

Conserved Energy Functions for Cellular Automata: Finding Nontrivials Faster Through a Complete Theory of the Trivials

LEEMON BAIRD^{1*}, BARRY FAGIN^{2†}

¹ *Geometric Modeling and Scientific Visualization Research Center, King Abdullah University of Science and Technology, Saudi Arabia, and also the Academy Center for Cyberspace Research, Department of Computer Science, US Air Force Academy, Colorado Springs, Colorado USA 80840*

² *Department of Computer Science, US Air Force Academy, Colorado Springs, Colorado USA 80840*

Received 31 October 2010

Key words: nontrivial, trivial, conserved, energy, function, cellular, automata, tensors, dimensions

Conserved quantities indicate fundamental physical properties of systems, and are therefore eagerly sought after in science. Conservation of energy and conservation of angular momentum, for example, are two fundamental principles that offer profound insights into the nature of the world we live in.

Cellular automata, as models of physical systems, can exhibit conserved functions of relevance to the system under study. Number conservation is a simple example, but far more sophisticated ones can be discovered with appropriate algorithmic techniques. Cellular automata with identical conserved functions are, in some sense, closely related to one another. Thus conserved functions can be used to classify cellular automata, and to identify connections between seemingly unrelated systems.

* email: Leemon@Leemon.com

† email: barry.fagin@usafa.edu

There are known algorithms to find all conserved energy functions of a given order for a given cellular automaton. They are exponentially slow, but can be made faster by eliminating the trivial conserved functions that classify all states as having the same energy. So we must find the basis set for the trivials. That problem is nontrivial.

We present here the first proofs for the basis set for all trivial conserved functions in the general case, and use this to derive a number of optimizations for reducing time and memory for the discovery of nontrivials.

We use these results to show the Game of Life has no nontrivials with any rectangular energy window containing 13 or fewer cells. Other 2D automata, however, do have nontrivials. We give the complete list of those functions for all life-like automata (i.e. 2D, 2-color, 3×3 neighborhood, outer totalistic) with rectangular energy windows with 9 or fewer cells, and discuss patterns we have observed.

1 PRELIMINARIES: BASIC DEFINITIONS

We consider a problem described in [6]. We consider cellular automata with k states in n dimensions. The *neighborhood* of a cellular automaton is the region of surrounding cells used to determine the next state of a given cell. The *window* of an energy function for a cellular automaton is the region of adjacent cells that contribute to the function. Both neighborhoods and windows are n -dimensional tensors, with the size of each dimension specified as a positive integer. Given the size of such a tensor, it is useful to define the following 3 sets of tensors:

Definition 1.1. Cellular automata are composed of cells, each of which is in one of k states (or *colors*) at any given time. The set \mathcal{C} is the set of such colors, and the set \mathcal{C}_* is that set augmented with another color named $*$. The color $*$ denotes a special color that is not used by the cellular automaton, but which simplifies our proofs. It is explained in more detail in the pages that follow.

$$\mathcal{C} = \{0, 1, 2, \dots, k - 1\} \tag{1.1}$$

$$\mathcal{C}_* = \mathcal{C} \cup \{*\} \tag{1.2}$$

It is sometimes useful to choose one color to be treated specially. In all such cases, the color 0 will be chosen.

Definition 1.2. An n -dimensional *cellular automaton rule* is a function R that gives the color of a given cell on the next time step as a function of a neighborhood of cells centered on that cell on the current time step. For symmetric neighborhoods, the neighborhood is an n -dimensional tensor of size $w_1 \times \cdots \times w_n$, where each w_i is an odd, positive integer.

$$R : \mathcal{C}^{w_1 \times \cdots \times w_n} \rightarrow \mathcal{C} \quad (1.3)$$

For a non-symmetric neighborhood, we can consider a larger, symmetric neighborhood that contains it, with an R that ignores the extra cells. Therefore, without loss of generality, we will assume the neighborhood is symmetric.

Definition 1.3. An n -dimensional *cellular automaton* is an n -dimensional tensor whose elements are in \mathcal{C} , and which is updated on each time step according to a cellular automaton rule R , applied to every cell in parallel. The rule is a function applied to each cell and its neighbors, where neighbors wrap toroidally (i.e. the top edge is considered adjacent to the bottom, the left edge is adjacent to the right, and so on for each dimension).

Definition 1.4. A *life-like* cellular automaton is one that is 2-dimensional, has 2 colors, has a 3×3 neighborhood, and is outer totalistic, meaning that the color of a cell on the next step is a function only of its color on the current step and the total number of its 8 neighbors that are of each color on the current step. The name derives from Conway's Game of Life, which is an example of this type of CA.

A cellular automaton as a whole works by applying R to every cell in parallel, with toroidal wrapping, giving a resulting universe the same size as the original. But it is also useful to consider the result of applying R to some smaller region, without wrapping, giving a result that is an even smaller region (smaller because the new values for cells on the edges can't be calculated). So we define the function T , which applies R to tensors that are a subset of the universe:

Definition 1.5. The *regional successor function*, R_{reg} , advances a region within a cellular automaton one time step by applying a rule R to all the windows within a region M of size $s_1 \times \cdots \times s_n$

$$R_{reg} : (\mathcal{C}^{w_1 \times \cdots \times w_n} \rightarrow \mathcal{C}) \times \mathcal{C}^{s_1 \times \cdots \times s_n} \rightarrow \mathcal{C}^{(s_1 - w_1 + 1) \times \cdots \times (s_n - w_n + 1)}$$

which is defined as:

$$R_{reg}(R, M) = M' \text{ where } M'_{i_1, \dots, i_n} = R(M_{\{i_1, \dots, i_1 + w_1 - 1\}, \dots, \{i_n, \dots, i_n + w_n - 1\}}) \quad (1.4)$$

The subscripts for M' range over all values for which the subscripts for M on the right side will be legal (i.e. not larger than its size).

Note that the ellipses on the right side of the equation are used in two different ways. Each element of the result comes from applying the R successor function to only that portion of the M tensor where the elements have a first coordinate in the range $\{i_1, \dots, i_1 + w_1 - 1\}$, and a second coordinate in the range $\{i_2, \dots, i_2 + w_2 - 1\}$, and so on up to an n th coordinate in the range $\{i_n, \dots, i_n + w_n - 1\}$.

There are various ways in which energy functions for cellular automata might be defined. We will use the following, common definition:

Definition 1.6. An *energy function* is a function $f : \mathcal{C}^{s_1 \times \dots \times s_n} \rightarrow \mathbb{R}$ that assigns a real number to a window of size $s_1 \times \dots \times s_n$ within a cellular automaton given its contents.

Definition 1.7. The *order* (or *size*) of an energy function with n -dimensional window of size $s_1 \times s_2 \times s_3 \times \dots \times s_n$ is defined to be the product of the sizes $s_1 s_2 s_3 \dots s_n$.

Definition 1.8. The *total energy* $E : \mathcal{C}^{u_1 \times \dots \times u_n} \rightarrow \mathbb{R}$ of a given state U of an entire cellular automaton universe with $u_1 \times \dots \times u_n$ cells, with respect to a given energy function f , is

$$E(U) = \sum_W f(U_W) \quad (1.5)$$

where U is the universe state for a cellular automaton, W is the position of the energy window within that universe, and U_W is that window within the universe, which wraps toroidally at the edges of the universe.

Definition 1.9. A *conserved energy function* (or a *conserved function*) for a given cellular automaton rule is an energy function that for a universe of any size, and for any given state of that universe, will assign the same total energy to that universe for both that state and its successor.

Definition 1.10. A *trivial conserved energy function* (or a *trivial*) is an energy function that for a universe of any size, will assign the same total energy to

that universe regardless of its state. A *nontrivial conserved energy function* (or a *nontrivial*) for a given cellular automaton rule is a conserved energy function that is not trivial.

Definition 1.11. Given n positive integers s_1, \dots, s_n defining the size of an n -dimensional tensor, the set $\mathcal{B}(s_1, \dots, s_n)$ is the set of all tensors over \mathcal{C} of that size. This set is partitioned into two sets, $\mathcal{Z}(s_1, \dots, s_n)$, the *zero-sided tensors*, which have at least one side that contains the origin element and is filled entirely with zero elements, and $\bar{\mathcal{Z}}(s_1, \dots, s_n)$, the *non-zero-sided tensors*, which do not have such a side. The *origin element* is the element of the tensor at location $(1, 1, \dots, 1)$. The $\mathcal{B}_0(s_1, \dots, s_n)$ is defined to be the subset of \mathcal{B} where all the elements of the tensor after the middle one (in row major order) are 0.

$$\mathcal{B}(s_1, \dots, s_n) = \mathcal{C}^{s_1 \times \dots \times s_n} \quad (1.6)$$

$$\mathcal{Z}(s_1, \dots, s_n) = \{M \in \mathcal{B}(s_1, \dots, s_n) \mid \exists i \forall j \forall s_j M_{s_1, \dots, s_{i-1}, 1, s_{i+1}, \dots, s_n} = 0\} \quad (1.7)$$

$$\bar{\mathcal{Z}}(s_1, \dots, s_n) = \mathcal{B}(s_1, \dots, s_n) \setminus \mathcal{Z}(s_1, \dots, s_n) \quad (1.8)$$

So in 1 dimension, the zero-sided vectors are those with a 0 as the first element. In 2 dimensions, the zero-sided matrices are those with a top row of all zeros, or a leftmost column of all zeros, or both. In n dimensions, the zero-sided tensors are those for which all the elements on side i are zero, for some i , where the side is an n -dimensional slice including the origin.

It is useful to define a matching function H that can be used in the construction of various functions over these tensors. The function returns 1 iff two tensors have elements that match, where the * symbol is treated as matching any color.

Definition 1.12. Given n -dimensional tensors over \mathcal{C}_* , the function $H : \mathcal{C}_*^{s_1 \times \dots \times s_n} \times \mathcal{C}_*^{s_1 \times \dots \times s_n} \rightarrow \{0, 1\}$ is defined as

$$H(A, B) = \begin{cases} 1 & \text{if } \forall i \forall s_i A_{s_1, \dots, s_n} = B_{s_1, \dots, s_n} \\ & \vee A_{s_1, \dots, s_n} = * \\ & \vee B_{s_1, \dots, s_n} = * \\ 0 & \text{otherwise} \end{cases} \quad (1.9)$$

Given an n -dimensional tensor, it is useful to unwrap it into a 1D string of characters. This will be done in *row major order*. For matrices, this means

the elements will be read from left to right across the top row, then left to right across the second row, and so on down to the bottom row. Tensors of other dimensionalities are unwrapped similarly, with the last dimension changing most quickly, and the first dimension changing most slowly. It is useful to have a function $V_{num}(T)$ that unwraps the elements of tensor T , then converts the resulting string to an integer by treating it as a number in base k , with the first element being the most significant digit, and the last being the least significant.

