based on an application of a special case of the parallel prefix computation (PPC) framework by Ladner and Fischer [LF80]. They observed that any transcription function can be efficiently computed as follows:

1. For an encoding of the space of functions from S to itself, encode for input string $x \in \Sigma^n$ the function $t_i := t(\cdot, x_i): S \rightarrow S$ for each $i \in \{1, \dots, n\}$.
2. Compute for each i (the encoding of) the function $\pi_i := t_i \circ t_{i-1} \circ \dots \circ t_1$.
3. Compute $s_i = \pi_i(s_0)$ for each i .
4. Compute the output $o(s_{i-1}, x_i)$ for each i .

For a constant-size Moore machine, all but the second step can be done in parallel for each i by circuits of size $\mathcal{O}(n)$ and depth $\mathcal{O}(1)$. The second step can be performed by a circuit of size $\mathcal{O}(n)$ and depth $\mathcal{O}(\log n)$ exploiting associativity of function composition. This celebrated result yields the only asymptotically optimal adder constructions known to date, cf. [SL15].

In [BLM18], it was possible to exploit highly convenient circumstances. First, it held that $S = \Sigma = \mathbb{B}^2$ and $t: \mathbb{B}^2 \times \mathbb{B}^2 \rightarrow \mathbb{B}^2$ turned out to be associative. Second, also the extension of $t: \mathbb{T}^2 \times \mathbb{T}^2 \rightarrow \mathbb{T}^2$ induced by a hazard-free circuit implementing t was associative.⁶ One can show in this setting that this operator can be used directly, yielding a very efficient solution that avoids hazards that would arise from the multi-step process laid out above.

Error Correcting and Snake-in-the-Box Codes. The properties we require from our codes are reminiscent of error detection and correction codes (see, e.g., [MS77]). Given an imprecision- p codeword x for $p \leq k$, a k -preserving code γ is sufficiently strong to guarantee that any $y \in \text{res}(x)$ is either recognized as a non-codeword or is in $\gamma(r_x)$. The latter is “correct” in the sense that this codeword is contained in the range of values that x represents. This corresponds to error detection; e.g., a (redundant) parity bit enables noticing, but not fixing, single-bit errors.

The stronger property of k -recoverability is akin to error correction. For a recoverable code, we can map any $y \in \text{res}(x)$ back to a codeword from $\gamma(r_x)$ via $\gamma(\gamma^{-1}(y))$. The matching property of an error correction code is the ability to recover the original codeword after up to k bit flips.

The requirements to our codes are less restrictive than for error detection and correction codes. Error detection and correction necessitate Hamming distance $k + 1$ and $2k + 1$, respectively, while a k -recoverable code may have codewords in Hamming distance one. This holds true even for codewords that are “far away” from each other, i.e., $\gamma(i)$ and $\gamma(j)$ for $|j - i| \gg k$. This explains why our codes are also less restrictive than (generalized) snake-in-the-box codes. Generalizing the definition from [Kau58], a distance- k snake-in-the-box code satisfies that the Hamming distance between $\gamma(i)$ and $\gamma(j)$ is at least $\min\{k + 1, |j - i|\}$, while consecutive codewords differ in a single bit only. With this definition, a distance- k snake-in-the-box code is k -preserving and $\lfloor k/2 \rfloor$ -recoverable, respectively. In contrast, our code is not even a distance-1 snake-in-the-box.

For comparison, an n -bit code that can detect errors after up to k bit flips can by the (tight) Hamming bound [MS77] achieve rate at most $(n - \ell)/n$, where $\ell := \log(\sum_{i=0}^t \binom{n}{i})$ with $t := \lfloor k/2 \rfloor$. For a distance- k snake-in-the-box code, a fraction of $1/(k + 1)$ of its codewords must be mutually in Hamming distance at least $k + 1$, so by the same argument its rate is at most $(n - \ell - \log(k + 1))/n$. In contrast, we show that for k -recoverable codes $\ell \leq 2k - 1$, regardless of n .

Generality of the Model. Kleene logic has been invented and reinvented many times in different contexts. The work by Goto [Got49] was almost certainly independent of Kleene’s [Kle38] and remained unnoticed by the western world for decades. Likewise, Huffman’s work [Huf57] makes

⁶This is not true in general, even if the original function is associative. A counter-example is given by addition (modulo 4) of 2-bit standard binary code: $(01 + 01) + 0u = 1u \neq uu = 01 + (01 + 0u)$.

no mention of Kleene logic, developing its own terms. The authors of [FFL18] were unaware of the connection to these articles while working on early versions of the manuscript, learning later that they reinvented the wheel, too. Similarly, related works on cybersecurity [TWM⁺09, HOI⁺12] appear to derive from coming up with the concept independently again, although the work by Hu et al. [HOI⁺12] draws the connection to Huffman via the influential work of Eichelberger on detecting hazards [Eic65]. See [BEI01] for a survey covering some of these articles and discussing different logics that also cover “dynamic hazards.”⁷

In our opinion, all of this goes to show that the questions we study in this paper are fundamental and of widespread interest. This suggests that our results are of interest to other areas as well.

3 Hazard-free Implementation of Transcription Functions

The PPC framework by Ladner and Fischer [LF80], in its most general form, computes transcription functions of Moore machine $(S, s_0, \Sigma, \Lambda, t, o)$ as follows. It identifies⁸ $s \in S$ with the (column) unit vector $e_s \in \mathbb{B}^{|S|}$, where

$$(e_s)_{s'} := \begin{cases} 1 & \text{if } s = s' \\ 0 & \text{else.} \end{cases}$$

Thus, functions $f: S \rightarrow S$ are represented by Boolean matrices $B_f \in \mathbb{B}^{|S|} \times \mathbb{B}^{|S|}$ in the natural way, i.e., $(B_f)_{\cdot s} = e_s$. We get that $e_{(f \circ g)(s)} = B_f \cdot B_g \cdot e_s$, where \cdot denotes the (associative) Boolean matrix multiplication map. In particular, for all $i \in [n + 1]$, it holds that $e_{s_i} = \left(\prod_{j=1}^i B_j \right) \cdot e_{s_0}$.

If $|S|$ is constant, there is a constant-sized circuit implementing \cdot , and exploiting the associativity of the operator, Ladner and Fischer give an $\mathcal{O}(\log n)$ -depth $\mathcal{O}(n)$ -size circuit computing e_{s_i} for all i from x . Evaluating $o(s_{i-1}, x_i)$ in parallel for all $i \in \{1, \dots, n\}$ then yields the output.

Our goal in this section is to adapt this approach such that we obtain hazard-free implementations of transcription functions. To simplify notation, we introduce the following definition.

Definition 3.1 (Hazard-free Extensions). *For function $f: \mathbb{B}^n \rightarrow \mathbb{B}^m$, denote by $f_u: \mathbb{T}^n \rightarrow \mathbb{T}^m$ its hazard-free extension defined by $f_u(x) := \ast_{y \in \text{res}(x)} f(y)$.*

Note that circuit C is a hazard-free implementation of $f: \mathbb{B}^n \rightarrow \mathbb{B}^m$ iff $C(x) = f_u(x)$ for all $x \in \mathbb{T}^n$. In other words, our goal can be rephrased as implementing τ_u for the given Moore machine.

3.1 Naïve Approach

At first glance, applying the same pattern to derive a hazard-free implementation of transcription functions looks promising. First, using the above encoding and reinterpreting input symbols as Boolean matrices “works,” in the sense that the resulting matrices over $(\mathbb{T}, \text{or}, \text{and})$ faithfully represent the hazard-free extensions of the respective functions on elements of S .

Observation 3.2. *Denote by F the space of functions from S to itself and let $E: \mathbb{B}^\ell \rightarrow F$ be arbitrary. Then we have for all $x \in \mathbb{T}^\ell$ and $s \in S$ that $(B_E \cdot e_s)_u(x) = (\ast_{y \in \text{res}(x)} B_{E(y)}) \cdot_u e_s$.*

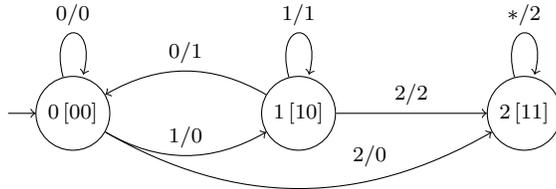
Proof. The proof follows by using Definition 3.1 twice, and since $\text{res}(s) = \{s\}$.

$$(B_E \cdot e_s)_u(x) = \ast_{\substack{y \in \text{res}(x) \\ s' \in \text{res}(s)}} B_{E(y)} \cdot e_{s'} = \ast_{y \in \text{res}(x)} B_{E(y)} \cdot e_s = \left(\ast_{y \in \text{res}(x)} B_{E(y)} \right) \cdot_u e_s. \quad \square$$

⁷What we refer to as hazards in our paper is called “static hazards” elsewhere, as for (static) hazards the output is supposed to undergo no change.

⁸This is a choice of encoding, which is implicit in the description in [LF80].

Figure 1: A toy example. This transducer simply outputs its state with a delay of one step. The set of states, input alphabet, and output alphabet are the same, i.e., $S = \Sigma = \Lambda \subseteq \mathbb{B}^2$, These sets are encoded via standard unary encoding.



Second, we can apply the machinery for associative operators developed in [LF80].

Observation 3.3 (\cdot_u is associative). *For all $A \in \mathbb{T}^{\alpha \times \beta}$, $B \in \mathbb{T}^{\beta \times \gamma}$, and $C \in \mathbb{T}^{\gamma \times \delta}$, we have that $(A \cdot_u B) \cdot_u C = A \cdot_u (B \cdot_u C)$.*

Proof. As **or** and **and** are associative also on \mathbb{T} , this follows by the same straightforward calculation as for matrices over arbitrary (semi)rings.⁹ \square

Finally, we know that \cdot_u can be implemented by a circuit. In fact, it is straightforward to do so with a polynomially-sized circuit.

Corollary 3.4 (of [IKL⁺18, Lemma 4.2]). *There is a circuit of size less than $2\alpha\beta\gamma$ and depth $1 + \lceil \log \beta \rceil$ that computes $A \cdot_u B$ for matrices $A \in \mathbb{T}^{\alpha \times \beta}$ and $B \in \mathbb{T}^{\beta \times \gamma}$.*

Proof. The standard algorithm for Boolean matrix multiplication is monotone, i.e., does not use negations, and requires for each of the $\alpha\gamma$ entries of $A \cdot_u B$ a binary tree of β **and** gates (the leaves) and $\beta - 1$ **or** gates; monotone circuits are hazard-free. \square

So where do things go wrong? The culprit is the encoding of restricted transition functions, as we can see from the toy example given in Figure 1. There, $\Sigma = S = \Lambda \subseteq \mathbb{B}^2$ encoded via $\gamma(i) = 1^i 0^{2-i}$ on domain $[3]$, $s_0 = \gamma(0)$, and $o(\gamma(i), \gamma(j)) = \gamma(i)$, i.e., the transducer simply outputs its state with a delay of one step. Note that γ is standard unary encoding, so one might expect it to enable resilience to hazards.

Nonetheless, given input $x = (\gamma(0) * \gamma(1))\gamma(0)\gamma(2)\gamma(0)$, we have that $*_{y_1 \in \text{res}(x_1)} B_{t(\cdot, y_1)} \cdot e_{s_0} = (\mathbf{u}, \mathbf{u}, 0)^T$. As $(0, 0, 0)^T \in \text{res}(\mathbf{u}, \mathbf{u}, 0)^T$, but the vector e_s representing state s contains a 1 for any $s \in S$. As the all-0 vector is mapped to itself by any matrix, it follows that continuing the computation can never lead to a vector with an entry of 1. This is bad news, as $*_{y \in \text{res}(x)} \tau(y) = \gamma(0)(\gamma(0) * \gamma(1))\gamma(0)\gamma(2)$ — the second input makes certain that we are in state $\gamma(0)$ again, and this is reproduced by the output. Note that even when going to state $\gamma(2)$, the encoding does not reflect this, despite the fact that an input symbol $\gamma(2)$ results in state $\gamma(2)$ no matter what.

The underlying issue is that the chosen function encoding is not able to capture that after processing a resolution the first input symbol ($\gamma(0) * \gamma(1)$), the state machine is *certainly* in either state $\gamma(0)$ or $\gamma(1)$. However, there is no certainty regarding either of the individual states, so the chosen encoding cannot reflect this; as $(0, 0, 0)^T$ represents no state whatsoever, we have no way to recover the correct output symbols later.¹⁰ To overcome the obstacle illustrated by this example, we need to add further redundancy to the encoding.

⁹Note that $(\mathbb{T}, \mathbf{or}, \mathbf{and})$ is only a (commutative) semiring, as its “addition,” i.e., **or**, has no inverses.

¹⁰One may interpret $(0, 0, 0)^T$ as a single other state, but this static assignment cannot always be correct.

A Universal Encoding for Functions

For Boolean $f: \mathbb{B}^n \rightarrow \mathbb{B}^m$, we obtain $f_u(x)$ by mapping each $y \in \text{res}(x)$ using f and then taking the $*$ operation over the resulting set. The latter may —depending on x and the encoding— lose information, as $\text{res}(f_u(x))$ may be a strict superset of $f(\text{res}(x))$. This becomes problematic when we subsequently apply some function g that is constant on $f(\text{res}(x))$, but not on $\text{res}(f_u(x))$; we then get that $\mathbf{u} = g_u(f_u(x)) \neq (g \circ f)_u(x) \in \mathbb{B}$.

The key idea underlying our approach is to maintain the information that f maps $\text{res}(x)$ to $f(\text{res}(x))$ in the encoding, by storing for each $A \subseteq \mathbb{B}^n$ and $B \subseteq \mathbb{B}^m$ whether $f(A) \subseteq B$. When concatenating functions, we then can retrieve the information that $g \circ f$ is constant on $\text{res}(x)$ and use the resulting stable 1 to “suppress” the uncertainty the above simpler encoding cannot handle using an **or** operation.

Definition 3.5 (Universal Function Encoding). *Given function $f: S \rightarrow T$, define*

$$\forall A \subseteq S, B \subseteq T: (\mathcal{M}_f)_{BA} := \begin{cases} 1 & \text{if } f(A) \subseteq B \\ 0 & \text{else.} \end{cases}$$

Thus, \mathcal{M}_f is a Boolean $2^{|T|} \times 2^{|S|}$ matrix. Moreover, for $s \in S$ and $A \subseteq S$, define $e_A^{(s)}$ via $e_A^{(s)} := 1$ if $s \in A$ and $e_A^{(s)} := 0$ else. Hence, for all $B \subseteq T$ we have that $(\mathcal{M}_f \cdot e^{(s)})_B = 1$ iff $f(s) \in B$.