Definition 1.13. An n -dimensional tensor M with elements in \mathcal{C} can be converted to an integer by the function $V_{num} : \mathcal{C}^{s_1 \times \dots \times s_n} \rightarrow \mathbb{N}$, which treats the elements of the tensor as digits base k , where the elements are taken in row major order, treating the first as the least significant digit, and the last as the most significant.

$$V_{num}(M) = \sum_{i_1=1}^{s_1} \sum_{i_2=1}^{s_2} \dots \sum_{i_n=1}^{s_n} M_{i_1, i_2, \dots, i_n} \prod_{j=1}^n k^{(i_j-1) \prod_{m=j+1}^n s_m} \quad (1.10)$$

For this definition, the rightmost product is understood to be 1 for all cases where the lower bound exceeds the upper.

Definition 1.14. An n -dimensional tensor with elements in \mathcal{C} can be converted to a binary vector by the function $V_t : \mathcal{C}^{s_1 \times \dots \times s_n} \rightarrow \{0, 1\}^{(k^{s_1 s_2 \dots s_n})}$, which is defined as

$$V_t(M) = v \text{ where } v_i = \begin{cases} 1 & \text{if } i = V_{num}(M) + 1 \\ 0 & \text{otherwise} \end{cases} \quad (1.11)$$

The vector $V_t(M)$ has one element for each possible color pattern for a tensor of the same size as M . That vector will be all zeros, except for a 1 in the position corresponding to the pattern M .

Definition 1.15. A function $f : \mathcal{C}^{s_1 \times \dots \times s_n} \rightarrow \mathbb{R}$ can be converted to a real vector with $k^{s_1 s_2 \dots s_n}$ elements by the function $V : (\mathcal{C}^{s_1 \times \dots \times s_n} \rightarrow \mathbb{R}) \rightarrow \mathbb{R}^{k^{s_1 s_2 \dots s_n}}$, which is defined as

$$V(f) = \sum_{M \in \mathcal{B}(s_1, \dots, s_n)} f(M) \cdot V_t(M) \quad (1.12)$$

This vector is a convenient way to represent an energy function. It completely specifies the energy function, by listing the output of the function for every possible input. We will define various classes of energy functions by simultaneous linear equations, treating the elements of this vector as the variables.

Note that the energy function window is independent of the CA neighborhood. Energy functions can be defined over regions different from the scope of the transition rule of the CA. Our work with 1D CAs in [2], for example, has identified conserved energy functions with windows with sizes as small as 1×1 and as large as 1×14 , for CAs that have neighborhoods of size 1×3 .

Definition 1.16. Given tensor M of size $m_1 \times \cdots \times m_n$, which is a region within an n -dimensional universe, and given an energy window size of $s = (s_1, \dots, s_n)$, a vector representing the total energy of all energy windows that fit within M can be found with the function

$$e : \mathbb{N}^n \times \mathcal{C}^{m_1 \times \cdots \times m_n} \rightarrow \mathbb{N}^{k^{s_1 s_2 \cdots s_n}}$$

which is defined as

$$e(s, M) = \sum_{i_1=1}^{m_1-s_1+1} \sum_{i_2=1}^{m_2-s_2+1} \cdots \sum_{i_n=1}^{m_n-s_n+1} V(M_{\{i_1, \dots, i_1+s_1-1\}, \dots, \{i_n, \dots, i_n+s_n-1\}}) \quad (1.13)$$

The $e(s, M)$ function slides the energy window to all possible positions that fit entirely within the tensor M , and finds the energy at each position. It then sums all the energies coming from identical patterns, and constructs a vector with the total energy derived from each possible pattern. The sum of the elements of this vector would simply be the total energy of M . But it is useful to maintain the vector of separate values when generating sets of linear equations that define the trivials, the nontrivials, or the conserved functions.

Definition 1.17. For a positive integer n , the function $N : \mathbb{N}^n \rightarrow \mathbb{N}$ is defined as

$$N(s_1, \dots, s_n) = \sum_{b=1}^{2^n-1} k^{\prod_i (s_i - b_i)} (-1)^{1 + \sum_i b_i} \quad (1.14)$$

where b_i is the i th bit of integer b written in binary, with bit 1 being least significant and bit n being most.

In 1 and 2 dimensions this reduces to:

$$N(c) = k^{c-1} \quad (1.15)$$

$$N(r, c) = k^{(r-1)c} + k^{r(c-1)} - k^{(r-1)(c-1)} \quad (1.16)$$

It will be proved below that this gives the cardinality of many of the sets that will be considered here. It equals the number of zero-sided tensors of a

given size, the number of trivials, and the number of unit complements. And when subtracted from a simple power of 2, it yields the number of non-zero-sided tensors, the number of equations defining the conserved functions, and the number of equations defining the nontrivials. These terms are defined and the counts proved below.

Definition 1.18. In n dimensions, the seven transforms that operate on tensors of size $s_1 \times \cdots \times s_n$ are:

$$P_C : \mathcal{C}_*^{s_1 \times \cdots \times s_n} \rightarrow \mathcal{C}_*^{s_1 \times \cdots \times s_n} \quad (1.17)$$

$$P_* : \{1, 2, \dots, n\} \times \mathcal{C}_*^{s_1 \times \cdots \times s_n} \rightarrow \mathcal{C}_*^{s_1 \times \cdots \times s_n} \quad (1.18)$$

$$P_{rot} : \{1, 2, \dots, n\} \times \mathcal{C}_*^{s_1 \times \cdots \times s_n} \rightarrow \mathcal{C}_*^{s_1 \times \cdots \times s_n} \quad (1.19)$$

$$P_{LD} : \{1, 2, \dots, n\} \times \mathcal{C}_*^{s_1 \times \cdots \times s_n} \rightarrow \mathcal{C}_*^{s_1 \times \cdots \times s_n} \quad (1.20)$$

$$P_{RD} : \{1, 2, \dots, n\} \times \mathcal{C}_*^{s_1 \times \cdots \times s_n} \rightarrow \mathcal{C}_*^{s_1 \times \cdots \times s_n} \quad (1.21)$$

$$P_L : \mathcal{C}_*^{s_1 \times \cdots \times s_n} \rightarrow \mathcal{C}_*^{s_1 \times \cdots \times s_n} \quad (1.22)$$

$$P_R : \mathcal{C}_*^{s_1 \times \cdots \times s_n} \rightarrow \mathcal{C}_*^{s_1 \times \cdots \times s_n} \quad (1.23)$$

and are defined to be:

$$P_C(M) = M' \text{ where } M'_{i_1, \dots, i_n} = \begin{cases} 0 & \text{if } \forall j \ i_j = \lceil s_j/2 \rceil \\ M_{i_1, \dots, i_n} & \text{otherwise} \end{cases} \quad (1.24)$$

$$P_*(d, M) = M' \text{ where } M'_{i_1, \dots, i_n} = \begin{cases} * & \text{if } i_d = 1 \\ M_{i_1, \dots, i_n} & \text{otherwise} \end{cases} \quad (1.25)$$

$$P_{rot}(d, M) = M' \text{ where } M'_{i_1, \dots, i_n} = M_{i_1, \dots, i_{d-1}, 1+(i_d \bmod s_d), i_{d+1}, \dots, i_n} \quad (1.26)$$

$$P_{LD}(d, M) = \begin{cases} P_*(d, M) & \text{if } \forall j \forall i_j \ M_{i_1, \dots, i_{d-1}, 1, i_{d+1}, \dots, i_n} \in \{0, *\} \\ M & \text{otherwise} \end{cases} \quad (1.27)$$

$$P_{RD}(d, M) = \begin{cases} P_{rot}(d, P_*(d, M)) & \text{if } \forall j \forall i_j \ M_{i_1, \dots, i_{d-1}, 1, i_{d+1}, \dots, i_n} \in \{0, *\} \\ M & \text{otherwise} \end{cases} \quad (1.28)$$

$$P_L(M) = P_{LD}(1, P_{LD}(2, \dots P_{LD}(n, M) \dots)) \quad (1.29)$$

$$P_R(M) = P_{RD}(1, P_{RD}(2, \dots P_{RD}(n, M) \dots)) \quad (1.30)$$

The function $P_{rot}(d, M)$ rotates the elements of tensor M along dimension d , so that one side that included the origin moves to the opposite side. The function P_C sets the central element to zero. The function P_L transforms a zero-sided tensor by replacing the 0 elements on each all-zero, origin-containing side with * elements. And P_R does the same, then rotates it so each modified side moves to the opposite side. The functions P_* , P_{LD} , and P_{RD} are only used here to define the other functions, and won't be used again.

The following gives three examples of P_L and P_R applied to zero-sided matrices of size 3×5 . In each example, M is a zero-sided matrix, where the all-zero side is on the left, top, and both, respectively:

$$M = \begin{array}{ccccc} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \quad P_L(M) = \begin{array}{ccccc} * & 1 & 1 & 1 & * \\ * & 0 & 1 & 0 & * \\ 0 & 1 & 0 & 0 & 0 \end{array} \quad P_R(M) = \begin{array}{ccccc} 0 & 1 & 1 & 1 & * \\ 0 & 1 & 0 & 1 & * \\ 0 & 1 & 0 & 0 & 0 \end{array} \quad (1.31)$$

$$M = \begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array} \quad P_L(M) = \begin{array}{ccccc} * & * & * & * & * \\ * & 0 & 1 & 1 & * \\ 0 & 0 & 0 & 0 & 1 \end{array} \quad P_R(M) = \begin{array}{ccccc} 0 & 0 & 1 & 1 & * \\ 0 & 0 & 0 & 1 & * \\ * & * & * & * & * \end{array} \quad (1.32)$$

$$M = \begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array} \quad P_L(M) = \begin{array}{ccccc} * & * & * & * & * \\ * & 0 & 0 & 0 & * \\ * & 0 & 1 & 1 & 0 \end{array} \quad P_R(M) = \begin{array}{ccccc} 0 & 0 & 0 & * & * \\ 0 & 0 & 1 & 1 & * \\ * & * & * & * & * \end{array} \quad (1.33)$$

Definition 1.19. The function $P_Z : C^{s_1 \times \dots \times s_n} \rightarrow C^{(2s_1-1) \times \dots \times (2s_n-1)}$ takes a small n -dimensional tensor and pads it with zero elements on many of its sides to create a large n -dimensional tensor. In each dimension, if the small tensor was of size s_i in that dimension, then the large tensor will be of size $2s_i - 1$ in that dimension. The zero elements are added in such a way that the last nonzero element in the original tensor becomes the center element in the new tensor.

For example,

$$P_Z \left(\begin{array}{ccccc} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) = \begin{array}{ccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \quad (1.34)$$

In this 2D example, the small matrix M is of size 3×5 , and $P_Z(M)$ is of size $(2 \cdot 3 - 1) \times (2 \cdot 5 - 1) = 5 \times 9$. Note that this M happens to have 4 nonzero elements, arranged in a sort of V shape. If the elements of M are read in row major order (i.e. left to right across the top row, then left to right on the second row, etc.), then the last nonzero element to be read is the bottom of the V. The P_Z function pads with zeros in such a way as to put that last nonzero element in the exact center of a large matrix of the correct size, padded with enough zeros on each side to accomplish that.