To see how this encoding helps, consider the example from Figure 1 once more, with the same input string. The matrix representation of $f := t_u(\cdot, x_1) = t(\cdot, \gamma(0)) * t(\cdot, \gamma(1))$ has an entry of $1 = 1 * 1$ at row index $B = \{0, 1\}$ and column index $A = \{0\}$, as both $t(\cdot, \gamma(0))$ and $t(\cdot, \gamma(1))$ map 0 to an element of B . Moreover, $g := t_u(\cdot, x_2) = t(\cdot, \gamma(0))$ maps B back to A , implying that the respective matrix has a 1 at row index A and column index B . A simplified version of the matrices can be seen in Figure 2. When multiplying the matrices, the resulting matrix has at index pair (A, A) the “sum” (i.e., the **or**) over all sets C of the “product” (i.e., the **and**) of whether f maps A to C and also g maps C back to A . As the **or** is stable when having at least one stable 1 as input due to the above observation on B , we “learn” that $g \circ f$ maps the initial state 0 to itself, i.e., the product has a stable 1 at index pair (A, A) .

Similar considerations show that we get stable entries of 0 whenever it is *certain* that some set is not mapped to a subset of another. This is captured by the following theorem, which is the our stepping stone to proving Theorem 3.13.

Theorem 3.6. *Let $f_i: S \rightarrow T$ for all $i \in [\ell]$, $g_j: T \rightarrow U$ for all $j \in [\ell']$, $A \subseteq S$, and $C \subseteq U$. Then*

$$\left(\left(\begin{matrix} * \\ i \in [\ell'] \end{matrix} \mathcal{M}_{g_i} \right) \cdot_u \left(\begin{matrix} * \\ j \in [\ell] \end{matrix} \mathcal{M}_{f_j} \right) \right)_{CA} = \left(\begin{matrix} * \\ (i,j) \in [\ell'] \times [\ell] \end{matrix} \mathcal{M}_{g_j \circ f_i} \right)_{CA}.$$

Theorem 3.6 implies that we can decompose computation of $\tau_u(x)$ into mapping each symbol x_i to $*_{y_i \in \text{res}(x_i)} \mathcal{M}_{t(\cdot, y_i)}$, multiplying the resulting matrices, and then using the result to determine the output symbols from the starting state and the input symbols. By Observation 3.3, we can exploit the associativity of the matrix product to obtain an efficient circuit in the vein of [LF80], yielding Theorem 3.13. The remainder of this section is dedicated to proving the above theorem and this conclusion.

We first show that, also for this encoding, function composition maps to matrix multiplication.

Lemma 3.7. *For functions $f: S \rightarrow T$, $g: T \rightarrow U$, it holds that $\mathcal{M}_g \cdot \mathcal{M}_f = \mathcal{M}_{g \circ f}$.*

and, analogously,

$$\left(G^{(1)} \cdot F^{(1)}\right)_{ij} \geq (G'' \cdot F'')_{ij} = 1.$$

Together, this implies that

$$\left(* \left\{G^{(0)} \cdot F^{(0)}, G^{(1)} \cdot F^{(1)}\right\}\right)_{ij} = * \left\{\left(G^{(0)} \cdot F^{(0)}\right)_{ij}, \left(G^{(1)} \cdot F^{(1)}\right)_{ij}\right\} = *\{0, 1\} = \mathbf{u}. \quad \square$$

Equipped with these technical statements, we are now in position to prove the key technical result of this section.

Theorem 3.6. *Let $f_i: S \rightarrow T$ for all $i \in [\ell]$, $g_j: T \rightarrow U$ for all $j \in [\ell']$, $A \subseteq S$, and $C \subseteq U$. Then*

$$\left(\left(\left(*_{i \in [\ell']} \mathcal{M}_{g_i}\right) \cdot_{\mathbf{u}} \left(*_{j \in [\ell']} \mathcal{M}_{f_i}\right)\right)_{CA} = \left(*_{(i,j) \in [\ell] \times [\ell']} \mathcal{M}_{g_j \circ f_i}\right)_{CA}.$$

Proof. By Lemma 3.7, for each i and j we have that $\mathcal{M}_{g_j \circ f_i} = \mathcal{M}_{g_j} \cdot \mathcal{M}_{f_i}$. Observe that $*A \preceq *B$ for $A \subseteq B$, where \preceq is defined by the partial order $b \prec \mathbf{u}$ for $b \in \mathbb{B}$. Moreover, $X \subseteq \text{res}(*X)$ for any set X . Thus,

$$\begin{aligned} \left(\left(\left(*_{i \in [\ell']} \mathcal{M}_{g_i}\right) \cdot_{\mathbf{u}} \left(*_{j \in [\ell']} \mathcal{M}_{f_i}\right)\right)_{CA} &= * \left(\text{res} \left(*_{i \in [\ell']} \mathcal{M}_{g_i}\right) \cdot \text{res} \left(*_{j \in [\ell']} \mathcal{M}_{f_i}\right)\right)_{CA} \\ &\succeq \left(*_{(i,j) \in [\ell] \times [\ell']} \mathcal{M}_{g_j} \cdot \mathcal{M}_{f_i}\right)_{CA} \\ &= \left(*_{(i,j) \in [\ell] \times [\ell']} \mathcal{M}_{g_j \circ f_i}\right)_{CA}. \end{aligned}$$

the claimed equality follows if the l.h.s. equals $b \in \{0, 1\}$. Hence, assume that the l.h.s. equals \mathbf{u} .

We apply Lemma 3.9 with $G = *_{i \in [\ell']} \mathcal{M}_{g_i}$ and $F = *_{j \in [\ell']} \mathcal{M}_{f_i}$, yielding that

$$\left(G^{(0)} \cdot F^{(0)} = 0\right) \wedge \left(G^{(1)} \cdot F^{(1)} = 1\right),$$

which by Definition 3.8 is equivalent to

$$\forall B \subseteq T \exists (i, j) \in [1, \ell] \times [1, \ell'] : (\mathcal{M}_{g_j})_{CB} = 0 \vee (\mathcal{M}_{f_i})_{BA} = 0 \quad (3.10)$$

$$\exists B_1 \subseteq T \exists (i, j) \in [1, \ell] \times [1, \ell'] : (\mathcal{M}_{g_j})_{CB_1} = (\mathcal{M}_{f_i})_{B_1A} = 1. \quad (3.11)$$

Set $B_0 := \bigcup_{i \in [\ell]} f_i(A)$. As thus $f_i(A) \subseteq B_0$ by construction, it holds that $(\mathcal{M}_{f_i})_{B_0A} = 1$ for all i . Equation (3.10) thus entails that

$$\exists j \in [\ell'] : (\mathcal{M}_{g_j})_{CB_0} = 0 \Leftrightarrow g_j(B_0) \not\subseteq C.$$

Hence, there is an $x \in B_0$ satisfying that $g_{j_0}(x) \notin C$ for a $j_0 \in [\ell']$. By construction, $x \in f_{i_0}(A)$ for some $i_0 \in [\ell]$, yielding that $(g_{j_0} \circ f_{i_0})(A) = g_{j_0}(f_{i_0}(A)) \not\subseteq C$. We conclude that $\left(\mathcal{M}_{g_{j_0} \circ f_{i_0}}\right)_{CA} = 0$.

Now consider Equation (3.11), which says that there are indices $i_1 \in [\ell]$ and $j_1 \in [\ell']$ such that $g_{j_1}(B_1) \subseteq C$ and $f_{i_1}(A) \subseteq B_1$. This immediately yields that $(g_{j_1} \circ f_{i_1})(A) \subseteq C$ and thus $\left(\mathcal{M}_{g_{j_1} \circ f_{i_1}}\right)_{CA} = 1$. Thus

$$\left(*\{\mathcal{M}_{g_j \circ f_i} \mid i \in [\ell] \wedge j \in [\ell']\}\right)_{CA} = * \left\{\left(\mathcal{M}_{g_j \circ f_i}\right)_{CA} \mid i \in [\ell] \wedge j \in [\ell']\right\} = *\{0, 1\} = \mathbf{u}. \quad \square$$

The following corollary readily follows by inductive application of Theorem 3.6.

Corollary 3.12. *Suppose for $i \in [n]$, we are given mappings $E_i: \mathbb{B}^{\ell_i} \rightarrow F_i$ to function spaces F_i , such that for all $i \in [n-1]$ the codomain of functions from F_i equals the domain of functions from F_{i+1} . Let E denote a function that maps $x \in \prod_{i=0}^{n-1} \mathbb{B}^{\ell_i}$ to $\circ_{i=0}^{n-1} E_i(x_i)$. Then, for all $x \in \prod_{i=0}^{n-1} \mathbb{T}^{\ell_i}$,*

$$(\mathcal{M}_E)_u(x) = (\mathcal{M}_{E_{n-1}})_u(x_{n-2}) \cdot_u (\mathcal{M}_{E_{n-1}})_u(x_{n-2}) \cdot_u \dots \cdot_u (\mathcal{M}_{E_0})_u(x_0).$$

Applying the corollary to the transducer's transcription function yields the decomposition we need to apply the framework from [LF80] to obtain an efficient circuit, i.e., prove Theorem 3.13.

Theorem 3.13. *For any $\ell, n \in \mathbb{N}$ and Moore machine $T = (S, s_0, \Sigma \subseteq \mathbb{B}^\ell, \Lambda, t, o)$, there is a hazard-free circuit of size $\mathcal{O}((2^{2\ell+|S|} + 2^{3|S|})n)$ and depth $\mathcal{O}(\ell + |S| \log n)$ implementing $\tau_{T,n}$.*

Proof. W.l.o.g., assume that $\Sigma = \mathbb{B}^\ell$; otherwise fix an arbitrary extension of t to the domain $S \times \mathbb{B}^\ell$. Moreover, assume for now that $\Lambda = \mathbb{B}$; the generalization is straightforward and follows at the end of the proof. Thus, for input $x \in \Sigma^n$, the i -th output symbol of a hazard-free implementation of the transcription function equals

$$\tau_u(x)_i = \ast_{y \in \text{res}(x)} o(s_{i-1}(y), y_i) = \ast_{y \in \text{res}(x)} o(\circ_{j=1}^{i-1} t(\cdot, y_j))(s_0), y_i) = \ast_{y \in \text{res}(x)} o(\cdot, y_i) \circ \left(\circ_{j=1}^{i-1} t(\cdot, y_j) \right) (s_0).$$

We apply Corollary 3.12 to $E_i(y_i) = o(\cdot, y_i)$ and $E_j(y_j) = t(\cdot, y_j)$ for $j \in \{1, \dots, i-1\}$, yielding that

$$(\mathcal{M}_E)_u(x) = (\mathcal{M}_{E_i})_u(x_i) \cdot_u (\mathcal{M}_{E_{i-1}})_u(x_{i-1}) \cdot_u \dots \cdot_u (\mathcal{M}_{E_1})_u(x_1), \quad (3.14)$$

where, by construction, $(E(y))(s_0) = \left(\circ_{j=1}^{i-1} E_i(y_j) \right) (s_0) = \tau(y)_i$ for all $y \in \prod_{j=1}^{n-1} \mathbb{B}^{\ell_j}$. Recall that by Definition 3.5, for function $f: S \rightarrow \Lambda$ and $\{1\} \subset \Lambda = \mathbb{B}$, $(\mathcal{M}_f \cdot e^{(s_0)})_{\{1\}} = 1$ iff $f(s_0) = 1$. Thus,

$$\tau_u(x)_i = \left(\ast_{y \in \text{res}(x)} \mathcal{M}_{E(y)} \cdot e^{(s_0)} \right)_{\{1\}} = \left((\mathcal{M}_E)_u(x) \cdot_u e^{(s_0)} \right)_{\{1\}} = ((\mathcal{M}_E)_u(x))_{\{1\}\{s_0\}},$$

where the second equality follows analogously to Observation 3.2. It is hence sufficient to provide a circuit of the claimed size and depth computing this value for all $i \in [n]$. To this end, we compute the entries of the corresponding matrices of the rhs of Equation (3.14), that is, we compute for each $j \in \{1, \dots, i-1\}$ and each $A, B \subseteq S$ by using a hazard-free multiplexer (cf. [IKL⁺18, Lemma 5.1]) the mapping of x_j to

$$\left((\mathcal{M}_{E_j})_u(x_j) \right)_{BA} = \ast_{y_j \in \text{res}(x_j)} \begin{cases} 1 & \text{if } t(A, y_j) \subseteq B \\ 0 & \text{else.} \end{cases}$$

Note that it suffices to perform this computation once for each index $j \in \{1, \dots, n-1\}$. As the size of each such multiplexer is $\mathcal{O}(|\Sigma|) = \mathcal{O}(2^\ell)$, the total size of these subcircuits is therefore in $\mathcal{O}(2^\ell 2^{2|S|} n)$; the depth of each such multiplexer is $\mathcal{O}(\ell)$.

Similarly, we use hazard-free multiplexers to compute for each $A \subseteq S$

$$\left((\mathcal{M}_{E_i})_u(x_i) \right)_{\{1\}A} = \ast_{y_i \in \text{res}(x_i)} \begin{cases} 1 & \text{if } o(A, y_i) = 1 \\ 0 & \text{else.} \end{cases}$$

This results total size $\mathcal{O}(2^\ell 2^{2|S|} n)$ and depth $\mathcal{O}(\ell)$.

Finally, we need to perform the matrix multiplication operations. By Corollary 3.4, we can multiply $2^{|S|}$ by $2^{|S|}$ matrices using $\mathcal{O}(2^{3|S|})$ gates and depth $\mathcal{O}(|S|)$, and by Observation 3.3, $\circ_{\mathbf{u}}$ is associative. Thus, we can plug the matrix multiplication circuit into the construction from [LF80], yielding a circuit of total size $\mathcal{O}(2^{3|S|}n)$ and depth $\mathcal{O}(|S| \log n)$ computing all required products except for the final multiplication with $(\mathcal{M}_{E_i})_{\mathbf{u}}(x_i)$, the matrix representing the output function. Here, $\beta = \gamma = |S|$ and $\alpha = 2^{|\mathbb{B}|} = 4$ in the application of Corollary 3.4, so the total size is bounded by $\mathcal{O}(2^{2|S|}n)$ and the depth bound is the same.