Definition 1.20. For a given tensor size $s_1 \times \dots \times s_n$, the set \mathcal{T} is defined to be the set of all trivial conserved energy functions with energy windows of that size.

That defined \mathcal{T} to be a set of functions. But it is often more convenient to convert each of those functions to a vector, by building a vector of the coefficients of the H terms in the definition of f_M . This is what the V function in Definition 1.15 does. The definition of f_M above has only two H terms, but these are for M tensors that can have the $*$ symbol as elements. The $*$ symbol can be eliminated by expanding each such H term into multiple terms, for each possible combination of colors assigned to the $*$ symbols. This is what the V function does, and so it can give vectors with more than 2 nonzero elements. The set \mathcal{T} is a space of functions, and \mathcal{T}_c is the space of coefficient vectors corresponding to each of those functions.

Definition 1.21. For a given tensor size $s_1 \times \dots \times s_n$, the set \mathcal{T}_c is defined to be the following set of real vectors:

$$\mathcal{T}_c(s_1, \dots, s_n) = \{V(f) \mid f \in \mathcal{T}\} \quad (1.35)$$

2 THEORETICAL RESULTS

The following theorems derive a basis set for the trivials, then show how that can speed the search for nontrivials by allowing the energy for certain patterns to be set to zero without loss of generality. It is interesting that the concept of the *set of zero sided tensors* arises several times and plays a crucial role in the theory of the trivials.

Theorem 2.1. *The cardinality of the set $\mathcal{Z}(s_1, \dots, s_n)$ is $N(s_1, \dots, s_n)$.*

Proof: Definition 1.17 defines N as a sum over integers b . This proof will group those integers by Hamming weight, then apply the inclusion-exclusion principle. Let G_w be the set of all integers in $\{0, \dots, 2^n - 1\}$ that have a Hamming weight of w (i.e. have exactly w bits equal to 1 when written in binary). Then the sum from Definition 1.17 can be broken into sets of terms

with equal Hamming weight:

$$N(s_1, \dots, s_n) = \sum_{b=1}^{2^n-1} k^{\prod_i (s_i - b_i)} (-1)^{1 + \sum_i b_i} \quad (2.1)$$

$$= \sum_{w=1}^n \sum_{b \in G_w} k^{\prod_i (s_i - b_i)} (-1)^{1 + \sum_i b_i} \quad (2.2)$$

$$= \sum_{w=1}^n (-1)^{w+1} \sum_{b \in G_w} k^{\prod_i (s_i - b_i)} \quad (2.3)$$

For a given b , the exponent on the k represents the number of unconstrained elements remaining in the tensor after certain sides have been constrained to be entirely zero. The 1 bits in b select which sides are constrained to be all zero. So if the 1st, 3rd, and 7th bits of b equal 1, then the size of the tensor is decremented in the 1st, 3rd, and 7th dimensions, which reflects that the origin-containing sides in the 1st, 3rd, and 7th dimensions have all elements constrained to equal zero.

The expression of k raised to the number of unconstrained elements gives a count of how many tensors over \mathcal{C} exist, subject to the constraint that certain sides must have all elements equal to 0 (where the sides are chosen by the 1 bits in b).

So the entire double sum is a sum of counts of tensors that have been filled with bits in various ways. Some tensors are included more than once in that sum, because one side might be constrained to be all zeros, or that side might be unconstrained, but have every element chosen to be zero, which yields the same tensor. So there is a need to subtract off those that are counted multiple times.

The power of -1 means that some tensors are added to the total, while some are subtracted from the total. Since $\mathcal{Z}(s_1, \dots, s_d)$ is the set of all tensors with at least one side set to zero, it must be shown that tensors with no all-zero sides are not included in the count. And it must be shown that each tensor with at least one all-zero side is counted exactly once (i.e. will be positively counted exactly one more time than it is negatively counted). These two cases will now be shown.

The first case is obvious. If a tensor has no all-zero sides, then it would only be counted by a term of $b = 0$, which has a Hamming weight of $w = 0$. But the sum is for $w > 0$, so such tensors are never counted.

For the second case, consider a tensor M with exactly z of its origin-containing sides having all zero elements. The double sum will count that

tensor several times for each value of w . When $w = z$, it is counted exactly once, with a b that has w of its bits set to 1, corresponding to the origin-containing, all-zero sides of M . When $w = z - 1$, it is counted w times, where b has only $z - 1$ bits set to 1, corresponding to $z - 1$ of the z all-zero sides, and with the remaining bits filled in to match M . When $w = z - 2$, it is counted $\binom{z}{2}$ times, and in general, for each $w \leq z$ it is counted $\binom{z-w}{z-w} = \binom{z}{w}$ times, once for each way of setting w of the bits of b to 1, corresponding to w of the z all-zero, origin-containing sides of M . Each of these counts has a coefficient that is a power of -1, so the total contribution of M to the count is

$$\sum_{w=1}^z (-1)^{w+1} \binom{z}{w} \quad (2.4)$$

$$= -(-1)^{0+1} \binom{z}{0} + \sum_{w=0}^z (-1)^{w+1} \binom{z}{w} \quad (2.5)$$

$$= 1 + \sum_{w=0}^z (-1)^{w+1} \binom{z}{w} \quad (2.6)$$

$$= 1 + 0 \quad (2.7)$$

$$= 1 \quad (2.8)$$

So, we see that every M that has at least one all-zero, origin-containing side will contribute a value of exactly 1 to the original sum, and every M that lacks such a side will contribute nothing. Therefore the sum will give exactly the count of how many tensors have the desired property, and therefore $N(s_1, \dots, s_n)$ does give the size of the set $\mathcal{Z}(s_1, \dots, s_n)$. \square

Corollary 2.2. *The cardinality of the set $\bar{\mathcal{Z}}(s_1, \dots, s_n)$ is $k^{s_1 s_2 \dots s_n} - N(s_1, \dots, s_n)$.*

Proof: There are $k^{s_1 s_2 \dots s_n}$ tensors over \mathcal{C} of the given size, of which $N(s_1, \dots, s_n)$ are zero-sided tensors (by Theorem 2.1), so there must be $k^{s_1 s_2 \dots s_n} - N(s_1, \dots, s_n)$ non-zero-sided tensors. \square

Lemma 2.3. *The set $\mathcal{T}_c^\perp(s_1, \dots, s_n)$ is spanned by the set $\{e(s, M) - e(s, P_C(M)) \mid M \in \mathcal{B}(2s_1 - 1, \dots, 2s_n - 1)\}$*

Proof: A trivial is defined to be an energy function such that given any size for a universe, all universe states of that size will be assigned the same energy. Equivalently, if all the possible states are listed in some order, each state must have the same energy as the next on the list. If the states are listed in Gray code order, then each will differ from the next in only a single cell. Since total

energy is unchanged by rotation of the entire universe, we can rotate the state in each equation so that the cell that changes is in the center of the universe. If there are more than two colors, then the energy with the central cell set to each nonzero color can be equated to the energy with that central cell set to zero. Thus, for any given size s , the trivials can be defined as those functions that satisfy

$$E(s, M) = E(s, P_C(M)) \quad (2.9)$$

or

$$E(s, M) - E(s, P_C(M)) = 0 \quad (2.10)$$

for each possible choice of universe state M . Recall that $E(s, M)$ is the total energy of the universe, found by summing the energy function over all possible windows of size s , and P_C sets the central cell in the universe to zero. Each E is defined as a sum of the energies for all possible positions of the window. But for any window position that does not include the central cell, the corresponding terms in the two E functions will cancel. Therefore, rather than considering all universes of all sizes, it is only necessary to consider small tensors M that are just large enough to contain all the windows that contain the cell that is zeroed, without any toroidal wrapping at the edges. The equations will be satisfied for all such tensors if and only if they are satisfied for all possible universe states of any size. Given that the energy window is of size $s = (s_1, \dots, s_n)$, we need only consider the matrices over \mathcal{C} of size $(2s_1 - 1) \times (2s_2 - 1) \times \dots \times (2s_n - 1)$.

There are only finitely many possible patterns of cells that fit within the energy function for a given window size. So the value that the function assigns to each input pattern can be considered a variable, and the set of equations can be viewed as a set of linear equations in those variables.

For example, consider a 1D universe with 2 colors and a window size of 2. In general, an energy function f for this case can be written as

$$f(M) = \begin{cases} v_1, & \text{if } M = \mathbf{00} \\ v_2, & \text{if } M = \mathbf{01} \\ v_3, & \text{if } M = \mathbf{10} \\ v_4, & \text{if } M = \mathbf{11} \end{cases} \quad (2.11)$$

Equivalently, it can be written in terms of the matching function H

$$f(M) = v_1 H(M, \mathbf{00}) + v_2 H(M, \mathbf{01}) + v_3 H(M, \mathbf{10}) + v_4 H(M, \mathbf{11}) \quad (2.12)$$

It is also useful to convert f from a function to a corresponding coefficient vector using the V function defined earlier

$$V(f) = (v_1, v_2, v_3, v_4) \quad (2.13)$$

The energy functions correspond to the linear vector space of all such coefficient vectors. The conserved functions correspond to a subspace of that, and the trivials correspond to a subspace of that. The subspace corresponding to the trivials can be found by solving the set of linear equations described above. For this example, those equations are:

$$E(s, \dots \mathbf{010}\dots) - E(s, \dots \mathbf{000}\dots) = 0 \quad (2.14)$$

$$E(s, \dots \mathbf{011}\dots) - E(s, \dots \mathbf{001}\dots) = 0 \quad (2.15)$$

$$E(s, \dots \mathbf{110}\dots) - E(s, \dots \mathbf{100}\dots) = 0 \quad (2.16)$$

$$E(s, \dots \mathbf{111}\dots) - E(s, \dots \mathbf{101}\dots) = 0 \quad (2.17)$$

In each equation, the ellipses represent cells that have no effect on the result, since the state passed to the first E differs from the second only in the central cell. Since the energy window is of size 2, the region that has an effect is of size $2 \cdot 2 - 1 = 3$. Given the definition of f , this set of four simultaneous equations can be rewritten as

$$(v_2 + v_3) - (v_1 + v_1) = 0 \quad (2.18)$$

$$(v_2 + v_4) - (v_1 + v_2) = 0 \quad (2.19)$$

$$(v_4 + v_3) - (v_3 + v_1) = 0 \quad (2.20)$$

$$(v_4 + v_4) - (v_3 + v_2) = 0 \quad (2.21)$$

Each of the original E terms becomes only two terms here, because a window of size 2 can only be in two different positions in the universe and still overlap the central cell. The full definition of E would add more terms within each set of parentheses, but all of those additional terms will cancel in the subtraction because they represent windows that are not overlapping with the only cell that is different. Within each set of parentheses, the first v represents an energy window looking at the first and second of the 3 cells, and the second v represents an energy window looking at the second and third cell. These four simultaneous equations can be combined into a single equation in the obvious way:

$$U = \text{*****} \mathbf{1} 000 \text{*****} \quad (2.25)$$

Assume that f is an energy function that satisfies all the equations generated by tensors in the form of the M shown, where each $*$ is an arbitrary color, the 1 is an arbitrary nonzero color, and each 0 is the 0 color. If f satisfies all these equations, then by the definitions of the equations, whenever M can be superimposed on a larger universe at a position such that each 0 in M corresponds to a cell in that universe that is also 0 , then the cell in that universe corresponding to the 1 in M can be changed to an arbitrary color without changing the total energy of the universe. In this illustration, this means that the cell in U marked as 1 could be changed to 0 without affecting the total energy of the universe. If that change is made, then in the new universe, M can slide one more step to the left and still match. Therefore the cell in U to the left of the 1 can also be set to 0 . Continuing, every cell in the universe can eventually be set to 0 without changing the total energy. So by induction, we can conclude that any universe containing an island of s_1 adjacent 0 cells is guaranteed to have the same total energy as the all-zeros universe.