Altogether, we arrive at size $\mathcal{O}((2^{3|S|} + 2^{2|S|+\ell})n)$ and depth $\mathcal{O}(\ell + |S| \log n)$, completing the proof for the special case $\Lambda = \mathbb{B}$. For the general case, the only change is that we have to assemble separate matrices for each bit of each output symbol and multiply them with the computed products $(\mathcal{M}_{E_{i-1}})_{\mathbf{u}}(x_{i-1}) \circ_{\mathbf{u}} \dots \circ_{\mathbf{u}} (\mathcal{M}_{E_1})_{\mathbf{u}}(x_1)$. However, the number of possible distinct functions resulting from mapping a stable input symbol to a (restricted) output function is bounded by $|\Sigma| = 2^\ell$, implying that this is the maximum number of bits we need to compute; further bits are simply copies of others that can be provided by additional output wires without adding gates. Thus, we replace additive $\mathcal{O}(2^\ell 2^{3|S|}n)$ and $\mathcal{O}(2^{2|S|}n)$ terms by $\mathcal{O}(2^{2\ell+|S|}n)$ and $\mathcal{O}(2^{2|S|+\ell}n)$ terms, respectively, while the depth of the circuit does not change. This shows the claimed bounds for the general case, completing the proof. \square

We remark that in the case of $\Sigma \subset \mathbb{B}^\ell$ the choice how to extend t to the domain $S \times \mathbb{B}^\ell$ matters. Despite the fact that we “merely” choose how to treat non-input symbols, it is possible that an unstable input resolves to such a symbol. If this happens, the choice of where t maps this resolution to affects the value the hazard-free extension takes. A “bad” extension may result in unstable output without need, decreasing the usefulness of the constructed circuit.

Fortunately, when using a k -recoverable encoding of the input, the extension of the decoding function to non-codewords induces a suitable extension of t in a natural way. We see this as another indicator that k -recoverability is a central property of codes in the context of hazard-free circuits.

We conclude this section with a slight generalization of Theorem 3.13 that will be useful in Section 5. Instead of allowing only a single pass over the input and having to generate the output right away, we give our transducers the additional power to perform some “preprocessing,” by traversing the input string to compute, e.g., a parity bit. Such preprocessing runs can be simulated by a “standard” transducer, by first copying the input and then separating it by a new input symbol ‘|’ that is added to the input alphabet. The transducer is then allowed to produce arbitrary output symbols during the initial pass over the first copy of the input, and we later only use the output generated in the second pass.

Corollary 3.15. *Let $M = (S_M, s_M, \Sigma, t_M)$ be a finite state machine, i.e.,*

- S_M is the state space,
- s_M is the starting state,
- $\Sigma \subseteq \mathbb{B}^\ell$ is the input alphabet, and
- $t_M: S_M \times \Sigma \rightarrow S_M$ is the state transition function.

Moreover, let $T = (S, s_0, \Sigma, \Lambda, t, o)$ be a Moore machine, $\Delta: S_M \rightarrow S$, and σ an arbitrary permutation on $\{1, \dots, n\}$, for some $n \in \mathbb{N}$. Consider the Boolean function $\tau: \Sigma^n \rightarrow \Lambda^n$ defined as follows. Given $x \in \Sigma^n$, let M process the input string $x_{\sigma(1)}x_{\sigma(2)} \dots x_{\sigma(n)}$, resulting in final state $s \in S_M$. Then let T process x with starting state $\Delta(s)$, yielding the output. Then $\tau_{\mathbf{u}}$ can be implemented by a circuit of size $2^{\mathcal{O}(|S_M|+|S|+\ell)}n$ and depth $\mathcal{O}(\ell + (|S_M| + |S|) \log n)$.

Proof. Set $\Sigma' := \mathbb{B}^{\ell+1}$ and embed Σ into Σ' by via $\iota(x)_i := x_i 0$, i.e., by padding each input symbol with a trailing 0. Define the new symbol $| := 0^\ell 1 \in \Sigma'$. Define $x' := \iota(x_{\sigma(1)}x_{\sigma(2)} \dots x_{\sigma(n)})| \iota(x)$. Define a new Moore machine processing input x' in the following way. The state space is $S_M \cup S$ and

the starting state s_M . The input alphabet is Σ' and the output alphabet Λ . The state transition function is given by the state transition functions of the individual machines for all inputs with a trailing 0 and by Δ on input symbol $|$. The output function is given by o on states from S (ignoring the trailing 0 of input symbols) and arbitrary on other states.

As can be easily verified, the resulting Moore machine implements τ on Boolean inputs, if we simply ignore all outputs until $|$ has been processed. As the length of x' is up to a constant factor that of x and x' can be obtained by a trivial circuit, the claim now is immediate from Theorem 3.13. \square

4 Preserving and Recoverable Codes

In this work, we focus on non-lossy codes. For this reason, we require encoding functions to be injective. We now state the core definition characterizing the codes we need to efficiently avoid the k -bit hazards of relevance to us.

Definition 4.1 (*k*-preserving & recoverable Codes). *A code is an injective function $\gamma: [M] \rightarrow \mathbb{B}^n$ (where $[M] := \{0, \dots, M-1\}$), whose image is the set of codewords. By γ^{-1} we denote the decoding function, i.e., the inverse of γ on its image. For range $r = \langle i, i+k \rangle_M$, where $\langle i, i+k \rangle_M$ denotes $k+1$ consecutive integers modulo M starting at i , define $x_r := * \gamma(r)$ as the corresponding extended codeword and define its range as $r_{x_r} := r$.*

- (i) *For extended codeword x , $p_x := |r_x| - 1$ is its imprecision (i.e., codewords are 0-imprecise).*
- (ii) *Let $k \in \mathbb{N}_0$. Code γ is k -preserving, iff for all $p \leq k$ and p -imprecise extended codewords x it holds that $\gamma([M]) \cap \text{res}(x) = \gamma(r_x)$. That is, $\text{res}(x)$ does not contain “new” codewords which are not in $\gamma(r_x)$.*
- (iii) *Let $k \in \mathbb{N}_0$. Code γ is k -recoverable, iff there is an extension of γ^{-1} to \mathbb{B}^n (i.e., the entire codomain of γ) such that for all extended codewords x of imprecision $p_x \leq k$ and all $y \in \text{res}(x)$ it holds that $\gamma^{-1}(y) \in r_x$.*

Note that γ being k -recoverable is at least as restrictive as it being k -preserving: the latter is a necessary condition for the extension of γ^{-1} required by the former to exist.

Put simply, a k -preserving code “preserves” an uncertainty of up to k , which corresponds to detecting an error in the terminology of error correcting codes. If the code is also recoverable, we can also “correct” errors, in the sense that we can map the string after resolution back to a codeword from the original range of represented values.

In order to guarantee k -recoverability, a code must have at least k (almost) redundant bits.

Theorem 4.2. *For $k, n \in \mathbb{N}$, no code $\gamma: [(k+1) \cdot 2^n] \rightarrow \mathbb{B}^{n+k}$ is p -recoverable for $p > k$.*

Proof. Consider any k -recoverable code γ on domain $[M]$ for some $M \in \mathbb{N}$. For $i \in [M-k]$, set $x := \gamma(i)$. For $j \in [k+1]$, we show by induction that $\gamma(i)$ and $\gamma(i+j)$ have Hamming distance j . The base case $j = 0$ is trivial. For the step from j to $j+1$, consider $y = x_{\langle i, i+j \rangle_M}$. As γ is k -preserving, the only codewords that $\text{res}(y)$ contains are $\gamma(i), \dots, \gamma(i+j)$. Accordingly, $\gamma(i+j+1) \notin \text{res}(y)$ must differ from y in a stable bit. However, $y = \gamma(i) * \gamma(i+1) * \dots * \gamma(i+j)$, implying that $\gamma(i), \dots, \gamma(i+j)$ agree on this bit. Hence, by the induction hypothesis, $\gamma(i)$ and $\gamma(i+j+1)$ have Hamming distance $j+1$.

Using that the code is k -preserving once more, we get that $\text{res}(x_{\langle i, i+k \rangle_M})$ contains exactly the codewords i to $i+k$. Moreover, as shown above, $\gamma(i)$ and $\gamma(i+k)$ have Hamming distance k , implying that $|\text{res}(x_{\langle i, i+k \rangle_M})| = 2^k$.

Now fix $M = (k + 1) \cdot 2^n$. The sets $S_\ell := \{\gamma((k + 1)\ell), \dots, \gamma((k + 1)(\ell + 1) - 1)\}$, where $\ell \in [2^n]$, form a partition of the codewords. Because γ is k -recoverable, γ^{-1} can be extended such that for each $\ell \in [2^n]$, $\gamma^{-1}(\text{res}(*S_\ell)) \subseteq \{(k + 1)\ell, \dots, (k + 1)(\ell + 1) - 1\}$. Hence, it follows that $\text{res}(*S_\ell) \cap \text{res}(*S_{\ell'}) = \emptyset$ for all $\ell \neq \ell' \in [2^n]$. As $|\text{res}(*S_\ell)| = 2^k$ for each $\ell \in [2^n]$, the codomain of γ has size at least $2^k \cdot 2^n$. \square

In particular, unary code on $M - 1$ bits is the shortest code that has M codewords and guarantees $(M - 1)$ -recoverability. Once we fix these constraints, the only choice left is to pick $\gamma(0)$ and the order in which bits are flipped. Bounding k is thus necessary to obtain codes with less redundancy.

The next lemma sheds some light on why k -recoverability is the “right” property to ask for. For $p \leq k$, with a k -recoverable code we can implement any Function $f: [M] \rightarrow \mathbb{B}^m$ by a circuit mapping imprecision- p codewords x to $*_{i \in r_x} f(i)$. This should be read as having information on a range the input is from and obtaining the most precise output that is possible in Kleene logic under this assumption.

Lemma 4.3. *Let $f: [M] \rightarrow \mathbb{B}^m$. For any k -recoverable code $\gamma: [M] \rightarrow \mathbb{B}^n$, there is a combinational circuit C satisfying for all $p \leq k$ and p -imprecise codewords x that*

$$C(x) = *_{i \in r_x} f(i).$$

Proof. Since γ is k -recoverable, there is an extension of γ^{-1} to \mathbb{B}^n such that for all p -imprecise extended codewords x we have that $\gamma^{-1}(\text{res}(x)) \subseteq r_x$ (see Definition 4.1, Item iii). As $x = *_{i \in r_x} \gamma(i)$, we also have that $\gamma(i) \in \text{res}(x)$ for all $i \in r_x$ and hence $\gamma^{-1}(\text{res}(x)) = r_x$.

We take a hazard-free multiplexer circuit with n control bits and, accordingly 2^n selectable inputs. The input corresponding to control string $x \in \mathbb{B}^n$ is fed the (constant, as known a priori) string $f(\gamma^{-1}(x))$. We then use the input $x \in \mathbb{B}^n$ to the circuit as control string of the multiplexer. As the multiplexer is hazard-free, it outputs for p -imprecise codeword x (where $p \leq k$)

$$*_{y \in \text{res}(x)} f(\gamma^{-1}(y)) = *_{i \in \gamma^{-1}(\text{res}(x))} f(i) = *_{i \in r_x} f(i). \quad \square$$

In particular, any two k -recoverable codes with the same number of M codewords can be mapped to each other by a circuit that translates for $p \leq k$ all imprecision- k codewords to their counterparts.

Corollary 4.4. *Let $\gamma: [M] \rightarrow \mathbb{B}^n$ and $\gamma': [M] \rightarrow \mathbb{B}^m$ be k -recoverable codes. There is a circuit mapping for all $r := \langle i, i + p \rangle$, where $p \leq k$ and $i \in [M - p]$, the imprecision- p codeword $*_{i \in r} \gamma(i)$ to $*_{i \in r} \gamma'(i)$.*

Proof. Apply Lemma 4.3 to code γ and the function γ' . \square

These statements together essentially tell that k -recoverable codes permit to jump back and forth between encodings freely without losing precision, so long as the imprecision is limited to k . Note, however, that the above implementation is not efficient, as the multiplexer circuit is of size exponential in n , even if k is small. In Section 5, we will use the machinery from Section 3 to obtain an efficient circuit for addition. The remainder of this section is dedicated to developing a recoverable code for which we can do so.

4.1 Constructing a k -Preserving, $\lceil k/2 \rceil$ -Recoverable code

Roughly speaking, our code is a hybrid of binary reflected Gray code and standard unary code. We start by introducing these codes individually.

Table 1: 4-bit binary reflected Gray code. Note that this code is already 1-preserving.

#	rg ₄	#	rg ₄	#	rg ₄	#	rg ₄
0	0 000	4	0 110	8	1 100	12	1 010
1	0 001	5	0 111	9	1 101	13	1 011
2	0 011	6	0 101	10	1 111	14	1 001
3	0 010	7	0 100	11	1 110	15	1 000

Reflected Gray Code.

Definition 4.5 (rg_n). We define n -bit binary reflected Gray code (cf. Table 1), $\text{rg}_n : [2^n] \rightarrow \mathbb{B}^n$, recursively. A 1-bit code is given by $\text{rg}_1(0) = 0$ and $\text{rg}_1(1) = 1$. For $n > 1$, we start with the first bit fixed to 0 and counting with $\text{rg}_{n-1}(\cdot)$ (for the first 2^{n-1} codewords), then toggle the first bit to 1, and finally “count down” $\text{rg}_{n-1}(\cdot)$ while fixing the first bit again, cf. Table 1. Formally, this yields for $x \in [N]$

$$\text{rg}_n(x) := \begin{cases} 0 \text{rg}_{n-1}(x) & \text{if } x \in [2^{n-1}] \\ 1 \text{rg}_{n-1}(2^n - 1 - x) & \text{if } x \in [2^n] \setminus [2^{n-1}]. \end{cases}$$

As each n -bit string is a codeword, the code is a bijection. Observe that exactly one bit is flipped in each two consecutive codewords, making the code a Gray code. This implies that the code is 1-recoverable, and it matches the bound from Lemma 4.2. Our goal is to generalize to a construction providing a k -recoverable code.

The parity of a bit string string is defined as the **xor** over all bits.

Definition 4.6. The parity $\text{par} : \mathbb{B}^n \rightarrow \mathbb{B}$ of a bit string $z \in \mathbb{B}^n$, where $n \in \mathbb{N}$, is

$$\text{par}(z) := \bigoplus_{i=1}^n z_i,$$

where for $x, y \in \mathbb{B}$, $x \oplus y := \text{xor}(x, y)$.

Equivalently, reflected Gray code can be defined in a left-recursive manner.

Lemma 4.7. For $n \in \mathbb{N}$, the reflected Gray code of $\ell \in [2^n]$, is given by

$$\text{rg}_n(\ell) = \text{rg}_{n-1}(\lfloor \ell/2 \rfloor)(\text{par}(\text{rg}_{n-1}(\lfloor \ell/2 \rfloor)) \oplus (\ell \bmod 2)).$$

Proof. Taking every second codeword and removing the last bit reproduces the list of codewords of rg_{n-1} , i.e., $\text{rg}_n(\ell)_{1,n-1} = \text{rg}_{n-1}(\lfloor \ell/2 \rfloor)$. Concerning the last bit, observe that the last bit toggles on every up-count to an odd encoded number. As the parity of $(n-1)$ -bit prefix changes on every up-count to an even encoded number, the claim follows. \square

This gives a simple left-recursive characterization of the decoding function, which will come in handy later.

Corollary 4.8. For all $n \in \mathbb{N}$ and $x \in \mathbb{B}^n$,

$$\text{rg}_n^{-1}(x) = 2 \text{rg}_{n-1}^{-1}(x_{1,n-1}) + \text{par}(x).$$

Table 2: A 4-bit 1-recoverable code. The front part of the code consists of 2-bit Gray code rg_2 , and the rear part of the code, ξ_2 , toggles between 2-bit standard unary code and its complement. The 1-imprecise extended codewords are denoted by ‘-’ and appear in between their corresponding codewords, e.g., $00\mathbf{u}0 = *\gamma(\langle 0, 1 \rangle_M)$.

rg_2, ξ_2	#						
00 00	0	01 11	3	11 00	6	10 11	9
00 $\mathbf{u}0$	–	01 $\mathbf{u}1$	–	11 $\mathbf{u}0$	–	10 $\mathbf{1u}$	–
00 10	1	01 01	4	11 10	7	10 10	10
00 $\mathbf{1u}$	–	01 $\mathbf{0u}$	–	11 $\mathbf{1u}$	–	10 $\mathbf{u}0$	–
00 11	2	01 00	5	11 01	8	10 00	11
$\mathbf{0u}$ 11	–	$\mathbf{u1}$ 00	–	$\mathbf{1u}$ 11	–	–	–

Unary Code. From Theorem 4.2, we know that for $k > 1$, we need to add redundancy to the code. The proof of the theorem actually guides the design further: We must toggle $k + 1$ distinct bits on $k + 1$ consecutive up-counts, or the code cannot be k -recoverable. This implies that the subcode for such consecutive codewords is actually unary. It turns out¹¹ that one can reuse the “same” unary subcode repeatedly, with the modification that the bits are toggled in the same order every time the code is reused.

Definition 4.9 (ξ_n and $\bar{\xi}_n$). For $n \in \mathbb{N}_0$ and $i \in [n + 1]$, define

$$\xi_n(i) := 1^i 0^{n-i} \quad \text{and} \quad \bar{\xi}_n(i) := \overline{\xi_n(i)} = 0^i 1^{n-i}.$$

Clearly, both ξ_n and $\bar{\xi}_n$ are Gray codes.

Observation 4.10. Concatenate the list of codewords of $\xi_n([n])$ and $\bar{\xi}_n([n])$, the result being a cyclic code of $2n$ words. When taking $*$ over any segment of up to n words, the resolution contains only those words.

Proof. The claim is readily verified for the segment $1^{n-j}0^j$ for $j \in [n]$. The general case of segments starting at some $x \in \mathbb{B}^n$ follows by observing that the bit-wise **xor** with x rotates the cycle so that x becomes 0^n . \square

The Hybrid Code. We define our n -bit k -preserving code $\gamma_{n,k} : [(k + 1) \cdot 2^{n-k}] \rightarrow \mathbb{B}^n$ by combining the two above codes as follows.

Definition 4.11 ($\gamma_{n,k}$). $\gamma_{n,k}$ consists of two parts, where the front part is a binary reflected Gray code of $n - k$ bits rg_{n-k} and the rear part keeps repeating ξ_k and $\bar{\xi}_k$ alternately. The front part “counts up” on every $(k + 1)$ -th up-count of the compound code, and its current parity decides on the values of the rear part for these consecutive up-counts. If it is 0, then we use ξ_k , if it is 1, we use $\bar{\xi}_k$. That is,

$$\gamma_{n,k}(i) := \text{rg}_{n-k}(\lfloor i/(k + 1) \rfloor) \begin{cases} \xi_k(i \bmod (k + 1)) & \text{if } \text{par}(\text{rg}_{n-k}(\lfloor i/(k + 1) \rfloor)) = 0 \\ \bar{\xi}_k(i \bmod (k + 1)) & \text{if } \text{par}(\text{rg}_{n-k}(\lfloor i/(k + 1) \rfloor)) = 1. \end{cases}$$

See Table 2 for an example.

¹¹In hindsight, the proof of Theorem 4.2 strongly suggests this.

Observation 4.12. For every $i \in [M]$, where $M := (k+1) \cdot 2^{n-k}$, and $j \in [k+1]$, the Hamming distance of $\gamma_{n,k}(i)_1 \dots \gamma_{n,k}(i)_{n-k}$ and $\gamma_{n,k}(i+j \bmod M)_1 \dots \gamma_{n,k}(i+j \bmod M)_{n-k}$ is at most 1.

It is easy to verify that this code is k -preserving.

Lemma 4.13. $\gamma_{n,k}$ is k -preserving.

Proof. Consider imprecision- p codeword x for $p \leq k$. By Observation 4.12, at most one bit among the first $n-k$ bits of x equals \mathbf{u} . If there is no such bit, the claim is immediate from Observation 4.10. If there is such a bit, we can interpret this bit (or its complement) as an additional bit in the unary code (resulting in a ring of $2k+2$ words), and the claim follows from Observation 4.10 as well. \square

The resolution of a bit string may result in non-codewords. Hence, we extend the decoding function to all bit strings.

Definition 4.14 (Extension of $\gamma_{n,k}^{-1}$). First, we define mappings $u: \mathbb{B}^k \times \mathbb{B} \rightarrow \xi_k([k+1])$ and $\bar{u}: \mathbb{B}^k \times \mathbb{B} \rightarrow \bar{\xi}_k([k+1])$ from bit strings to our unary encodings. For $z \in \mathbb{B}^k$, the mappings produce a codeword sharing either the longest prefix or suffix with z , depending on the second input. That is, for $b \in \mathbb{B}$ let ℓ_x^{\max} be the maximal index such that $z_{\ell_x^{\max}} = b$ and ℓ_b^{\min} be the minimal index such that $z_{\ell_b^{\min}} = b$. Then we define

$$\begin{aligned} u(z, 0) &:= 1^{\ell_0^{\min}-1} 0^{k-\ell_0^{\min}+1}, & u(z, 1) &:= 1^{\ell_1^{\max}} 0^{k-\ell_1^{\max}}, \\ \bar{u}(z, 1) &:= 0^{\ell_1^{\min}-1} 1^{k-\ell_1^{\min}+1}, & \bar{u}(z, 0) &:= 0^{\ell_0^{\max}} 1^{k-\ell_0^{\max}}. \end{aligned}$$

Note that $\bar{u}(z, x) = \overline{u(\bar{z}, \bar{x})}$.

For an arbitrary bit string $x \in \mathbb{B}^n$, we define

$$u(x) := u(x_{n-k+1,n}, x_{n-k+\lfloor k/2 \rfloor+2}) \quad \text{and} \quad \bar{u}(x) := \bar{u}(x_{n-k+1,n}, x_{n-k+\lfloor k/2 \rfloor+2})$$

and

$$\gamma_{n,k}^{-1}(x) = x_{1,n-k} \begin{cases} u(x) & \text{if } \text{par}(x_{1,n-k}) = 0 \\ \bar{u}(x) & \text{if } \text{par}(x_{1,n-k}) = 1. \end{cases}$$

Note that for codewords x , $\gamma_{n,k}(\gamma_{n,k}^{-1}(x)) = x$, as then $u(x_{k+1,n}, 0) = u(x_{k+1,n}, 1) = x_{k+1,n}$ if $\text{par}(x_{1,n-k}) = 0$ and $\bar{u}(x_{k+1,n}, 0) = \bar{u}(x_{k+1,n}, 1) = x_{k+1,n}$ if $\text{par}(x_{1,n-k}) = 1$, i.e., the above definition indeed provides an extension of the decoding function $\gamma_{n,k}^{-1}$ to \mathbb{B}^n . We now prove that this extension shows γ to be $\lceil k/2 \rceil$ -recoverable.

Theorem 4.15. $\gamma_{n,k}: [(k+1) \cdot 2^{n-k}] \rightarrow \mathbb{B}^n$ is $\lceil k/2 \rceil$ -recoverable.

Proof. Throughout this proof, write $x = yz$ for all $x \in \mathbb{T}^n$ with $y \in \mathbb{T}^{n-k}$ and $z \in \mathbb{T}^k$. We need to show that for any $p \leq \lceil k/2 \rceil$ and imprecision- $\lceil k/2 \rceil$ codeword x , we have that $\gamma_{n,k}^{-1}(x) \subseteq r_x$, where $\gamma_{n,k}^{-1}$ is extended to \mathbb{B}^n according to Definition 4.14. Accordingly, fix $r = \langle i, i+p \rangle_{(k+1)2^n}$ for any $i \in (k+1)2^n$ and $p \leq \lceil k/2 \rceil$. W.l.o.g., assume that $i+p < (k+1)2^n$; otherwise, simply take all respective values modulo $(k+1)2^n$. Set $x(j) := \gamma(i+j)$ for all $j \in [p+1]$ and $x := x_r = *_{i \in r} x(j)$ for the remainder of the proof. Moreover, for $i \in \mathbb{N}$, abbreviate $\underline{i} := \lfloor i/(k+1) \rfloor$.

Suppose first that $\underline{i} = \underline{i+p}$. Then $y(j) = \text{rg}_{n-k}(\underline{i}) = y$ for all $j \in [p+1]$. If $\underline{i} \in 2\mathbb{N}_0$, $z(j) = 1^{\ell_j} 0^{k-\ell_j}$, where $\ell_j = i - k\underline{i} + j$. Thus, $x = y 1^{\ell} \mathbf{u}^p 0^{k-\ell-p}$, where $\ell = i - k\underline{i}$. It is now immediate that for any resolution $x' \in \text{res}(x)$, we have that $u(z', z'_{1+\lfloor k/2 \rfloor}) = z(j)$ for some $j \in [p+1]$, implying that $\gamma_{n,k}^{-1}(x') \in r$. Analogous reasoning shows the same for $y \in 2\mathbb{N}_0 + 1$.

Hence, consider the case that $\underline{i} = i + p - 1$, i.e., the front part (i.e., the $(n - k)$ -bit prefix) of $\gamma_{n,k}$ performs an up-count within r . We claim that this implies that $z_{1+\lfloor k/2 \rfloor} = b \in \mathbb{B}$, i.e., all resolutions of x agree on $z_{1+\lfloor k/2 \rfloor}$. To see this, observe that after changing the bit in the sequence given by r , only bits z_i with $i \leq p - 1 \leq \lfloor k/2 \rfloor < 1 + \lfloor k/2 \rfloor$ can change, and before changing the bit in the front part, only bits z_i with $i \geq k - (p - 2) \geq k - \lfloor k/2 \rfloor + 2 > \lfloor k/2 \rfloor + 1$ can change.

Suppose that the bit in the front part changes on the up-count from $\gamma(i + j_0)$ to $\gamma(i + j_0 + 1)$, where $j_0 \in [p]$. Consider the case that $\underline{i} \in 2\mathbb{N}_0$ first. Then, reasoning as above, we get that $z = \mathbf{u}^{p-(j_0+1)}\mathbf{1}^{k-p+1}\mathbf{u}^{j_0}$. In particular, $z_{1+\lfloor k/2 \rfloor} = 1$. Thus, for any stabilization $z' \in \text{res}(z)$, we have that $\xi_k^{-1}(u(z, 1)) \geq k - j_0$. Thus, if for $x' = y'z' \in \text{res}(x)$ we have that $y' = \gamma(i)_1 \dots \gamma(i)_{n-k}$, then $\gamma^{-1}(x') \in \{k\underline{i} + k - j_0, \dots, k\underline{i} + k\}$. As the up-count in the front part occurs from $\gamma(i + j_0)$ to $\gamma(i + j_0 + 1)$, we have that $k\underline{i} + k = i + j_0$, we conclude that $\gamma_{n,k}^{-1}(x') \in r$.

Similarly, any stabilization $z' \in \text{res}(z)$ satisfies that $\bar{\xi}_k^{-1}(\bar{u}(z, 1)) \leq p - (j_0 + 1)$, as there can be at most this many leading zeroes. Thus, $\text{rg}_{n-k}^{-1}(y') \in 2\mathbb{N}_0 + 1$ entails that $\gamma_{n,k}^{-1}(x') \in \{k(\underline{i} + 1) + 1, \dots, k(\underline{i} + 1) + p - j_0\} \subset r$.

Finally, it remains to consider the case that $\underline{i} \in 2\mathbb{N}_0 + 1$. Then $z = \mathbf{u}^{p-(j_0+1)}\mathbf{0}^{k-p+1}\mathbf{u}^{j_0}$, implying that $z_{1+\lfloor k/2 \rfloor} = 0$. We can deduce that $\gamma_{n,k}^{-1}(x') \in r$ for each $x' \in \text{res}(x)$ analogously to the case that $\underline{i} \in 2\mathbb{N}_0$. \square

The lower bound from Theorem 4.2 matches exactly the number of codewords of $\gamma_{n,k}$, but we can only show that the code is k -preserving. Theorem 4.15 shows the matching property of k' -recoverability, but only for $k' = \lfloor k/2 \rfloor$. This leaves open the question whether the upper bounds are tight or some improvement is feasible.

5 k -Recoverable Addition

In this section, we utilize the results from Section 3 to provide circuits for k -recoverable addition of $\gamma_{n,k}$, the k -preserving, $\lfloor k/2 \rfloor$ -recoverable code given in Section 4.

Definition 5.1 (*k -recoverable Addition*). For code $\gamma: [M] \rightarrow \mathbb{B}^n$, we define γ -addition in the natural way: For codewords $x, y \in \gamma[M]$, $x +_\gamma y = \gamma(\gamma^{-1}(x) + \gamma^{-1}(y) \bmod M)$. If γ is k -recoverable for $k \in \mathbb{N}$, we extend this definition to all $x, y \in \mathbb{B}^n$ by the (not necessarily unique) extension of γ^{-1} to \mathbb{B}^n k -recoverability guarantees. A circuit implements k -recoverable addition w.r.t. k -recoverable code γ , iff it implements (the extended) operator $+_\gamma$ on extended codewords x and y satisfying that $p_x + p_y \leq k$ without hazards.