Now consider a universe $U_{1,1}$ with an arbitrary state, and 3 other universes that are identical except that one or two islands of s_1 contiguous cells have been set to zero, where the islands are far enough apart that no single energy window can touch two islands simultaneously.

$$U_{0,0} = \text{*****} 000 \text{*****} 000 \text{*****} \quad (2.26)$$

$$U_{0,1} = \text{*****} 000 \text{*****} \text{*****} \quad (2.27)$$

$$U_{1,0} = \text{*****} \text{*****} 000 \text{*****} \quad (2.28)$$

$$U_{1,1} = \text{*****} \text{*****} \quad (2.29)$$

The first 3 universes each contain at least one island, and so must all have the same total energy as the all-zeros universe. By the definition of total energy, it must be true that $(E(U_{0,0}) - E(U_{0,1})) - (E(U_{1,0}) - E(U_{1,1})) = 0$, because all the terms cancel out. Therefore $U_{1,1}$ must have the same energy as the all-zeros universe. Since $U_{1,1}$ is completely arbitrary, all universes must have the same energy, so f is trivial. Therefore this subset of equations is still large enough to guarantee that its solution space is still the trivials.

This 1D island argument generalizes to higher dimensions by applying it once in each dimension. For example, in 2 dimensions, M is a matrix of size $2s_1 - 1$ by $2s_2 - 1$ with all zeros after the center element (in row major order),

and a universe with a 2D island looks like U .

$$M = \begin{array}{c} \begin{array}{cccccc} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \\ \end{array} \quad (2.30)$$

$$U = \begin{array}{c} \begin{array}{cccccccccccccccc} * & * \\ * & * \\ * & * \\ * & * & * & * & * & * & * & * & 1 & 0 & 0 & 0 & * & * & * & * & * & * & * & * & * & * & * \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ * & * \\ * & * \\ * & * \end{array} \\ \end{array} \quad (2.31)$$

Now U has an island of zeros consisting of a ribbon of s_1 all-zero rows, topped by a bump of s_2 zeros. If M is superimposed on U such that the cells marked 1 correspond, then every 0 in M corresponds to a 0 in U . Therefore, for any f satisfying the equations, that 1 can be set to 0 in U without changing the total energy of U . Then the cell to the left can be zeroed, and so on, until the entire row is set to zero. Applying the 1D island argument horizontally to that row, we can conclude that for any universe that contains a ribbon of s_1 all-zero rows, the row above can be set to all zeros without affecting the total energy. Repeating that to the row above, and so on, the entire universe can be zeroed out. So applying the 1D island argument vertically to rows, we can conclude that every universe has the same energy as the all-zero universe. Thus f is trivial. So the restricted set of equations is again sufficient.

The same argument in 3D would have a 3D M with zeros in its second half (in row major order). The starting U would have a 3D island of zeros. Such an island consists of s_1 layers of all zeros, topped by a 2D island of zeros. The 2D island of zeros is a matrix with s_2 rows of zeros, topped by a 1D island of zeros. The 1D island of zeros is a row of s_3 zeros. The proof in 3D is the same as before, starting in the last dimension and working back. The argument generalizes to n dimensions in the same way. Therefore there is no loss of generality if M is restricted to be a large tensor with zeros for the second half of its elements. The restricted equations will span the same space as the original ones in the previous lemma, and so this lemma is proved. \square

Lemma 2.5. *The set $\mathcal{T}_c^\perp(s_1, \dots, s_n)$ is spanned by the set $\{e(s, M) - e(s, P_C(M)) \mid M = P_Z(A), A \in \bar{\mathcal{Z}}(s_1, \dots, s_n)\}$*

Proof: This lemma is identical to the previous one, but restricting M to an even smaller subset of the possible tensors. In the preceding lemma, M could

be any large tensor that is half zeros. In this case, M can only be those tensors that come from taking a smaller, non-zero-sided tensor, and using P_Z to pad it with zeros. Such padding always leaves the second half entirely zero, so the set of M tensors in this lemma is a subset of that in the last lemma.

We will show that the same space is spanned, even if M is restricted to only those tensors where all the nonzero elements fit within a single energy window. To do this, we will show that for any M that contains two nonzero elements that are too far apart, the corresponding equation is not linearly independent of the other equations, and so can be removed.

Consider the case where M is such a matrix. Of the two nonzero elements that are far apart, arbitrarily choose one and call it the "first element", and call the other the "second element". Let $M_{0,1}$ be M with the first element set to 0, let $M_{1,0}$ be M with the second element set to zero, let $M_{0,0}$ be M with both set to zero, and let $M_{1,1}$ be M with neither element changed (i.e., $M_{1,1}$ is another name for M). The set of linear equations that defines the trivial conserved functions will therefore include these 4 equations:

$$E(s, M_{0,0}) - E(s, P_C(M_{0,0})) = 0 \quad (2.32)$$

$$E(s, M_{0,1}) - E(s, P_C(M_{0,1})) = 0 \quad (2.33)$$

$$E(s, M_{1,0}) - E(s, P_C(M_{1,0})) = 0 \quad (2.34)$$

$$E(s, M_{1,1}) - E(s, P_C(M_{1,1})) = 0 \quad (2.35)$$

An energy function satisfying these equations should also satisfy any linear combination of them, so it must satisfy the sum of the middle two minus the other two, which gives this equation:

$$\begin{aligned} & E(s, M_{0,1}) - E(s, P_C(M_{0,1})) + E(s, M_{1,0}) - E(s, P_C(M_{1,0})) \\ & - E(s, M_{0,0}) + E(s, P_C(M_{0,0})) - E(s, M_{1,1}) + E(s, P_C(M_{1,1})) = 0 \end{aligned} \quad (2.36)$$

We assumed there is a large enough distance between the two nonzero elements being considered, that no single energy window can contain both of them at once. Those terms in the sum $E(s, M_{0,0})$ that include the first element will therefore cancel out with those terms in the negative sum $-E(s, M_{0,1})$ that contain the first element. This is because those two sums differ only in the terms that include the second element, none of which include the first element. Similarly, the terms in $E(s, M_{0,0})$ that contain the second element

cancel with those in $-E(s, M_{1,0})$ that contain the second element. Note that there are 8 summations in the above equation, and they can all be paired up similarly so that all terms including the first element cancel, as do all containing the second element. In addition, if there is a window that contains neither the first nor second element, then the term for that window will occur in all 8 summations, being positive in 4 and negative in 4, and so will cancel out. Therefore, all terms in all the summations cancel, and the equation reduces to the equation $0 = 0$.

Given a set of linear equations, if some subset of the equations sums to the $0 = 0$ equation, that means that they are linearly dependent, and so any one of them can be removed from the set without loss of generality. We will choose to remove the equation involving $M_{1,1}$. This can be repeated for each M matrix that contains any two nonzero elements that are too far apart to fit in a single energy window. If the M tensors with highest number of nonzero elements are deleted first, then this procedure will always be possible, because when it is time to delete the tensor $M_{1,1}$, the 3 tensors $M_{0,1}$, $M_{1,0}$, and $M_{0,0}$ will all have fewer nonzero elements, and so cannot have been deleted yet. Therefore, without loss of generality, we can remove all of the equations based on M tensors with nonzero elements that are too far apart to cover with a single energy window.

At this point, the set of tensors being considered consists of those M tensors that have all 0 elements in their second half, and have all of their nonzero elements clustered together in a region that fits within a single energy window. For example, in 2D, if the energy window is 3×5 then the entire M matrix will be 5×9 . One example of an M that would generate one of the equations that remains in the final set is:

$$M = \begin{matrix} 000000000 \\ 000000000 \\ 000000000 \\ 000000000 \\ 000000000 \end{matrix} \quad (2.37)$$

This is a legal M because there is a nonzero element in the center, all the elements are 0 in the second half, and all of the nonzero elements fit within a single energy window. That last fact is shown by shading a 3×5 region that is the size of the energy window, and includes all the nonzero elements. There are actually several positions that the shaded window could have been drawn. For uniqueness, we will always choose to draw that window so that there is a nonzero element in the top row and in the leftmost column of the window. In other words, the window will be chosen so that its contents are not a zero-sided tensor. If the contents of that window are called A , then

$M = P_Z(A)$, and $A \in \bar{\mathcal{Z}}(s_1, \dots, s_n)$. Any A in that set will generate a legal M corresponding to an equation that is kept. Therefore the set given in the lemma is sufficient to span the desired space, and the lemma is proved. \square

Theorem 2.6. *One basis set for the coefficient vectors of the equations defining the trivial conserved functions with energy windows of size $s = (s_1, \dots, s_n)$ is $\{e(s, M) - e(s, P_C(M)) \mid M = P_Z(A), A \in \bar{\mathcal{Z}}(s_1, \dots, s_n)\}$.*

Proof: This theorem is equivalent to the last lemma, except that the set must not only span the desired space, but it must also be minimal, with no linearly dependent vectors.

If a matrix is upper triangular with no zeros on the diagonal, then its rows will be linearly independent. One way to prove that the rows of an arbitrary matrix are linearly independent is to reorder the columns and then reorder the rows to put it in that form. More generally, to show that all the vectors in a set are linearly independent, it is sufficient to give a name to each position in the vector, and give a name to each vector, where there is a total order on names, and the vector named A has a nonzero element in the position named A , and it has zero elements in all the positions with names less than A . This will now be done with the coefficient vectors for the equations defining the trivials.

Each equation is generated from a tensor M . It was shown in the last lemma that it is sufficient for each M to be generated from a non-zero-sided tensor A by the equation $M = P_Z(A)$. Therefore it is natural to use A as the name of the coefficient vector for its corresponding equation. The terms in the equation refer to energy functions defined over energy windows the same size and shape as A . Therefore it is natural to use A as the name of the position in the vector corresponding to the energy function applied to a window whose contents are A . It remains only to find some total order over the tensors that has the desired property.

Consider the following ordering over the tensors $A \in \bar{\mathcal{Z}}(s_1, \dots, s_n)$. In this ordering, the tensors are sorted by the number of nonzero elements. Those tensors with more nonzero elements come before those with fewer. In the case of a tie, the tensors are ordered lexicographically, with color 0 coming after all other colors, with all the elements unrolled into a single list in row major order. So in 2D, if two matrices have the same number of nonzero elements, then their order is determined by their upper-left element. If that element is a tie, then their order is determined by the element to the right of it. If that is a tie, comparison continues across the top row, then left-to-right

across the second row, and so on down to the bottom row. It will now be shown that this ordering has the desired property.

Consider the M shown in equation (2.37), and the A which is shown shaded. This M will generate an equation with a term for each possible position of the energy window, both with and without setting the center cell to 0. In the ordering just described, the window position that is shown as shaded will come before any other possible position, and before any version with the central element set to 0. It must come before any position where the window is moved some distance to the right, or down, or both, because that will reduce the number of nonzero elements in the window, because we assumed A is a non-zero-sided matrix, and therefore has nonzero elements on the top row and leftmost column, which will leave the window if it moves right or down. The shaded window shown will also come before positions where it is moved up or left or both, because these moves will increase the number of leading zeros in the row-major-order unrolling of the contents of the window, and so those shifted windows will come later than the shaded window. And the shaded window must come before the shaded window in the same position when M is modified to set its central cell to 0, because such a modification will decrease the number of nonzero element in the shaded window.

Therefore, for any equation generated by an $M = P_Z(A)$, there will be a term referring to A , and all other terms will refer to tensors that come later in the ordering. So the coefficient vector named A has a nonzero number in position A and zeros in all positions less than A . Thus this ordering has the desired properties, and so the equations are linearly independent.

So it has been shown that the proposed set of equations are sufficient to define the trivials, and are a minimal set, because they are linearly independent. Thus they form a basis set, and the theorem is proved. \square

Lemma 2.5 gave a small set of equations defining the trivial conserved functions. Each trivial is conserved for *all* possible cellular automata. It is natural to seek the equivalent for the conserved functions for any *particular* CA. Such a set is given by the following theorem.

Theorem 2.7. *The set of coefficient vectors for one set of linear equations that defines the conserved functions with energy windows of size $s = (s_1, \dots, s_n)$ for cellular automaton rule R with neighborhood of size $w = (w_1, \dots, w_n)$ is*

$$\{e(s, M) - e(s, P_C(M)) - e(s, R_{reg}(R, M)) + e(s, R_{reg}(R, P_C(M))) \\ | M = P_Z(A), A \in \bar{Z}(s_1 + w_1 - 1, \dots, s_n + w_n - 1)\}$$

Proof: In this set, M ranges over only those big tensors that can be gen-

erated by padding certain smaller tensors with zeros. If M were to instead range over all big tensors, then this set of equations would become the same as the unoptimized, common definition of the equations defining the conserved functions, which has been given in practically every paper that has been published on conserved functions for CAs. It states that the “conserved functions” can be defined by imagining an arbitrarily-large universe with an arbitrary state, and one cell defined to be the “origin”, which has a nonzero color. If the CA is run for one step, the total energy of the universe must remain unchanged. If the origin is instead set to zero and the universe is run for one step, then again the energy must remain unchanged. These four energies (the universe before and after that step, with the origin set to 0 or not), must therefore satisfy the equation $(E(M) - E(\text{successor}(R, M))) - (E(P_C(M)) - E(\text{successor}(R, P_C(M)))) = 0$, for all universe states M . Because of canceling terms, it is sufficient to draw M from the set of tensors of a size related to the sum of the CA neighborhood and energy window sizes, rather than drawing from the infinite set of all universes of all sizes. That yields a statement identical to the theorem, except for the mentioned restriction.

This theorem restricts M to a subset of the tensors analogous to that used in Lemma 2.5 for the trivials. In fact, the proof here goes through in an identical way to the proofs in Lemmas 2.4 and 2.5. Where proof by induction works for the trivial lemmas, it also works for this conserved theorem. Where terms in sums cancel for the trivial lemmas, they also cancel for the conserved theorem. Where equations are shown to be linearly dependent and are deleted in the trivial lemmas, the same can be done for the conserved theorem. So the proof for this theorem will not be shown in detail here, since it is essentially the same. \square

Now that a basis set for \mathcal{T}^\perp has been shown, it is possible to give a basis set for the trivials, \mathcal{T} . The following definition defines such a set, and is followed by a theorem proving that it actually is such a basis.

Definition 2.8. For a given tensor size $s_1 \times \dots \times s_n$, the set \mathcal{T}_b is defined to be the following set of functions, and \mathcal{T}_{bc} is the set of coefficient vectors corresponding to those functions.

$$\mathcal{T}_b(s_1, \dots, s_n) = \{f_M \mid M \in \mathcal{Z}(s_1 \times \dots \times s_n)\} \quad (2.38)$$

$$\mathcal{T}_{bc}(s_1, \dots, s_n) = \{V(f) \mid f \in \mathcal{T}_b(s_1 \times \dots \times s_n)\} \quad (2.39)$$

where for any given M , the function $f_M : C^{s_1 \times \dots \times s_n} \rightarrow \mathbb{R}$ is defined as

$$f_M(x) = \begin{cases} 1 & \text{if } M = \mathbf{0} \\ H(x, P_L(M)) - H(x, P_R(M)) & \text{otherwise} \end{cases} \quad (2.40)$$

Theorem 2.9. *The set $\mathcal{T}_b(s_1, \dots, s_n)$ is a basis set for the space of all trivial conserved functions with energy windows of size $s_1 \times \dots \times s_n$.*

Proof:

To prove \mathcal{T}_b is a basis for the trivials, it is sufficient to show three things: that every element of \mathcal{T}_b is a trivial, that its elements are linearly independent, and that the number of elements of \mathcal{T}_b equals the dimensionality of the space of trivials. Each of these will now be shown.

It is clear that each element of \mathcal{T}_b is a trivial. One of them is the constant function $f(x) = 1$, which is trivial because it assigns to each universe an energy equal to the number of cells, independent of the state of those cells. The rest of the functions are of the form $f(x) = H(x, A) - H(x, B)$ where tensors A and B are two different patterns formed by taking a smaller tensor C and padding it on one or more sides with * symbols. If a universe state contains n different (possibly overlapping) copies of the pattern C , then as f is scanned over the entire universe, it will match A exactly n times and match B exactly n times, yielding a total energy for the universe of $n - n = 0$. Thus all universe states are given a total energy of zero regardless of the state, so the function is trivial.

We will show that the functions in \mathcal{T}_b are linearly independent by showing that the vectors in \mathcal{T}_{bc} are linearly independent. To check for linear independence, the H function should first be expanded by summing over all possible ways to replace each * symbol with a color from \mathcal{C} . For example, in 2D with the colors $\{0, 1\}$, the trivial:

$$f(x) = H(x, \begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix}) - H(x, \begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix}) \quad (2.41)$$

can be expanded as:

$$\begin{aligned} f(x) = & H(x, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}) + H(x, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}) + H(x, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}) + H(x, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}) \\ & - H(x, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}) - H(x, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}) - H(x, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}) - H(x, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}) \end{aligned} \quad (2.42)$$

These are equivalent by the definition of the H function. If \mathcal{C} contains k colors, and x is a tensor with n elements, then there are k^n possible patterns

that can appear within the H function, and so each trivial can be represented by a vector of k^n coefficients. As in the earlier proof, we can prove these trivials are linearly independent by assigning each possible vector a name, and assigning each position within the vector a name, and choosing an ordering over names, such that vector A has a nonzero in position A and has zeros in all earlier elements.

It is natural to name each position in the vector by the tensor that is inside H for that position. Since \mathcal{T}_b is defined to have a trivial for each $M \in \mathcal{Z}(s_1, \dots, s_n)$, it is natural to use that M as the name for each vector. Let the ordering be an ordinary lexicographical ordering, where the elements of a tensor are taken in row major order, and where 0 comes before all other colors (which is the opposite convention as in the earlier proof).

If these choices of naming and ordering are applied to the example in equations (2.41) and (2.42), it is clear that the name of the trivial will equal the tensor in the first term shown in equation (2.42), and the first coefficient (according to the chosen ordering) is the same tensor. Therefore, it has the desired properties. This will be true in general. The two tensors in (2.41) from applying P_L and P_R to a zero-sided tensor M . Therefore, M can be recovered by replacing the * elements in P_L with zeros. That same tensor will always appear as the first term (in the chosen ordering) in the expanded function. Clearly, the top row of (2.42) will always have its lexicographically-first element being the one where all the * elements were replaced with zeros. That is because the replacement of a * with a 0 will always come before the replacement of a * with a nonzero. A similar argument shows that the lexicographically-first tensor on the second row will also be first. And when those two are compared, the one on the first row will always precede the one on the second row, because the former is simply the latter shifted one space to the right by inserting zeros on the left.

Thus the vectors have the property that the first nonzero element in each vector is in the position whose name matches the name of the vector. Therefore the vectors are linearly independent.

It remains only to show that the space spanned by the trivials in \mathcal{T}_b has the same dimensionality as the space of trivials. The dimensionality of the space of energy functions equals the number of tensors the size of the energy window, because each energy function is defined by a vector of coefficients, with one coefficient per possible tensor. It was proved that a basis set for the linear equations defining the trivials has the same number of equations as there are non-zero-sided tensors. Therefore, the solution space for those equations (i.e. the space of all trivials) must have a dimensionality that is

equal to the number of possible tensors minus the number of non-zero-sided tensors. This difference is simply the number of zero-sided tensors (because the zero-sided and non-zero-sided tensors partition the set of all tensors).

Thus, \mathcal{T}_b contains a number of trivials equal to the dimensionality of the entire space of trivials. Since they are linearly independent, and are all trivials, this proves that \mathcal{T}_b is a basis set for the space of all trivials. \square

In [2], we defined *comp* as a function from matrices to matrices, such that given a matrix T with c columns, the matrix $comp(T)$, called the *complement* of matrix T , is a matrix with c columns, with the property that the rows of $comp(T)$ are linearly independent of each other and of the rows of T , and that together with the rows of T they span all of \mathbb{R}^c .

A *unit complement matrix* is a complement matrix that has a single 1 on each row, with all other elements zero. It was proved in that paper that if T is a matrix formed from the equations defining the trivials, and if $comp(T)$ is a unit complement, then it is particularly useful. If C is the matrix of coefficients of the equations defining the conserved functions, then we can zero out (i.e., set all elements to zero) every column of C that corresponds to a column with a 1 element in the unit complement $comp(T)$. The resulting set of equations will have a solution that spans only a subset of the space of the nontrivials. But the union of that solution with the trivials will span exactly the space of all conserved functions. See [2] for the proofs of all those results.