This definition enforces that the circuit faithfully implements interval addition on the ranges of the operands.

Observation 5.2. For extended codewords x and y satisfying $p_x + p_y \leq k$, a k -recoverable adder circuit outputs

$$x(+_\gamma)_u y = \underset{i \in r_x + r_y \bmod M}{*} \gamma(i).$$

Proof. Using that γ^{-1} maps resolutions of extended codewords back to the respective range, we get that

$$\begin{aligned} x(+_\gamma)_u y &= \underset{(x', y') \in \text{res}(x) \times \text{res}(y)}{*} \gamma(\gamma^{-1}(x') + \gamma^{-1}(y') \bmod M) \\ &= \underset{(i, j) \in r_x \times r_y}{*} \gamma(i + j \bmod M) \\ &= \underset{i \in r_x + r_y \bmod M}{*} \gamma(i). \end{aligned} \quad \square$$

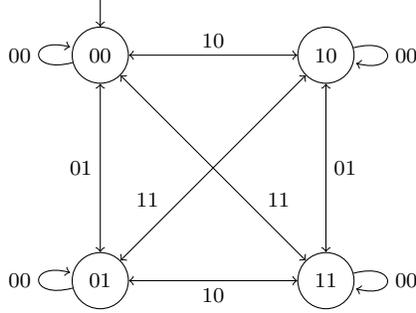


Figure 3: Finite state machine T_{par} , computing the parity of two input strings.

Note that this observation shows that on inputs with sufficiently small imprecision, the output of the circuit does not depend on the chosen extension of γ^{-1} to non-codewords. Combining this insight with Lemma 4.3 and Corollary 4.4, arguably we get a universal tool for handling addition up to imprecision k in Kleene logic. We now proceed to deriving circuits realizing k -recoverable addition.

5.1 Transducers for Adding Stable Strings

In order to obtain the desired circuit, we will apply Corollary 3.15, which reduces our task to designing (due to inductive application) a sequence of finite state machines and a single final transducer generating the output *on stable strings*. We cannot emphasize enough the degree to which this simplifies the task; in fact, most of the steps turn out to be very simple. We first describe the individual steps and then plug together the final machine.

5.1.1 Parity

The finite state machine that determines the parity of two inputs $x, y \in \mathbb{B}^\ell$ is given by $M_{\text{par}} := (S_{\text{par}}, 00, \mathbb{B}^2, t_{\text{par}})$, where the state space is $S_{\text{par}} = \mathbb{B}^2$. The transition function t_{par} is defined as follows:

$$\begin{aligned} (t_{\text{par}}(s_{i-1}, x_i y_i))_1 &:= (s_{i-1})_1 \oplus x_i, \\ (t_{\text{par}}(s_{i-1}, x_i y_i))_2 &:= (s_{i-1})_2 \oplus y_i. \end{aligned}$$

This is a simple product machine, where the first bit of the state encoding denotes the parity of the first input string and the second bit the parity of the second string, respectively. The i -th input symbol is $x_i y_i$. A visualisation of M_{par} is given in Figure 3.

Observation 5.3. *After processing the sequence $(x_1 y_1)(x_2 y_2) \dots (x_\ell y_\ell)$ for inputs $x, y \in \mathbb{B}^\ell$, M_{par} is in state $\text{par}(x) \text{par}(y)$.*

5.1.2 Unary Addition

Finite State Machine Computing the Sum. Consider $x, y \in \xi_k([k+1])$. We would like to determine $\xi_k^{-1}(x) + \xi_k^{-1}(y)$. A simple $(2k+1)$ -state machine is given by $M_{\text{un}} = ([2k+1], 0, \mathbb{B}^2, t_{\text{un}})$,

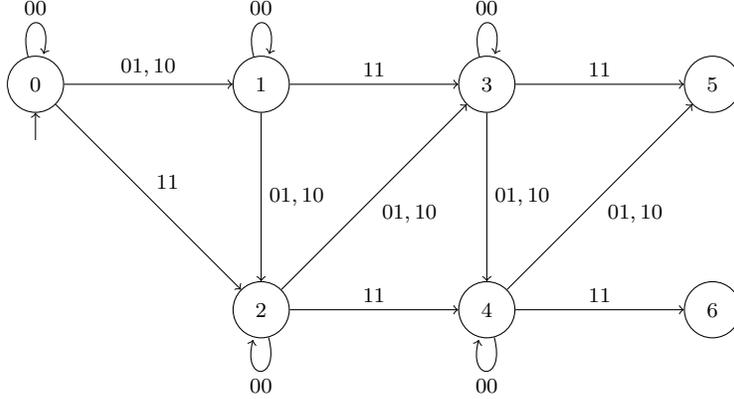


Figure 4: Finite state machine M_{un} for $k = 3$. Transitions for states 5 and 6 are omitted, as these states can only be reached after processing the last input symbol.

where the transition function is

$$\begin{aligned}
 t_{un}(s_{i-1}, 00) &:= s_{i-1}, \\
 t_{un}(s_{i-1}, 01) &:= (s_{i-1} + 1) \bmod (2k + 1), \\
 t_{un}(s_{i-1}, 10) &:= (s_{i-1} + 1) \bmod (2k + 1), \\
 t_{un}(s_{i-1}, 11) &:= (s_{i-1} + 2) \bmod (2k + 1).
 \end{aligned}$$

The i -th input symbol is $x_i y_i$. The modulo operation is, in fact, irrelevant, as an overflow is not possible; hence, t_{un} could be chosen arbitrarily on state/input pairs for which an overflow would occur. An example for $k = 3$ is given in Figure 4.

Observation 5.4. *After processing the sequence $(x_1 y_1)(x_2 y_2) \dots (x_k y_k)$ for inputs $x, y \in \xi_k([k+1])$, M_{un} is in state $s_k = \xi_k^{-1}(x) + \xi_k^{-1}(y)$.*

Generalization to Non-Codewords. We defined the mappings u and \bar{u} to map non-codewords to codewords. We adapt the state machine M_{un} to implicitly map the inputs in the same sense such that we are able to add arbitrary bit strings that are no codewords as follows.

Consider a single input string $z \in \mathbb{B}^k$ and suppose first that $z_{1+\lfloor k/2 \rfloor}$ is known a priori. If $z_{1+\lfloor k/2 \rfloor} = 0$, we interpret any input symbol as 0 once the first 0 occurred. This results in behavior identical to processing $u(z, 0)$. If $z_{1+\lfloor k/2 \rfloor} = 1$, we process the input in reverse order, interpreting any input symbol as 1 once the first 1 occurs. As the above state machine counts the total number of 1s, not their order, this results in identical output to processing $u(z, 1)$.

To enable equivalent behavior for arbitrary input strings, we reorder the input and (formally) concatenate two state machines. The first state machine is presented with initially presented with $z_{1+\lfloor k/2 \rfloor}$, which it stores. Subsequently, it is fed the string $z_1 \dots z_{\lfloor k/2 \rfloor}$. If $z_{1+\lfloor k/2 \rfloor} = 0$, it counts the 1s until the first 0 as above. If $z_{1+\lfloor k/2 \rfloor} = 1$, it simply ignores the input. The second state machine receives input string $z_k z_{k-1} \dots z_{2+\lfloor k/2 \rfloor}$. It is initialized to the state of the previous one if $z_{1+\lfloor k/2 \rfloor} = 0$; in this case it will ignore its input. Otherwise it is initialized to state $(1, k)$, meaning that $z_{1+\lfloor k/2 \rfloor} = 1$, and performs down-counts until it observes the first input of 1.

Observation 5.5. *Given $z \in \mathbb{B}^k$, processing z as described above results in the final state of the second state machine to be $(z_{1+\lfloor k/2 \rfloor}, \xi_k^{-1}(u(z, z_{1+\lfloor k/2 \rfloor}))$.*

Based on this insight, it is straightforward to compute $\xi_k^{-1}(u(x, x_{1+[k/2]})) + \xi_k^{-1}(u(y, y_{1+[k/2]}))$ for inputs $x, y \in \mathbb{B}^k$: processing the input in the right order and memorizing $x_{1+[k/2]}$ and $y_{1+[k/2]}$, we can perform up- and down-counts in the same manner to compute the correct output.

Corollary 5.6. *There is a state machine with $\mathcal{O}(k)$ states that computes $\xi_k^{-1}(u(x, x_{1+[k/2]})) + \xi_k^{-1}(u(y, y_{1+[k/2]}))$ for $x, y \in \mathbb{B}^n$ when fed the input sequence*

$$(x_{1+[k/2]}y_{1+[k/2]})(x_1y_1)(x_2y_2) \dots (x_{[k/2]}y_{[k/2]})(x_ky_k)(x_{k-1}y_{k-1}) \dots (x_{2+[k/2]}y_{2+[k/2]}).$$

Generalization to Arbitrary Input Parities. The addition above yields only the “right” output if the parities of the leading binary reflected Gray code strings of inputs $x = \gamma_{n,k}(i)$ and $y = \gamma_{n,k}(j)$ have both parity 0, as for parity 0 the unary part is encoded using $\bar{\xi}_k$. Fortunately, all that is necessary to adapt the addition for other input parities is to take the complement of the inputs’ unary parts in case the respective parity is 1. As we have that $\bar{u}(z, z_{1+[k/2]}) = \overline{u(\bar{z}, \bar{z}_{1+[k/2]})}$ and $\bar{\xi}_k(i) = \xi_k(\bar{i})$ for all $i \in [k+1]$, we get that

$$\bar{\xi}_k^{-1}(\bar{u}(z, z_{1+[k/2]})) = \bar{\xi}_k^{-1}(\overline{u(\bar{z}, \bar{z}_{1+[k/2]})}) = \xi_k^{-1}(u(\bar{z}, \bar{z}_{1+[k/2]})).$$

Hence, to determine the sum correctly, it is sufficient first run the parity machine M_{par} on the $(n-k)$ -bit prefixes, store the parities, and apply the state machine from Corollary 5.6, where we apply the **xor** with the respective parity to each input symbol before processing it.

Corollary 5.7. *Given inputs $x = \gamma_{n,k}(i)$ and $y = \gamma_{n,k}(j)$, there is a finite state machine M_{ξ} with $\mathcal{O}(k)$ states that determines $\pi_x := \text{par}(x_1 \dots x_{n-k})$, $\pi_y := \text{par}(y_1 \dots y_{n-k})$, and*

$$s_k(x, y) := \begin{cases} \xi_k^{-1}(u(x)) + \xi_k^{-1}(u(y)) & \text{if } \pi_x = \pi_y = 0 \\ \xi_k^{-1}(u(x)) + \bar{\xi}_k^{-1}(\bar{u}(y)) & \text{if } \pi_x = 0 \wedge \pi_y = 1 \\ \bar{\xi}_k^{-1}(\bar{u}(x)) + \xi_k^{-1}(u(y)) & \text{if } \pi_x = 1 \wedge \pi_y = 0 \\ \bar{\xi}_k^{-1}(\bar{u}(x)) + \bar{\xi}_k^{-1}(\bar{u}(y)) & \text{if } \pi_x = \pi_y = 1. \end{cases}$$

Note that π_x and π_y are the parities of the binary reflected Gray code parts and s_k is the correct sum of the unary part of x and y . The state s_k also shows whether the sum is at most k or larger, telling whether we have a carry that increases the sum of the $(n-k)$ -bit prefixes by 1.

5.1.3 Adding Reflected Gray Code

The following key observation permits a transducer performing the addition with a constant number of states.

Lemma 5.8. *Consider $i, j \in [2^n]$ and $c \in \mathbb{B}$. Set $x := \text{rg}_n(i)$ and $y := \text{rg}_n(j)$. Denote by x' and y' the $(n-1)$ -bit prefixes $\text{rg}_n(i)_{1,n-1}$ and $\text{rg}_n(j)_{1,n-1}$, respectively. Then*

$$\text{rg}_n(i + j + c)_{1,n-1} = \text{rg}_{n-1}(\text{rg}_{n-1}^{-1}(x') + \text{rg}_{n-1}^{-1}(y') + b),$$

where

$$b := \begin{cases} 1 & \text{if } \text{par}(x) + \text{par}(y) + c \geq 2 \\ 0 & \text{else.} \end{cases}$$

Moreover,

$$\text{rg}_n(i + j + c)_n = x_n \oplus y_n \oplus b \oplus c.$$

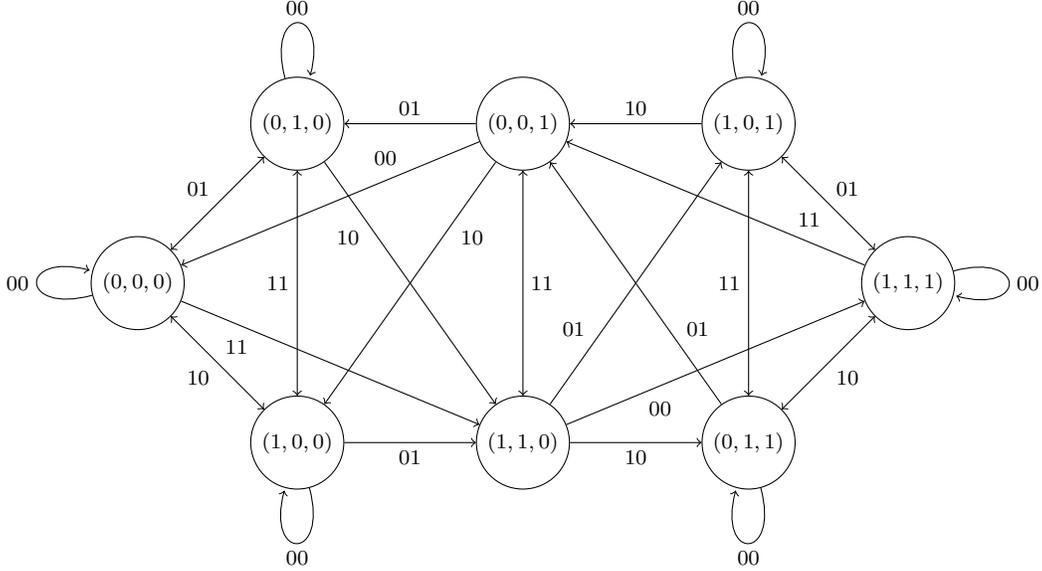


Figure 5: Visualization of the state space S_{rg} and the transition function t_{rg} of T_{rg} . The starting state is determined by M_{ξ} , as either $(0, 0, 0)$ (sum $\leq k$) or $(0, 0, 1)$ (sum $> k$).