This modified C matrix is extremely useful for empirical searches. The zeroed columns can be deleted before having the computer solve the system of equations, and the zeros can be inserted back in to the resulting answer. So the matrix to solve will be smaller. In fact it is only half the size, for a binary 1D CA. More importantly, if we want to know whether a given CA has any nontrivials, we can simply check whether the modified C is full rank. If it is, then no nontrivials exist. This allows for faster searches for CAs with nontrivials of a given size,

To use that result, we need to know the unit complement matrix in general. The earlier papers gave them only for 1D universes with 2 colors. The following theorem gives a unit complement matrix for the trivials for any universe with any arbitrary, finite number of dimensions, with any arbitrary number of colors, and any arbitrary energy window size. This was used in our exhaustive search of Life and of all the life-like CAs discussed in the next section. This search would have been impractically slow if we had not known the unit complement matrices in 2D, and used them to optimize our search.

Theorem 2.10. *The set of rows of a unit complement of the matrix of coeffi-*

cient vectors for the equations defining the trivials for energy windows of size $s_1 \times, \dots, \times s_n$ is $\{V_t(M) \mid M \in \mathcal{Z}\}$.

Proof: For the proposed matrix to be a complement matrix, its rows must be vectors that are linearly independent of each other, and are not in the space spanned by the trivials, and that together with the trivials span the entire space. To be a unit complement, they must have a 1 on each row, in a different position, with the rest zeros.

It is obvious that if the proposed set is a complement matrix then it will be a unit complement matrix, since by the definition of V_t , each vector consists of all zeros and a single 1, and the 1 is in a different location for each vector. This also guarantees that the rows of the matrix are linearly independent.

It is also obvious that each of these vectors is outside the space spanned by the trivials. That is because each vector in the proposed set corresponds to an energy function of the form $f(x) = H(x, M)$ for some M in \mathcal{Z} . Such a function will assign an energy to a universe equal to the number of copies of M in that universe's state, and so is not trivial, so its vector is not within the space spanned by the trivials.

Finally, it must be shown that the dimensionality of this space plus the dimensionality of the space of trivials equals the dimensionality of the space of all energy functions. The dimensionality of these three spaces clearly equals the sizes of the sets \mathcal{Z} , $\bar{\mathcal{Z}}$, and \mathcal{B} , respectively. Since the size of the last is the sum of the sizes of the first two, the desired relationship holds. Therefore, the theorem is proved. \square

Figure 1 summarizes all the theorems of this paper, giving four examples of the M matrix for each concept.

3 COMPUTATIONAL RESULTS FOR LIFE-LIKE CA'S

The challenge in identifying cellular automata with nontrivial conserved energy functions is the enumeration of the trivials and their elimination from the solution space. The actual calculation of the nontrivials can then be reduced to the calculation of the null space of the system of corresponding state space equations. Thus the theorems and definitions of the previous section may be used as the basis for computational identification of cellular automata with nontrivials of various orders. Computationally, this proceeds as follows:

- 1) Choose a CA and energy window size (s_1, s_2) .
- 2) For all possible matrices M given by Theorem 2.7, generate the corresponding state space equations.

Energy window matrix Size: $r \times c$ Count: k^{rc}	$\begin{matrix} 10010 \\ 00000 \\ 00100 \end{matrix}$	$\begin{matrix} 01010 \\ 10101 \\ 01010 \end{matrix}$	$\begin{matrix} 11111 \\ 11111 \\ 11111 \end{matrix}$	$\begin{matrix} 00000 \\ 00000 \\ 00000 \end{matrix}$
Zero-sided matrix Size: $r \times c$ Count: $N(r, c) = k^{(r-1)c} + k^{r(c-1)} - k^{(r-1)(c-1)}$	$\begin{matrix} 01111 \\ 00101 \\ 00100 \end{matrix}$	$\begin{matrix} 00000 \\ 10111 \\ 00001 \end{matrix}$	$\begin{matrix} 00000 \\ 01000 \\ 01010 \end{matrix}$	$\begin{matrix} 00000 \\ 00000 \\ 00000 \end{matrix}$
Unit complement function Size: $r \times c$ $f(x) = H(x, M)$	$\begin{matrix} 01111 \\ 00101 \\ 00100 \end{matrix}$	$\begin{matrix} 00000 \\ 10111 \\ 00001 \end{matrix}$	$\begin{matrix} 00000 \\ 01000 \\ 01010 \end{matrix}$	$\begin{matrix} 00000 \\ 00000 \\ 00000 \end{matrix}$
Trivial conserved function Size: $r \times c$ $f(x) = H(x, M) - H(x, M')$	$M = \begin{matrix} *1111 \\ *0101 \\ *0100 \end{matrix}$	$M = \begin{matrix} *1111 \\ *1011 \\ *0001 \end{matrix}$	$M = \begin{matrix} *1111 \\ *1000 \\ *1010 \end{matrix}$	$f(x) = 1$
	$M' = \begin{matrix} 1111* \\ 0101* \\ 0100* \end{matrix}$	$M' = \begin{matrix} 1011* \\ 0001* \\ *1111 \end{matrix}$	$M' = \begin{matrix} 1000* \\ 1010* \\ *1111 \end{matrix}$	
Non-zero-sided matrix Size: $r \times c$ Count: $k^{rc} - N(r, c)$	$\begin{matrix} 01000 \\ 10000 \\ 10001 \end{matrix}$	$\begin{matrix} 10000 \\ 11000 \\ 00000 \end{matrix}$	$\begin{matrix} 10111 \\ 00000 \\ 00000 \end{matrix}$	$\begin{matrix} 10000 \\ 00000 \\ 00000 \end{matrix}$
Equations defining the trivial conserved functions Size: $(2r - 1) \times (2c - 1)$ $0 = e(M) - e(M')$	$\begin{matrix} 01000000 & 00000000 & 00000000 & 00000000 \\ 10000000 & 00010000 & 00000000 & 00000000 \\ 10001000 & 00011000 & 10111000 & 00001000 \\ 00000000 & 00000000 & 00000000 & 00000000 \\ 00000000 & 00000000 & 00000000 & 00000000 \end{matrix}$			
Non-zero-sided matrix Size: $(r + 2) \times (c + 2)$ Count: $k^{(r+2)(c+2)} - N(r + 2, c + 2)$	$\begin{matrix} 0000101 \\ 0000100 \\ 1000000 \\ 0000100 \end{matrix}$	$\begin{matrix} 1000000 \\ 0000111 \\ 0000001 \\ 0000000 \end{matrix}$	$\begin{matrix} 1000000 \\ 0000000 \\ 0000001 \\ 1000000 \end{matrix}$	$\begin{matrix} 1000000 \\ 0000000 \\ 0000000 \\ 0000000 \end{matrix}$
Equations defining the conserved functions Size: $(2r + 3) \times (2c + 3)$ $0 = e(M) - e(M')$ $-e(s(M)) + e(s(M'))$	$\begin{matrix} 00001010000 & 00000000000 & 00000100000 & 00000000000 \\ 00000100000 & 00001000000 & 00000000000 & 00000000000 \\ 00000000000 & 00000000010 & 00000000000 & 00000000000 \\ 00000000000 & 00000000001 & 00000000000 & 00000000000 \\ 00000000000 & 00000000000 & 00000000000 & 00000000000 \\ 00000000000 & 00000000000 & 00000000000 & 00000000000 \\ 00000000000 & 00000000000 & 00000000000 & 00000000000 \\ 00000000000 & 00000000000 & 00000000000 & 00000000000 \end{matrix}$			

FIGURE 1

Summary of the main theoretical results of this paper, with four examples of each concept. The proofs are for arbitrary dimensions, neighborhood sizes, and number of colors, but the figure shows only 2D examples, for a CA with a 3×3 neighborhood, and $k = 2$ colors. In each case, M' is M with the central bit set to 0. For the equations, the large matrix is formed by padding the small matrix with zeros such that the last 1 bit ends up in the center of the large matrix (where "last" is the last 1 found when traversing the elements in row major order). In each of the four sections, the listed concepts all have the same count. For example, the number of zero-sided matrices of a given size equals the number of unit complement functions, which equals the number of trivials.

3) To remove the trivials from the solution space, delete the columns associated with the zero-sided tensors as determined by Theorem 2.10. This has the additional benefit of significantly reducing the size of the energy vectors and, therefore, the state space matrix as a whole.

4) Determine the rank of the resulting matrix. If it is full rank, the system of equations has no solution, and therefore no nontrivial exists for the given CA and window size. If the matrix is rank-deficient, a nontrivial exists. It is completely characterized by the basis vectors that are the columns of the matrix's null space.

In [2], we gave a complete taxonomy of nontrivials for 1D, 2-state cellular automata for energy windows up to size 16. Using the definitions and theorems previously presented, we now extended these results to 2D, 2-state automata, for energy windows up to size 9.

There are a total of $k^{(k^9)}$ cellular automata in 2D with k colors. This number is so large that any investigation other than a random sampling is effectively impossible. Accordingly, drawing substantive conclusions about unrestricted 2D cellular automata seems to the authors to be extraordinarily difficult. To reduce the scope of the problem and make a more complete investigation possible, we consider only *life-like* CA's (i.e. 2-D, 2-color, 3×3 neighborhood, outer totalistic).

Restricting the search space to life-like CA's significantly reduces the size of the problem. For such a CA, a cell can have from 0 to 8 outer neighbors of color 1, and as a function of that number the central cell can have one of four outcomes on the next step: (S)ame, (B)irth, (D)eath, and (F)lip. S leaves its color unchanged on the next step, B causes it to be set to alive (1), D causes it to be set to 0, and F causes it to flip from 0 to 1 or from 1 to 0. Thus any life-like CA can be represented as a string of characters from $\{S,B,D,F\}$. Using this notation, if we count the neighbors from 0 to 8 from left to right, Conway's Game of Life would be written as "DDSBDDDDDD". We refer to this description as the CA's *rule vector*.

There are 4^9 life-like CA's. It is known that renumbering the colors of a CA in reverse order and changing the outcomes correspondingly produces a CA identical to the original, up to isomorphism. Using the proposed notation, this corresponds to reversing the order of the letters and swapping B with D. The rule vector of every CA can be manipulated in this way to produce an isomorph. Of these CA's, 2^9 are their own isomorphs, so the total number of isomorphism classes of life-like CAs is $(4^9 - 2^9)/2 + 2^9 = 2^{17} + 2^8 = 131,328$. This is considerably smaller than the non-life-like case.

The definitions and theorems in this paper give the dimensions of the ma-

Energy window height (s_1)	Energy window width (s_2)	$\lceil \log_2 \text{rows} \rceil$	$\lceil \log_2 \text{cols} \rceil$
1	2	16	1
1	3	19	2
1	4	23	3
2	2	20	4
1	5	26	4
1	6	29	5
2	3	25	6
1	7	32	6
1	8	35	7
2	4	29	8
1	9	39	8
3	3	30	9

TABLE 1
State matrix sizes for various energy windows

trices to be analyzed as a function of the energy window (independent of the CA being analyzed). We show the matrix sizes for some 2D examples in Table 1.