Proof. By Lemma 4.8, $\text{rg}_n^{-1}(x) = 2\text{rg}_{n-1}^{-1}(x') + \text{par}(x)$ (and likewise for y). Therefore,

$$i + j + b = \text{rg}_n^{-1}(x) + \text{rg}_n^{-1}(y) + b = 2(\text{rg}_{n-1}^{-1}(x') + \text{rg}_{n-1}^{-1}(y')) + \text{par}(x) + \text{par}(y) + b.$$

As $\lfloor (i + j + b)/2 \rfloor = \text{rg}_{n-1}^{-1}(x') + \text{rg}_{n-1}^{-1}(y') + c$, the first claim holds by Lemma 4.7. Moreover, $i + j + c \bmod 2 = \text{par}(x) \oplus \text{par}(y) \oplus c$, yielding that

$$\text{rg}_n(i + j + c)_n = (\text{par}(x') \oplus \text{par}(y') \oplus b) \oplus (\text{par}(x) \oplus \text{par}(y) \oplus c) = x_n \oplus y_n \oplus b \oplus c,$$

as $\text{par}(z_{1,n-1}) \oplus \text{par}(z_{1,n}) = z_n$ for all $z \in \mathbb{B}^n$. \square

As we can determine the parity using a constant-sized finite state machine, we can use Corollary 3.15 to obtain a small circuit for the addition of the front part of our hybrid code.

Lemma 5.9. *There is a hazard-free circuit of size $2^{\mathcal{O}(k)}n$ and depth $\mathcal{O}(k \log n)$ that computes $(x +_{\gamma_{n,k}} y)_{1,n-k}$ from inputs $x, y \in \gamma_{n,k}([(k+1)2^n])$.*

Proof. First we apply Lemma 5.8 inductively to obtain the transducer T_{rg} , which is adding two reflected Gray code strings. A picture of the transducer is given in Figure 5. The state machine is processing $x_1 \dots x_{n-k}$ and $y_1 \dots y_{n-k}$ from back to front. At step, $i \in [1, n-k]$, it outputs $(x +_{\gamma_{n,k}} y)_{n-k+1-i} = x_{n-k+1-i} \oplus y_{n-k+1-i} \oplus c_i \oplus c_{i-1}$. Here $c_i = 1$ iff $\text{par}(x_{1,n-k+1-i}) + \text{par}(y_{1,n-k+1-i}) + c_{i-1} \geq 2$ and $c_i = 0$ otherwise, where we assume that the initial state¹² to be $s_0 = (\text{par}(x_{1,n-k}), \text{par}(y_{1,n-k}), c_0)$ for $c_0 \in \mathbb{B}$. Thus, for $i \in [k]$, state

$$s_{i+1} = t_{\text{rg}}(s_i, x_{n-k-i}y_{n-k-i}) := (\text{par}(x_{1,n-k-i}), \text{par}(y_{1,n-k-i}), c_{i+1})$$

can be readily obtained from state s_i and the $(i+1)$ -th input symbol $x_{n-k-i}y_{n-k-i}$. The output function is given by

$$o_{\text{rg}}(s_i, x_{i+1}y_{i+1}) = x_{1,n-k-i} \oplus y_{1,n-k-i} \oplus c_{i+1} \oplus c_i.$$

¹²We will ensure the “correct” initial state by using Corollary 3.15 with the state machine from Corollary 5.7

The above defines, up to the choice of $s_0 \in \mathbb{B}^3$, a Moore machine $T_{\text{rg}} = (\mathbb{B}^3, \cdot, \mathbb{B}^2, \mathbb{B}, t_{\text{rg}}, o_{\text{rg}})$.

Inductive application of Lemma 5.8 shows that T_{rg} with the specified inputs and starting state produces output $\text{rg}_{n-k}(\text{rg}_{n-k}^{-1}(x_{1,n-k}) + \text{rg}_{n-k}^{-1}(y_{1,n-k}) + c_0)$ (in reverse order). Using M_ξ from Corollary 5.7 to determine $\text{par}(x_{1,n-k})$, $\text{par}(y_{1,n-k})$, and $s_k(x, y)$, we can set $c_0 := 1$ iff $s_k > k$ and $c_0 := 0$ to determine s_0 . Noting that

$$(x +_{\gamma_{n,k}} y)_{1,n-k} = \begin{cases} \text{rg}_{n-k}(\text{rg}_{n-k}^{-1}(x_{1,n-k}) + \text{rg}_{n-k}^{-1}(y_{1,n-k})) & \text{if } s_k(x, y) \leq k \\ \text{rg}_{n-k}(\text{rg}_{n-k}^{-1}(x_{1,n-k}) + \text{rg}_{n-k}^{-1}(y_{1,n-k}) + 1) & \text{if } s_k(x, y) > k, \end{cases}$$

combining M_ξ and T using Corollary 3.15 yields the desired circuit. Here, the size and depth bounds follow from the fact that M_ξ has $\mathcal{O}(k)$ states and the transducer a constant number of states. \square

5.1.4 Transducer Outputting the Unary Part

It remains to output the correct trailing k bits, which are encoded in unary depending on the parity of the front string. Clearly, in parallel to outputting the leading $n - k$ bits of the output, we can determine its parity. Thus, based on the already determined value $s_k(x, y)$, it is straightforward to write the trailing k bits of the output.

Corollary 5.10. *There is a hazard-free circuit of size $2^{\mathcal{O}(k)}n$ and depth $\mathcal{O}(k \log n)$ that computes $(x +_{\gamma_{n,k}} y)_{n-k+1} \dots (x +_{\gamma_{n,k}} y)_n$ from inputs $x, y \in \gamma_{n,k}([(k+1)2^n])$.*

Proof. Knowing $s_k(x, y)$ and the parity of the $(n - k)$ -bit prefix of the output, we output $\xi_k(s_k(x, y))$ or $\bar{\xi}_k(s_k(x, y))$, respectively, using a total of $\mathcal{O}(k)$ states. \square

The following theorem is now an immediate consequence of Lemma 5.9 and Corollary 5.10.

Theorem 5.11. *There is a circuit of size $2^{\mathcal{O}(k)}n$ and depth $\mathcal{O}(k \log n)$ that implements $\lfloor k/2 \rfloor$ -recoverable addition on $\gamma_{n,k}$.*

5.2 Reducing the Dependency on k

We now reduce the maximum size of the state machines involved in the construction to $\mathcal{O}(1)$, resulting in a circuit for k -recoverable addition of size $\mathcal{O}(n + k \log k)$ and depth $\mathcal{O}(\log n)$. Roughly speaking, we can execute the addition of the unary part by sorting, as described in the introduction, and afterwards use a simple circuit to generate the correctly encoded result. However, the latter step requires to determine the parity of the reflected Gray code part of the output string, which implies that it has to be determined first — and computing this part of the output requires knowledge on the carry bit from the unary addition.

This may look like a chicken-and-egg problem, but we can break this cyclic dependence as follows. We first determine only the carry bit of the addition of the unary parts. As we will see, for stable inputs (including non-codewords) this can be done by a constant-sized state machine. Then we use the carry bit to perform the addition of the reflected Gray code parts as above, and to obtain the parity of the reflected Gray code part of the output.

Knowing the various parity bits, we know the encoding function of the inputs' unary parts as well as the output's unary part. We then use a construction based on a hazard-free multiplexer to eliminate hazards involving the bits that determine how the input is encoded, i.e., the inputs' parities and the bits on position $n - k + 2 + \lfloor k/2 \rfloor$. This simplifies the task to devising hazard-free circuits under the assumption that the respective bits are already known.

The final ingredient is to exploit that, by Observation 5.2, we may choose a different extension of $\gamma_{n,k}^{-1}$ to \mathbb{B}^n to define how our circuits behave on non-codewords, so long as the resolutions of an extended codeword (of small imprecision) are mapped to its range. We leverage this freedom by choosing an extension $\tilde{\gamma}_{n,k}^{-1}$ of $\gamma_{n,k}^{-1}$ that (after branching according to parity and the bit on position $n - k + 2 + \lfloor k/2 \rfloor$) essentially counts 1s or 0s. This enables usage of sorting as a suitable tool to obtain codewords from non-codewords, resulting in very small hazard-free circuits. Interestingly, it does not seem to be the case that $\tilde{\gamma}_{n,k}^{-1}$ permits a small state machine deciding whether the unary addition results in an overflow, thus necessitating the simultaneous use of both $\gamma_{n,k}^{-1}$ and $\tilde{\gamma}_{n,k}^{-1}$ to obtain Theorem 5.20.

To formalize the above reasoning, we fix some notation. Denote the inputs by $x^{(1)}, x^{(2)} \in \mathbb{B}^n$, set $\Sigma := \mathbb{B}^2$, and define $x \in \Sigma^n$ via $x_i := x_i^{(1)} x_i^{(2)}$. Moreover, throughout the remainder of this section, we write $x = yz$, where $y := x_1 \dots x_{n-k}$ and $z := x_{n-k+1} \dots x_n$; analogously, write $x^{(1)} = y^{(1)} z^{(1)}$ and $x^{(2)} = y^{(2)} z^{(2)}$. Our goal is to compute $x^{(1)} +_{\gamma_{n,k}} x^{(2)}$ according to Definition 5.1. We define the following functions.

1. $p^{(1)} := \text{par}(y^{(1)})$ and $p^{(2)} := \text{par}(y^{(2)})$.
2. For $i \in \{1, 2\}$, we will replace $z^{(i)}$ by $f(x^{(i)})$ such that $y^{(i)} f(x^{(i)}) = \gamma_{n,k}(\gamma_{n,k}^{-1}(x^{(i)}))$, i.e., we “correct” $x^{(i)}$ to its associated codeword. Thus,

$$f(x^{(i)}) := \begin{cases} u(z^{(i)}, z_k^{(i)}) & \text{if } \text{par}(y^{(i)}) = 0 \\ \bar{u}(z^{(i)}, z_k^{(i)}) & \text{if } \text{par}(y^{(i)}) = 1, \end{cases}$$

where u and \bar{u} are as in Definition 4.14.

3. By $c(x) := c(p^{(1)}, f(x^{(1)}), p^{(2)}, f(x^{(2)}))$ we denote the carry bit of the unary addition part, i.e.,

$$c(p^{(1)}, z^{(1)}, p^{(2)}, z^{(2)}) := \begin{cases} 1 & \text{if } p^{(1)} = p^{(2)} = 0 \text{ and } \xi_k^{-1}(z^{(1)}) + \xi_k^{-1}(z^{(2)}) > k \\ 1 & \text{if } p^{(1)} = 1 \neq p^{(2)} \text{ and } \bar{\xi}_k^{-1}(z^{(1)}) + \xi_k^{-1}(z^{(2)}) > k \\ 1 & \text{if } p^{(1)} = 0 \neq p^{(2)} \text{ and } \xi_k^{-1}(z^{(1)}) + \bar{\xi}_k^{-1}(z^{(2)}) > k \\ 1 & \text{if } p^{(1)} = p^{(2)} = 1 \text{ and } \bar{\xi}_k^{-1}(z^{(1)}) + \bar{\xi}_k^{-1}(z^{(2)}) > k \\ 0 & \text{else.} \end{cases}$$

5.2.1 Computing the Carry Bit

We now show that a state machine of constant size is sufficient to compute $c(x)$ for any input x , where we slightly modify the input by providing the unary part simultaneously in reverse. Moreover, we duplicate $x_{n-k+1+\lfloor k/2 \rfloor}$ and feed it to the state machine before the unary part of the input. That is, x' is defined via

$$x'_i := \begin{cases} x_i & \text{if } 1 \leq i \leq n - k \\ x_{n-k+1+\lfloor k/2 \rfloor} 0 & \text{if } i = n - k + 1 \\ x_{i-1} x_{2n-k-i+2} & \text{if } n - k + 2 \leq i \leq n + 1, \end{cases}$$

where the trailing 0 on index $n - k + 1$ serves as a marker to distinguish this index for the state machine, which has insufficient memory to store the current index.

Lemma 5.12. *There is a state machine of size $\mathcal{O}(1)$ that, given x' as input, computes $c(x)$.*

Proof. Our state machine first determines $p^{(1)}$ and $p^{(2)}$, which is trivial using two state bits, cf. Observation 5.3; these values then are stored without modification while processing the remaining input symbols.

Consider $\zeta_i^{(1)} := x_{n-k+1+i}^{(1)} \oplus p^{(1)}$ and $\zeta_i^{(2)} := x_{n-k+1+i}^{(2)} \oplus p^{(2)}$ for $i \in \{1, \dots, k\}$. Note that reinterpreting input symbols like this can be done on the fly, as $p^{(1)}$ and $p^{(2)}$ are known. This interpretation “translates” for all $i \in [k+1]$ from $\bar{\xi}_k(i)$ to $\xi_k(i)$ if the parity is 1 and leaves the string unmodified otherwise. Moreover, we have that $u(w, w_{1+\lfloor k/2 \rfloor}) = u(\bar{w}, \bar{w}_{1+\lfloor k/2 \rfloor})$ for all $w \in \mathbb{B}^k$. Therefore, reinterpreting the input symbol x_{n-k+1} (minus the trailing 0) in the same manner using the parities, w.l.o.g. it suffices to consider the case that $p^{(1)} = p^{(2)} = 0$ and then apply the resulting machine to inputs $\zeta_i^{(1)}$ and $\zeta_i^{(2)}$.

Under this assumption, we need to distinguish four cases, depending on $b_1 := z_{1+\lfloor k/2 \rfloor}^{(1)}$ and $b_2 := z_{1+\lfloor k/2 \rfloor}^{(2)}$. These values are fed to the machine on index $n-k+1$, marked by a trailing 0 indicating the transition from y to z . Recall that $u(w, 1)$ is the codeword of ξ_k sharing the longest suffix with $w \in \mathbb{B}^k$, while $u(w, 0)$ is the codeword sharing the longest prefix. Observe that always $u(w, w_{1+\lfloor k/2 \rfloor})_{1+\lfloor k/2 \rfloor} = w_{1+\lfloor k/2 \rfloor}$. Therefore, $w_{1+\lfloor k/2 \rfloor} = 1$ implies that $\xi_k^{-1}(u(w, w_{1+\lfloor k/2 \rfloor})) \geq 1 + \lfloor k/2 \rfloor$ and $w_{1+\lfloor k/2 \rfloor} = 0$ implies that $\xi_k^{-1}(u(w, w_{1+\lfloor k/2 \rfloor})) < 1 + \lfloor k/2 \rfloor$. We conclude that (i) $b_1 = b_2 = 1$ entails that

$$\xi_k^{-1}(u(z^{(1)}, b_1)) + \xi_k^{-1}(u(z^{(2)}, b_2)) \geq 2 \left(1 + \left\lfloor \frac{k}{2} \right\rfloor \right) > k$$

and, similarly, (ii) $b_1 = b_2 = 0$ entails that $\xi_k^{-1}(u(z^{(1)}, b_1)) + \xi_k^{-1}(u(z^{(2)}, b_2)) \leq 2\lfloor k/2 \rfloor \leq k$.