Column three shows the ceiling of the log base 2 of the maximum number of energy vectors needed to determine the existence of a nontrivial. Column four shows the ceiling of the log base 2 of the number of entries in each vector. This number is given by the total number of possible energy function values ($2^{s_1 s_2}$) minus the number of zero-sided tensors given by Definition 1.17.

Because these matrices have far more rows than columns, we might expect almost all of them to be full rank, and therefore few nontrivial conserved functions should exist over the range of cellular automata. Our empirical results were consistent with this. Full rank can be determined very quickly while rank deficiency cannot be known until all the possible state space vectors given by Theorem 2.7 have been examined for linear independence, so it would be inefficient to build the full state space matrix for each CA and then calculate its rank. Instead, we sift the sands of cellular automata through a three-stage computational sieve.

The first stage uses a "quick and dirty" algorithm to discard automata with no nontrivials. This eliminates over 99% of the candidates. The second stage takes automata that have passed the first stage and performs a little more work to try to drive the set of state space matrices to full rank. This eliminates about another 90% of the candidates it analyzes. The third stage operates only on

automata that have passed the first two stages, performing exact arithmetic using all the optimizations of Theorem 2.6 to determine whether or not a given CA has a nontrivial conserved function. If it does, its basis is calculated and reported. Each stage is implemented in MATLAB.

In stage I, we compute the energy vector of Definition 1.16 for one tensor at a time, attempting to add it to an existing energy vector set via Gaussian elimination to ensure that the rows in the state space matrix at any time are always linearly independent. Before such addition, however, we delete the columns corresponding to the zero-sided tensors for the indicated energy window. The total number of deleted columns is given by Definition 1.17. None of the optimizations discussed in the proof of Theorem 2.6 are performed at this stage. Instead, universe states are generated randomly, the energy vectors of their corresponding tensors are calculated, and Gaussian elimination is performed on each vector relative to those energy vectors already admitted into the state space matrix. When the number of linearly independent energy vectors is equal to the number of columns (the number of possible energy function values minus the number of zero-sided tensors), full rank has been achieved, and the (CA, energy window) pair under test is known not to correspond to a nontrivial conserved function.

Since states are generated randomly in this stage, as opposed to exhaustive enumeration of the appropriate tensors as given by Theorem 2.7, the number of states N to try before giving up on the possibility of reaching full rank is a user-definable parameter. Empirically, we have found that setting N to 32 times the maximum rank of the matrix gives a good tradeoff between quick computation on the one hand and admitting too many false positives on the other.

During this stage, all arithmetic is performed modulo a small prime, to eliminate the possibility of roundoff error or overflow. If full rank is reached, the matrix would be full rank in exact arithmetic as well, so the answer is correct. If full rank is not reached within the indicated time window, the matrix may or may not be rank-deficient, so the CA is marked as a candidate for stage II computation.

In stage II, candidate (CA, window) pairs that pass through the first stage are subject to repeated random state generation with a larger value of N for multiple attempts. No other optimizations are performed at this time. If no full rank matrix is produced (i.e. no linearly independent energy vector set of the cardinality given by Definition 1.16 is found), the pair is marked for analysis by stage III.

Stage III computation employs on-the-fly Gaussian elimination for one-

at-a-time energy vector generation, similar to the first two stages, but using integers in double precision arithmetic and enumerating the state space exactly as described in the proof of Theorem 2.6. To keep the computations from overflowing, vectors are reduced modulo the GCD of all their nonzero entries during this process, which means this stage is the most computationally intensive. If Gaussian elimination on the entire set of energy vectors does not produce a linearly independent set of Definition 1.16 cardinality, then the constructed state space matrix has a null space. That null space is calculated, and reported as the basis for all nontrivial conserved functions for that particular (CA, window) combination.

To guard against the possibility of numerical error, the largest value observed during stage III calculation is tracked and reported, to ensure that any possibility of overflow or loss of precision will be detected. For all calculations reported here, this maximum value has always been well below that which could induce error in double precision arithmetic. So we are confident our results are correct. Nonetheless, as an added safety check, we have implemented code which accepts as input a CA, an energy window, and a stage III basis set reported as characterizing a nontrivial. It tests each vector in the basis set over large numbers of randomly selected states by evaluating the energy function through brute force dot product calculation. In all cases, the resulting functions reported by stage III were conserved.

Table 2 shows the results of our computations for all life-like cellular automata up to isomorphism, for all energy functions up to order 9. It extends [2] to give a complete list of conserved functions for all automata of this type. Figures 2 and 3 are similar to Figure 5, extended to two dimensions. Figure 4 summarizes our current knowledge of the 1D conserved functions for the 256 *elementary* CA's (1D, 2-state, 3-neighborhood).

We have shown by exhaustive search that some CAs have no nontrivials up to order 16, though higher-order nontrivials might still exist for those CAs. However, for some CAs, it is possible to prove that no nontrivials can exist of any order. For example, that can be proved for the CAs $\{0, 8, 32, 40, 128, 136, 160, 168\}$ based on their long-term behavior for universe states starting with two adjacent 0 cells [1], and for $\{60, 90\}$ based on their long-term behavior for universes of a size that is a power of two [7].

In all cases, the basis functions shown in the figures were simplified using a Mathematica program. The program simplified the basis functions as much as it could, by combining terms using the * color, and by adding multiples of trivial functions, to reduce the number of terms in each basis function. This simplification process was heuristic, and so might be improved upon. But the

results shown are probably close to being as simple as possible, at least for those functions with few terms. It is an open question whether there is a much simpler way to represent some of the more complex functions.

Every life-like CA that has nontrivials of order 9 or less is shown in Table 2. Each row gives results for one such CA. The column *window* gives the size of the energy window, *basis* gives the number of basis functions that span the space of conserved functions, *CA#* gives the number for the CA using the standard numbering system used in Mathematica, *iso* gives the number for the isomorphic CA (if any), and *Rule* describes the rule for the CA using a common notation that lists the number of active neighbors needed for an inactive cell to become active (be Born), followed by the number of active neighbors needed for an active cell to remain active (to Survive). The *rule vector* describes what happens to a cell when it has 0 through 8 neighbors that are alive (set to 1). For each case, the letter indicates that the cell will be born (B), die (D), remain the same (S), or flip to the opposite (F). *Comments* describes the conserved function or lists another CA that has the same conserved function.

By symmetry, analogous conserved functions for any $m \times n$ window can also be found for one that is $n \times m$. Thus the only energy windows examined were those that were at least as wide as they were tall.

4 ANALYSIS

Figure 5 applies the ideas of this paper to the results of [2] and [4], expressing the basis functions as a linear sum of the matching H-functions of Definition 1.12.

Some patterns are clearly visible in Table 2, Figure 2, Figure 3 and Figure 4. For all life-like CA's for which nontrivial conserved functions exist, there is a great deal of homogeneity in the middle range of neighbor counts. For example, any given CA in the table has the same transition rules for neighbor counts 3-6, and most have identical transition rules for neighbor counts 2-7. We conjecture this is combinatorially driven. That is, for the middle range of neighbor counts, there are so many different ways to distribute a fixed number of neighbors among eight cells that a low-order conserved function cannot incorporate them all. By contrast, there is only one way to arrange zero or eight neighbors around a cell, eight ways to arrange one or seven, and so forth. Near the minimum and maximum of the neighbor count range, the number of possible configurations is sufficiently small that a low-order conserved function is more likely to emerge.

We also note that all life-like CA's with rule vectors of the form xFFFFFFFx,

window	basis	CA#	Iso	Rule	rule vector (num neighbors)								comments			
					0	1	2	3	4	5	6	7		8		
1 × 1	1	174762		S0123456789	S	S	S	S	S	S	S	S	S	S	identity, conserves all	
1 × 2	1	174760	240298	S12345678	D	S	S	S	S	S	S	S	S	S	conserves [111] pairs	
	1	87381		B012345678	F	F	F	F	F	F	F	F	F	F	conserves [10] pairs	
2 × 2	1	174730	178858	S01345678	S	S	D	S	S	S	S	S	S	S	conserves 2x2 pattern with all 1's	
	1	174728	244394	S1345678	D	S	D	S	S	S	S	S	S	S	identical to 174730	
	1	174722	195242	S0345678	S	D	D	S	S	S	S	S	S	S	identical to 174730	
	1	174720	260778	S345678	D	D	D	S	S	S	S	S	S	S	identical to 174730	
	4	191144	240290	B7/S12345678	D	S	S	S	S	S	S	S	S	B	S	
	4	240288	256680	B8/S234567	D	D	S	S	S	S	S	S	S	B	S	conserves 2x2 patterns with ≥ 3 1's
	4	21847		B01234567/S0	B	F	F	F	F	F	F	F	F	F	D	
	4	21845	87383	B01234567	F	F	F	F	F	F	F	F	F	F	D	Identical to 21847
5	174754	191146	S02345678	S	D	S	S	S	S	S	S	S	S	S	conserves 2x2 patterns with ≥ 3 1's	
5	174752	256682	S2345678	D	D	S	S	S	S	S	S	S	S	S	identical to 174754	
8	240296		B8/S12345678	D	S	S	S	S	S	S	S	S	S	B		
2 × 3	1	152918		B1234567/S8	S	F	F	F	F	F	F	F	F	S		
	1	152919	21846	B01234567/S08	B	F	F	F	F	F	F	F	F	S	identical to 152918	
	1	152916	218454	B1234567/S8	D	F	F	F	F	F	F	F	F	S	identical to 152918	
	1	152917	87382	B01234567/S8	F	F	F	F	F	F	F	F	F	S	identical to 152918	
	1	240291	60072	B08/S234567	B	D	S	S	S	S	S	S	S	B	conserves the difference between [101 111] and [111 101]	
	1	240289	125608	B08/S2345678	F	D	S	S	S	S	S	S	S	B	identical to 240291	
	1	256683	43680	B078/S012345678	B	S	S	S	S	S	S	S	S	B	B	
	1	256681	109216	B078/S12345678	F	S	S	S	S	S	S	S	S	B	B	identical to 256683
	1	218455	21844	B012345678/S08	B	F	F	F	F	F	F	F	F	F	B	identical to 152918
	1	218452		B12345678/S8	D	F	F	F	F	F	F	F	F	F	B	identical to 152918
	1	218435	87380	B012345678/S8	F	F	F	F	F	F	F	F	F	F	B	identical to 152918
2	43691		B0/S01234567	B	S	S	S	S	S	S	S	S	D			
2	43689	109227	B0/S1234567	F	S	S	S	S	S	S	S	S	D	identical to 43691		
2	109225		B08/S1234567	F	S	S	S	S	S	S	S	S	F	Identical to 43691		
7	240299	43688	B08/S012345678	B	S	S	S	S	S	S	S	S	B	B		
7	240297	109224	B08/S12345678	F	S	S	S	S	S	S	S	S	B	B	identical to 240299	
3 × 3	1	196607	2	B01234567/S8	B	B	B	B	B	B	B	B	B	S	conserves ring of 1's around a 0	
	1	174783	2730	B012/S012345678	B	B	B	S	S	S	S	S	S	S	S	
	1	240303	10920	B018/S012345678	B	B	S	S	S	S	S	S	S	B	conserves [001 011 010]	
	1	240295	27304	B018/S02345678	B	F	S	S	S	S	S	S	S	B	identical to 240303	
	1	43695	10923	B01/S01234567	B	B	S	S	S	S	S	S	S	D	identical to 240303	
	1	43687	27307	B01/S234567	B	F	S	S	S	S	S	S	S	D	identical to 240303	
	1	109231	10921	B018/S01234567	F	S	S	S	S	S	S	S	S	D	conserves [011 100 101]	
	1	109223	27305	B018/S0234567	B	F	S	S	S	S	S	S	S	F	identical to 109231	
	9	60075	43683	B07/S01234567	B	S	S	S	S	S	S	S	S	B	D	
	9	125611	43681	B078/S01234567	B	S	S	S	S	S	S	S	S	B	F	identical to 43683
	9	60073	109219	B07/S1234567	F	S	S	S	S	S	S	S	S	B	D	
9	60775	43683	B07/S01234567	B	S	S	S	S	S	S	S	S	B	D	Identical to 43683	
11	191145	109218	B07/S12345678	F	S	S	S	S	S	S	S	S	B	S		