Next, consider the case that $b_1 = 1 \neq b_2 = 0$. Denote by $i_1 \in \{1, \dots, k\}$ the largest index such that $z_{i_1}^{(1)} = 1$ and by $i_2 \in \{1, \dots, k\}$ the smallest index such that $z_{i_2+1}^{(2)} = 0$. Accordingly, $\xi_k^{-1}(u(z^{(1)}, b_1)) = i_1$ and $\xi_k^{-1}(u(z^{(2)}, b_2)) = i_2$. We need to check whether $i_1 + i_2 > k$ or not. To this end, examine $z^{(1)}$ index by index in reverse order, while simultaneously doing the same for $z^{(2)}$ front to back. The above condition is satisfied if and only if we encounter the first 0 in $z^{(2)}$ (which happens in step $i_2 + 1$) at least two steps after the first 1 when traversing $z^{(1)}$ in reverse (which occurs in step $k - i_1$). Clearly, this can be tested by a state machine of constant size receiving x' as input.

Finally, there is the case that $b_1 = 0 \neq b_2 = 1$; this case is analogous to the previous one, with the roles of $z^{(1)}$ and $z^{(2)}$ exchanged. As the input symbol at index $n-k+1$ enables distinguishing the above for cases, it is straightforward to design a state machine of constant size computing $c(x)$ given input x' . \square

As we already established that we can compute the reflected Gray code part of the output using a constant-sized transducer given $c(x)$ as input, we obtain as a corollary that the $(n-k)$ -bit prefix of the output can be computed by a small circuit.

Corollary 5.13. *There is a hazard-free circuit of size $\mathcal{O}(n)$ and depth $\mathcal{O}(\log n)$ that computes $(x +_\gamma y)_1 \dots (x +_\gamma y)_{n-k}$ from inputs $x, y \in \gamma_{n,k}([(k+1)2^n])$.*

Proof. We apply Corollary 3.15 together with Lemmas 5.8 and 5.12. \square

5.2.2 Computing the Sum of the Unary Parts

We first make an observation that simplifies our task, namely that even in face of unstable inputs, it is valid to “compress” the front part of the inputs to their parity when seeking to determining the trailing k bits of the output.

Observation 5.14. *For all $x^{(1)}, x^{(2)} \in \mathbb{T}^n$ and $i \in \{1, \dots, k\}$, we have that*

$$\left(x^{(1)} (+_{\gamma_{n,k}})_u x^{(2)} \right)_{n-k+i} = \left(\text{par}_u(y^{(1)}) z^{(1)} (+_{\gamma_{k+1,k}})_u \text{par}_u(y^{(2)}) z^{(2)} \right)_{1+i}.$$

Proof. We have that $\text{res}(\text{par}_u(w)) = \text{par}(\text{res}(w))$ for any $w \in \mathbb{T}^{n-k}$. Moreover, $\text{par}(u +_{\text{rg}_{n-k}} w) = \text{par}(u) + \text{par}(w)$ for all $u, w \in \mathbb{T}^{n-k}$. Using these identities and the definition of $\gamma_{n,k}$, we get that

$$\begin{aligned}
& \left(\text{par}_u(y^{(1)})z^{(1)}(+_{\gamma_{k+1,k}})_u \text{par}_u(y^{(2)})w^{(2)} \right)_{1+i} \\
&= * \left(\text{par}(\text{res}(y^{(1)})) \text{res}(z^{(1)}) +_{\gamma_{k+1,k}} \text{par}(\text{res}(y^{(2)})) \text{res}(z^{(2)}) \right)_{1+i} \\
&= * \left(\text{res}(y^{(1)}) \text{res}(z^{(1)}) +_{\gamma_{n,k}} \text{res}(y^{(2)}) \text{res}(z^{(2)}) \right)_{n-k+i} \\
&= \left(x^{(1)}(+_{\gamma_{n,k}})_u x^{(2)} \right)_{n-k+i}. \quad \square
\end{aligned}$$

Hence, we may first compute $\text{par}_u(y^{(1)})$ and $\text{par}_u(y^{(2)})$ using a simple `xor` tree, reducing the task to the special case that $n = k + 1$. Together with Corollary 5.13 and the fact that hazard-free circuits of exponential size always exist, this already yields a circuit of size $\mathcal{O}(2^{\mathcal{O}(k)} + n)$ and depth $\mathcal{O}(k + \log n)$. However, we will reduce the dependency on k further. Due to Observation 5.14, we may assume w.l.o.g. that $n = k + 1$ in the following, as the additive $\mathcal{O}(n)$ cost in size and $\mathcal{O}(\log n)$ depth do not affect the asymptotic bounds we obtain.

Our next step simplifies the task further. Given inputs $x, y \in \mathbb{T}^n$, we use a variant of “speculative computing” [TY12, TYM14]. The idea is to replace bits in fixed positions by stable inputs and, for each such combination, to evaluate the function. In [TY12, TYM14], the authors wait for the respective inputs to stabilize and only then complete the function evaluation by using the stabilized bits as control string in a multiplexer to choose the correct pre-computed value. By replacing the multiplexer with a hazard-free multiplexer, we obtain a circuit that is protected against hazards caused by unstable inputs at the respective positions [IKL⁺18].

Lemma 5.15. *Let $f: \mathbb{B}^{n_1} \times \mathbb{B}^{n_2} \rightarrow \mathbb{B}^m$. For $s \in \mathbb{B}^{n_1}$, let C_s be a hazard-free circuit of depth d_s implementing $f(s, \cdot)$. Then there is a hazard-free circuit of size $\mathcal{O}(m2^{n_1}) + \sum_{s \in \mathbb{B}^{n_1}} |C_s|$ and depth $\mathcal{O}(n_1) + \max_{s \in \mathbb{B}^{n_1}} \{d_s\}$ implementing f .*

Proof. W.l.o.g., suppose $m = 1$; for larger m , simply perform the construction for each output bit, increasing the size of the multiplexer by factor m . Let $(y, z) \in \mathbb{T}^{n_1} \times \mathbb{T}^{n_2}$. We use a hazard-free multiplexer with n_1 control bits whose select string is y and whose selectable input labeled by $s \in \mathbb{B}^{n_1}$ receives input $f_u(s, z)$. The size of the multiplexer is $\mathcal{O}(2^{n_1})$ and its depth is $\mathcal{O}(n_1)$ [IKL⁺18, Lemma 5.1]. To compute $f_u(s, z)$, we use C_s . The size and depth bounds are immediate.

For correctness, observe that $* \text{res}(* S) = * S$ for any set S (which the $*$ operator accepts). Hence,

$$\begin{aligned}
f_u(x) &= *_{\substack{s \in \text{res}(y) \\ z' \in \text{res}(z)}} f(s, z') \\
&= *_{s \in \text{res}(y)} \text{res} \left(*_{z' \in \text{res}(z)} f(s, z') \right) \\
&= *_{s \in \text{res}(y)} \text{res}(f(s, \cdot)_u(z)) \\
&= \text{MUX}_u(y, (f(s, \cdot)_u(z))_{s \in \mathbb{B}^{n_1}}). \quad \square
\end{aligned}$$

Note that we can re-index the input as needed, i.e., for any given Boolean function, the above lemma can be applied to an arbitrary subset of input indices.

We intend to apply Lemma 5.15 to the parities and the “center” bits of the unary strings, i.e., to $x_1^{(1)}$, $x_1^{(2)}$, $x_{2+\lfloor k/2 \rfloor}^{(1)}$, and $x_{2+\lfloor k/2 \rfloor}^{(2)}$. Accordingly, we need to devise for each stable assignment to

these four values a circuit computing the hazard-free extension of the output function, where the inputs $x \in \mathbb{T}^n$ satisfy the constraint that $x_1^{(1)}$, $x_1^{(2)}$, $x_{2+\lfloor k/2 \rfloor}^{(1)}$, and $x_{2+\lfloor k/2 \rfloor}^{(2)}$ are known stable values.

To this end, we will make use of hazard-free sorting circuits. However, sorting the unary part of a non-codeword leads to different results in terms of mapping non-codewords to codewords than $\gamma_{k+1,k}^{-1}$ would yield. Accordingly, we first need to specify a different extension of $\gamma_{k+1,k}^{-1}$ to non-codewords that is consistent with the requirements of k -recoverability, i.e., Definition 4.1. For this modified extension, we can implement $(+\gamma_{k+1,k})_{\mathbf{u}}$ using sorting, and Observation 5.2 shows that, on extended codewords of sufficiently small imprecision, the result is identical to using the original extension of $\gamma_{k+1,k}^{-1}$.

Definition 5.16. We define $\tilde{\gamma}_{k+1,k}^{-1}: \mathbb{B}^{k+1} \rightarrow [2(k+1)]$ as follows.

$$\tilde{\gamma}_{k+1,k}^{-1}(v) = \begin{cases} |\{i \in \{2, \dots, 1 + \lfloor k/2 \rfloor\} \mid v_i = 1\}| & \text{if } v_1 = v_{2+\lfloor k/2 \rfloor} = 0 \\ 1 + \lfloor k/2 \rfloor + |\{i \in \{3 + \lfloor k/2 \rfloor, \dots, k+1 \mid v_i = 1\}| & \text{if } v_1 = 0 \wedge v_{2+\lfloor k/2 \rfloor} = 1 \\ k+1 + |\{i \in \{2, \dots, 1 + \lfloor k/2 \rfloor\} \mid \bar{v}_i = 1\}| & \text{if } v_1 = v_{2+\lfloor k/2 \rfloor} = 1 \\ k+2 + \lfloor k/2 \rfloor + |\{i \in \{3 + \lfloor k/2 \rfloor, \dots, k+1 \mid \bar{v}_i = 1\}| & \text{if } v_1 = 1 \wedge v_{2+\lfloor k/2 \rfloor} = 0. \end{cases}$$

Lemma 5.17. Let w be an extended codeword of $\gamma_{k+1,k}$ of imprecision at most $\lfloor k/2 \rfloor$. Then $\tilde{\gamma}_{k+1,k}^{-1}(\text{res}(w)) \subseteq r_x$.

Proof. Denote $r_w = \{i_{\min}, \dots, i_{\max}\}$, where $i_{\min} \in [2(k+1)]$ and for notational convenience we interpret all values modulo $2(k+1)$ (recall that $\gamma_{k+1,k}$ has domain $[2(k+1)]$). We distinguish four cases.

- $i_{\max} \leq k$. Then $w = 0 \mathbf{1}^{i_{\min}} \mathbf{u}^{i_{\max}-i_{\min}} \mathbf{0}^{k-i_{\max}}$. If $v_{2+\lfloor k/2 \rfloor} = 0$, then

$$\tilde{\gamma}_{k+1,k}^{-1}(v) \in \left\{ i_{\min}, \dots, \min \left\{ \left\lfloor \frac{k}{2} \right\rfloor, i_{\max} \right\} \right\}.$$

If $v_{2+\lfloor k/2 \rfloor} = 1$, then

$$\tilde{\gamma}_{k+1,k}^{-1}(v) \in \left\{ \max \left\{ i_{\min}, 1 + \left\lfloor \frac{k}{2} \right\rfloor \right\}, \dots, i_{\max} \right\}.$$

Either way, $\tilde{\gamma}_{k+1,k}^{-1}(v) \in \text{res}(w)$.

- $i_{\min} \in \{1 + \lfloor k/2 \rfloor, \dots, k\}$ and $i_{\max} > k$. As w has imprecision at most $\lfloor k/2 \rfloor$, $i_{\max} < k + 2 + \lfloor k/2 \rfloor$, $w = \mathbf{u} \mathbf{u}^{i_{\max}-(k+1)} \mathbf{1}^{k-(i_{\max}-i_{\min}-1)} \mathbf{u}^{k-i_{\min}}$ satisfies that $w_{2+\lfloor k/2 \rfloor} = 1$. If $v_1 = 0$, then $\tilde{\gamma}_{k+1,k}^{-1}(v) \in \{i_{\min}, \dots, k\}$. If $v_1 = 1$, then $\tilde{\gamma}_{k+1,k}^{-1}(v) \in \{k+1, \dots, i_{\max}\}$.
- $i_{\min} \geq k+1$ and $i_{\max} < 2(k+1)$. Then $w = 1 \mathbf{0}^{i_{\min}-(k+1)} \mathbf{u}^{i_{\max}-i_{\min}} \mathbf{0}^{2k+1-i_{\max}}$. As in the first case, we get that $\tilde{\gamma}_{k+1,k}^{-1}(v) \in r_x$, regardless of whether $v_{2+\lfloor k/2 \rfloor} = 0$ or $v_{2+\lfloor k/2 \rfloor} = 1$.
- $k+2 + \lfloor k/2 \rfloor \leq i_{\min} \leq 2k+1$. As w has imprecision at most $\lfloor k/2 \rfloor$, $i_{\max} < 2k+1 + \lfloor k/2 \rfloor$, and $w = \mathbf{u} \mathbf{u}^{i_{\max}-(2k+2)} \mathbf{1}^{k-(i_{\max}-i_{\min}-1)} \mathbf{u}^{2k+1-i_{\min}}$ satisfies that $w_{2+\lfloor k/2 \rfloor} = 0$. As in the second case, we get that $\tilde{\gamma}_{k+1,k}^{-1}(v) \in r_x$, regardless of whether $v_1 = 0$ or $v_1 = 1$.

It is easy to verify that the above case distinction is exhaustive, as $i_{\max} - i_{\min} \leq \lfloor k/2 \rfloor$ due to w having imprecision at most $\lfloor k/2 \rfloor$. \square

Equipped with $\tilde{\gamma}_{k+1,k}^{-1}$, we are in the position to apply sorting to perform the desired hazard-free addition. For simplicity, we break this up into two steps, where we first produce a unary encoding of the sum.

Lemma 5.18. *There is a circuit of size $\mathcal{O}(k \log k)$ and depth $\mathcal{O}(\log k)$ that, given two extended codewords $x^{(1)}, x^{(2)} \in \mathbb{T}^{k+1}$ of $\gamma_{k+1,k}$ with total imprecision $p := p_{x^{(1)}} + p_{x^{(2)}} \leq \lceil k/2 \rceil$, outputs the imprecision- p extended unary codeword $\ast_{i \in r_{x^{(1)}} + r_{x^{(2)}}} \xi_{4k+2}(i)$.*

Proof. We first construct for each $b \in \mathbb{B}^4$ a hazard-free circuit computing the correct output provided that $x_1^{(1)} = b_1$, $x_{2+\lfloor k/2 \rfloor}^{(1)} = b_2$, $x_1^{(2)} = b_3$, and $x_{2+\lfloor k/2 \rfloor}^{(2)} = b_4$. Fix such a $b \in \mathbb{B}^4$. We specify inputs $z_1(x), \dots, z_{4k+2}$ to a sorting network for input universe \mathbb{B} ; note that the order of the inputs is irrelevant. $x^{(1)}$ and $x^{(2)}$ each determine $2k+1$ input values in the same way, so w.l.o.g. we explain how to do this for $x^{(1)}$ only. We distinguish four cases, depending on b_1 and b_2 :

- $b_1 = b_2 = 0$. Then $\lfloor k/2 \rfloor$ inputs are given by $x_2^{(1)}, \dots, x_{1+\lfloor k/2 \rfloor}^{(1)}$, and the remaining $2k+1 - \lfloor k/2 \rfloor$ inputs are 0.
- $b_1 = 0$ and $b_2 = 1$. Then $1 + \lfloor k/2 \rfloor$ inputs are 1, $\lfloor k/2 \rfloor - 1$ inputs are given by $x_{3+\lfloor k/2 \rfloor}^{(1)}, \dots, x_{k+1}^{(1)}$, and the remaining $2k+1 - 2\lfloor k/2 \rfloor$ inputs are 0.
- $b_1 = b_2 = 1$. Then $k+1$ inputs are fixed to 1, $\lfloor k/2 \rfloor$ inputs are given by $\bar{x}_2^{(1)}, \dots, \bar{x}_{1+\lfloor k/2 \rfloor}^{(1)}$, and the remaining $k - \lfloor k/2 \rfloor$ inputs are 0.
- $b_1 = 1$ and $b_2 = 0$. Then $k+2 + \lfloor k/2 \rfloor$ inputs are 1, $\lfloor k/2 \rfloor - 1$ inputs are given by $\bar{x}_{3+\lfloor k/2 \rfloor}^{(1)}, \dots, \bar{x}_{k+1}^{(1)}$, and the remaining $k - 2\lfloor k/2 \rfloor$ inputs are 0.

The construction is analogous for $x^{(2)}$, where b_3 and b_4 take the role of b_1 and b_2 , respectively. We can see from Definition 5.16 that the number of inputs that are 1 equals $\tilde{\gamma}_{k+1,k}^{-1}(x^{(1)}) + \tilde{\gamma}_{k+1,k}^{-1}(x^{(2)})$ by construction. Thus, sorting z yields the correct output on stable values, granted that indeed $x_1^{(1)} = b_1$, $x_{2+\lfloor k/2 \rfloor}^{(1)} = b_2$, $x_1^{(2)} = b_3$, and $x_{2+\lfloor k/2 \rfloor}^{(2)} = b_4$.

Next, we ensure that the circuit doing so is hazard-free. Observe that negation translates $\bar{\xi}_n$ to ξ_n (and back) for any n , and as double negation is the identity also in Kleene logic, this translation is hazard-free. Therefore, we can safely negate inputs as stated above and then perform the sorting; the resulting circuit has a hazard if and only if the sorting part has one. However, we can perform the sorting operation using comparators consisting of an **or** gate and an **and** gate each, resulting in a sorting network without any negation gates, i.e., a monotone circuit, which is hazard-free [IKL⁺18, Lemma 4.2]. Using an asymptotically optimal sorting network, the circuit has size $\mathcal{O}(k \log k)$ and depth $\mathcal{O}(\log k)$ [AKS83].

To complete the proof, we apply Lemma 5.15 (after suitable re-indexation) to the bits $x_1^{(1)}$, $x_{2+\lfloor k/2 \rfloor}^{(1)}$, $x_1^{(2)}$, and $x_{2+\lfloor k/2 \rfloor}^{(2)}$. Thus, $n_1 = 4$ and $n_2, m \in \mathcal{O}(k)$ in the application of the lemma, and the resulting circuit satisfies the claimed bounds on size and depth. As it is hazard-free, it computes

$$\ast_{\substack{y^{(1)} \in \text{res}(x^{(1)}) \\ y^{(2)} \in \text{res}(x^{(2)})}} \xi_{4k+2}(\tilde{\gamma}_{k+1,k}^{-1}(y^{(1)}) + \tilde{\gamma}_{k+1,k}^{-1}(y^{(2)})) = \ast_{\substack{i \in \tilde{\gamma}_{k+1,k}^{-1}(\text{res}(x^{(1)})) \\ j \in \tilde{\gamma}_{k+1,k}^{-1}(\text{res}(x^{(2)}))}} \xi_{4k+2}(i+j) = \ast_{i \in r_{x^{(1)}} + r_{x^{(2)}}} \xi_{4k+2}(i).$$

on inputs of imprecision at most $\lceil k/2 \rceil$, where in the last step we used that $\tilde{\gamma}_{k+1,k}^{-1}(\text{res}(x^{(1)})) = r_{x^{(1)}}$ and $\tilde{\gamma}_{k+1,k}^{-1}(\text{res}(x^{(2)})) = r_{x^{(2)}}$ by Lemma 5.17. \square

As our final step, we need to translate the obtained unary encoding to the encoding given by $\gamma_{k+1,k}$. This is straightforward, because each input bit affects a single output bit only, by flipping it on up-count corresponding to it. This permits a simple implementation based on **xor** trees, cf. [FKLP17, Lemma 3.5].

Lemma 5.19. *There is a hazard-free circuit of size $\mathcal{O}(n + k \log k)$ and depth $\mathcal{O}(\log n)$ that, given $x, y \in \mathbb{B}^n$, computes $(x +_{\gamma_{n,k}} y)_{n-k+1} \dots (x +_{\gamma_{n,k}} y)_n$.*

Proof. By Observation 5.14, we may w.l.o.g. assume that $n = k + 1$ (adding $\mathcal{O}(n)$ and $\mathcal{O}(\log n)$ to the size and depth of the circuit, respectively). By Lemma 5.18, we can obtain $*_{i \in r_x + r_y} \xi_{4k+2}(i)$ by a circuit of size $\mathcal{O}(k \log k)$ and depth $\mathcal{O}(\log k)$.

Next, note that $\gamma_{k+1,k}(0) = 0^{k+1}$ and for $i \in \{2, \dots, k + 1\}$, the bit at index i is “toggled” on up-counts $j(k + 1) + i$, where $j \in [4]$. Thus, $\gamma_{k+1,k}(a)_i = \bigoplus_{j \in [4]} \xi_{4k+2}(a)_{j(k+1)+i}$ for all such i and $a \in [4k + 2]$. Because the parity (i.e., **xor**) always depends on each unspecified input bit, even when fixing all other ones, this equality extends to Kleene logic. Therefore, adding $\mathcal{O}(k)$ **xor** gates and increasing the depth by $\mathcal{O}(1)$, we arrive at the desired circuit computing $(x +_{\gamma_{n,k}} y)_{n-k+1} \dots (x +_{\gamma_{n,k}} y)_n$. \square

Together with Corollary 5.13, we arrive at a circuit that is trivially optimal depth and for all $k \in \mathcal{O}(n/\log n)$ also in size.

Theorem 5.20. *There is a circuit of size $\mathcal{O}(n + k \log k)$ and depth $\mathcal{O}(\log n)$ that implements $\lfloor k/2 \rfloor$ -recoverable addition on $\gamma_{n,k}$.*

Proof. Follows immediately from Corollary 5.13 and Lemma 5.19. \square

References

- [AB87] Noga Alon and Ravi B. Boppana. The Monotone Circuit Complexity of Boolean Functions. *Combinatorica*, 7(1):1–22, 1987.
- [AKS83] M. Ajtai, J. Komlós, and E. Szemerédi. An $\mathcal{O}(n \log n)$ Sorting Network. In *15th Symposium on Theory of Computing (STOC)*, 1983.
- [BEI01] J. Brzozowski, Z. Esik, and Y. Iland. Algebras for Hazard Detection. In *Proc. 31st International Symposium on Multiple-Valued Logic*, 2001.
- [BLM18] Johannes Bund, Christoph Lenzen, and Moti Medina. Optimal Metastability-Containing Sorting Networks. In *Proc. Design, Automation and Test in Europe (DATE)*, 2018.
- [Cal58] Samuel H. Caldwell. *Switching Circuits and Logical Design*. John Wiley & Sons Inc, 1958.
- [DDT78] Marc Davio, Jean-Pierre Deschamps, and André Thaysé. *Discrete and Switching Functions*. McGraw-Hill International Book Co., 1978.
- [Eic65] E. B. Eichelberger. Hazard Detection in Combinational and Sequential Switching Circuits. *IBM J. Res. Dev.*, 9(2):90–99, 1965.
- [FFL18] Stephan Friedrichs, Matthias Függer, and Christoph Lenzen. Metastability-Containing Circuits. *IEEE Transactions on Computers*, 67, 2018.
- [FKLP17] Matthias Függer, Attila Kinali, Christoph Lenzen, and Thomas Polzer. Metastability-Aware Memory-Efficient Time-to-Digital Converters. In *23rd IEEE International Symposium on Asynchronous Circuits and Systems*, 2017.
- [Got49] M. Goto. Application of logical mathematics to the theory of relay networks (in Japanese). *J. Inst. Elec. Eng. of Japan*, 64(726):125–130, 1949.

- [HOI⁺12] W. Hu, J. Oberg, A. Irturk, M. Tiwari, T. Sherwood, D. Mu, and R. Kastner. On the Complexity of Generating Gate Level Information Flow Tracking Logic. *IEEE Transactions on Information Forensics and Security*, 7(3):1067–1080, 2012.
- [Huf57] David A. Huffman. The Design and Use of Hazard-Free Switching Networks. *Journal of the ACM*, 4(1):47–62, 1957.
- [IKL⁺18] C. Ikenmeyer, B. Komarath, C. Lenzen, V. Lysikov, A. Mokhov, and K. Sreenivasaiiah. On the complexity of hazard-free circuits. In *Symposium on the Theory of Computing (STOC)*, pages 878–889, 2018.
- [Kau58] W. H. Kautz. Unit-distance error-checking codes. *IRE Transactions on Electronic Computers*, EC-7(2):179–180, June 1958.
- [Kin08] David J. Kinniment. *Synchronization and Arbitration in Digital Systems*. Wiley, 2008.
- [Kle38] Stephen Cole Kleene. On Notation for Ordinal Numbers. *The Journal of Symbolic Logic*, 3(4):150–155, 1938.
- [Kle52] Stephen Cole Kleene. *Introduction to Metamathematics*. North Holland, 1952.
- [Kör66] Stephan Körner. *Experience and Theory: An Essay in the Philosophy of Science*. International library of philosophy and scientific method. Routledge & Kegan Paul, London, 1966.
- [LF80] Richard E Ladner and Michael J Fischer. Parallel Prefix Computation. *Journal of the ACM (JACM)*, 27(4):831–838, 1980.
- [Mar77] Leonard R. Marino. The Effect of Asynchronous Inputs on Sequential Network Reliability. *IEEE Transactions on computers*, (11):1082–1090, 1977.
- [Mar81] Leonard Marino. General Theory of Metastable Operation. *IEEE Transactions on Computers*, C-30(2):107–115, 1981.
- [MG76] K. Mehlhorn and Z. Galil. Monotone Switching Circuits and Boolean Matrix Product. *Computing*, 16(1):99–111, Mar 1976.
- [Moo56] Edward F. Moore. Gedanken-Experiments on Sequential Machines. In Claude Shannon and John McCarthy, editors, *Automata Studies*, pages 129–153. Princeton University Press, Princeton, NJ, 1956.
- [MS77] Florence J. MacWilliams and Neil J. A. Sloane. *The Theory of Error Correction Codes*. Elsevier, 1977.
- [Pat75] Michael S. Paterson. Complexity of Monotone Networks for Boolean Matrix Product. *Theoretical Computer Science*, 1(1):13–20, 1975.
- [Pec76] Miroslav Pechoucek. Anomalous Response Times of Input Synchronizers. *IEEE Transactions on Computers*, 100(2):133–139, 1976.
- [SC79] M. J. Stucki and J. R. Cox. Synchronization Strategies. In *Proceedings of the Caltech Conference On Very Large Scale Integration*, pages 375–393, 1979.

- [SKK⁺02] Premkishore Shivakumar, Michael Kistler, Stephen W. Keckler, Doug Burger, and Lorenzo Alvisi. Modeling the Effect of Technology Trends on the Soft Error Rate of Combinational Logic. In *International Conference on Dependable Systems and Networks (DSN)*, pages 389–398, 2002.
- [SL15] Earl E. Swartzlander and Carl E. Lemonds, editors. *Computer Arithmetic*, volume I–III. World Scientific Publishing Co, 2015.
- [Tar88] É. Tardos. The Gap between Monotone and Non-monotone Circuit Complexity is Exponential. *Combinatorica*, 8(1):141–142, 1988.
- [TWM⁺09] Mohit Tiwari, Hassan M.G. Wassel, Bitu Mazloom, Shashidhar Mysore, Frederic T. Chong, and Timothy Sherwood. Complete Information Flow Tracking from the Gates Up. *SIGARCH Comput. Archit. News*, 37(1):109–120, 2009.
- [TY12] G. Tarawneh and A. Yakovlev. An RTL Method for Hiding Clock Domain Crossing Latency. In *Electronics, Circuits, and Systems (ICECS)*, pages 540–543, 2012.
- [TYM14] Ghaith Tarawneh, Alex Yakovlev, and Terrence S. T. Mak. Eliminating Synchronization Latency Using Sequenced Latching. *IEEE Transactions on VLSI Systems*, 22(2):408–419, 2014.
- [Wor77] E. G. Wormald. A Note on Synchronizer or Interlock Maloperation. *IEEE Transactions on Computers*, C-26(3):317–318, 1977.

This figure "transducer.png" is available in "png" format from:

<http://arxiv.org/ps/1811.12369v1>