TABLE 2
Conserved functions of order ≤ 9 for life-like CA's (i.e. 2D, 2-color, outer totalistic)

CA Basis	CA Basis	CA Basis
174762 $f(x) = H(x, \begin{smallmatrix} 1 \\ 1 \end{smallmatrix})$	240288 $f_1(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$	174752 $f_1(x) = H(x, \begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix})$
87381 $f(x) = H(x, \begin{smallmatrix} 1 \\ 0 \end{smallmatrix})$	$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	$f_2(x) = H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$
174760 $f(x) = H(x, \begin{smallmatrix} 1 \\ 1 \end{smallmatrix})$	$f_2(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$	$f_3(x) = H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$
174720 $f(x) = H(x, \begin{smallmatrix} 1 \\ 1 \end{smallmatrix})$	$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	$f_4(x) = H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$
174722	$f_3(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$	$f_5(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$
174728	$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	
174730	$f_4(x) = H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	
21845 $f_1(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$	$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	
21847	$f_4(x) = H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	
$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	$+ 3H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	
$f_2(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$		
$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$		
$f_3(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$	240296 $f_1(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$	218453 $f(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix})$
$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	218452 $H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	$f_2(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$	218455 $+ H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
$f_4(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$	$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	152917 $H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
$+ H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$	$f_3(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$	152916 $H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
	$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	152919 $H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
	$f_4(x) = H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	152918 $H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
191144 $f_1(x) = H(x, \begin{smallmatrix} 0 \\ 0 \end{smallmatrix})$	$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	256681 $f(x) = H(x, \begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix})$
$- H(x, \begin{smallmatrix} 0 \\ 0 \end{smallmatrix})$	$f_4(x) = H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	256683 $+ H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
$f_2(x) = H(x, \begin{smallmatrix} 0 \\ 0 \end{smallmatrix})$	$f_5(x) = H(x, \begin{smallmatrix} 0 \\ 0 \end{smallmatrix})$	191145 $- H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
$- H(x, \begin{smallmatrix} 0 \\ 0 \end{smallmatrix})$	$f_6(x) = H(x, \begin{smallmatrix} 0 \\ 0 \end{smallmatrix})$	$- H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
$f_3(x) = H(x, \begin{smallmatrix} 0 \\ 1 \end{smallmatrix})$	$f_7(x) = H(x, \begin{smallmatrix} 0 \\ 1 \end{smallmatrix})$	240289 $f(x) = H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
$- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$	$+ 3H(x, \begin{smallmatrix} 0 \\ 1 \end{smallmatrix})$	240291 $- H(x, \begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix})$
$f_4(x) = H(x, \begin{smallmatrix} 0 \\ 1 \end{smallmatrix})$	$f_8(x) = H(x, \begin{smallmatrix} 0 \\ 1 \end{smallmatrix})$	
$+ H(x, \begin{smallmatrix} 1 \\ 1 \end{smallmatrix})$		
$- H(x, \begin{smallmatrix} 1 \\ 1 \end{smallmatrix})$		
$+ 2H(x, \begin{smallmatrix} 1 \\ 0 \end{smallmatrix})$		
$+ 2H(x, \begin{smallmatrix} 1 \\ 0 \end{smallmatrix})$		
		109225 $f_1(x) = H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
		43689 $f_2(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$
		43691 $f_2(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$
		240297 $f_1(x) = H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
		240299 $- H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
		$f_2(x) = H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
		$f_3(x) = H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
		$f_4(x) = H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
		$f_5(x) = H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
		$f_6(x) = H(x, \begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix})$
		$f_7(x) = H(x, \begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix})$

FIGURE 2

2D Basis functions for life-like CA's. For each CA, this lists the lowest-order non-trivial conserved functions. The functions are simplified in some cases by rewriting in terms of *, and by linear combinations with some of the trivials. The given functions, combined with the trivials, constitute a basis set for the space of all conserved functions for that CA. The table contains all of the non-isomorphic, life-like CAs that have nontrivials of size 2×3 or smaller (the 3×3 nontrivials are shown in Figure 3).

CA Basis	CA Basis
109223 $f(x) = H(x, 0\mathbb{1}0)$	174783 $f(x) = H(x, 000) + H(x, 00\mathbb{1}) + H(x, 000) - H(x, 000)$
109231 $\mathbb{1}0\mathbb{1}$	$00\mathbb{1}$
43687 $\mathbb{1}0\mathbb{1}$	$00\mathbb{1}$
43695 $\mathbb{1}0\mathbb{1}$	$00\mathbb{1}$
240295 $\mathbb{1}0\mathbb{1}$	$00\mathbb{1}$
240303 $\mathbb{1}0\mathbb{1}$	$00\mathbb{1}$
196607 $f(x) = H(x, \mathbb{1}0\mathbb{1})$	$-H(x, 00\mathbb{1}) - H(x, 000) + H(x, 000) - H(x, 00\mathbb{1})$
125609 $f_1(x) = H(x, \mathbb{1}0\mathbb{1})$	$00\mathbb{1}$
125611 $\mathbb{1}0\mathbb{1}$	$00\mathbb{1}$
60075 $\mathbb{1}0\mathbb{1}$	$00\mathbb{1}$
$f_2(x) = H(x, \mathbb{1}0\mathbb{1})$	$-H(x, 00\mathbb{1}) + H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - 2H(x, 00\mathbb{1})$
$f_3(x) = H(x, \mathbb{1}0\mathbb{1})$	$00\mathbb{1}$
$f_4(x) = H(x, \mathbb{1}0\mathbb{1})$	$-H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - 2H(x, 00\mathbb{1})$
$f_5(x) = H(x, \mathbb{1}0\mathbb{1})$	$00\mathbb{1}$
$f_6(x) = H(x, \mathbb{1}0\mathbb{1})$	$+2H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - H(x, 00\mathbb{1})$
$f_7(x) = H(x, 00\mathbb{1})$	$00\mathbb{1}$
$f_8(x) = H(x, 00\mathbb{1})$	$-H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - H(x, 00\mathbb{1})$
$f_9(x) = H(x, 000)$	$00\mathbb{1}$
60073 $f_1(x) = H(x, \mathbb{1}0\mathbb{1})$	$00\mathbb{1}$
$f_2(x) = H(x, \mathbb{1}0\mathbb{1})$	$-H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - H(x, 00\mathbb{1})$
$f_3(x) = H(x, \mathbb{1}0\mathbb{1})$	$00\mathbb{1}$
$f_4(x) = H(x, \mathbb{1}0\mathbb{1})$	$-H(x, 00\mathbb{1}) + H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - H(x, 00\mathbb{1})$
$f_5(x) = H(x, \mathbb{1}0\mathbb{1})$	$00\mathbb{1}$
$f_6(x) = H(x, \mathbb{1}0\mathbb{1})$	$-H(x, 00\mathbb{1}) + H(x, 00\mathbb{1}) + H(x, 00\mathbb{1}) - 2H(x, 00\mathbb{1})$
$f_7(x) = H(x, \mathbb{1}0\mathbb{1})$	$00\mathbb{1}$
$f_8(x) = H(x, \mathbb{1}0\mathbb{1})$	$-H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) - H(x, 00\mathbb{1}) + H(x, 00\mathbb{1})$
$f_9(x) = H(x, 00\mathbb{1})$	$00\mathbb{1}$

FIGURE 3
2D Basis functions (continued). The 3×3 nontrivials, continued from figure 2.

Order	CA(isomorphs)	Rule		Order	CA(isomorphs)	Rule	
∞	0(255)	0		2	12(68,207,221)	Xy	
∞	8(64,239,253)	Xyz		2	14(84,143,213)	X+XYZ	
∞	30(86,135,149)	X+YZ		2	15(85)	X	
∞	32(251)	xYz		2	34(48,187,243)	Yz	
∞	40(96,235,249)	xz+yZ		2	35(49,59,115)	xYZ+Y	
∞	60(102,153,195)	x+y		2	42(112,171,241)	xyz+z	
∞	90(165)	x+z		2	43(113)	xY+Xz+YZ	
∞	106(120,169,225)	xy+z		2	51	Y	
∞	128(254)	xyz		2	140(196,206,220)	xyZ+y	
∞	136(192,238,252)	yz		2	142(212)	xy+Xz+yZ	
∞	150	x+y+z		2	200(236)	XyZ+y	
∞	154(166,180,210)	xY+z		3	2(16,191,247)	XyZ	
∞	160(250)	xz		3	3(17,63,119)	Xy	
∞	168(224,234,248)	XYz+z		3	4(223)	XyZ	
>16	6(20,159,215)	Xy+Xz		3	10(80,175,245)	Xz	
>16	9(65,111,125)	Xy+XZ		3	56(98,185,227)	xY+Xyz	
>16	13(69,79,93)	X+XYZ		3	76(205)	xyz+y	
>16	18(183)	xY+Yz		3	138(174,208,244)	xYZ+z	
>16	22(151)	X+Xyz+YZ		3	172(202,216,228)	Xy+xz	
>16	25(61,67,103)	Xyz+YZ		4	1(127)	XYZ	
>16	26(82,167,181)	xYZ+Xz		4	11(47,81,117)	X+XyZ	
>16	28(70,157,199)	Xy+xYZ		4	27(39,53,83)	Xz+YZ	
>16	37(91)	xYZ+XZ		4	29(71)	Xy+YZ	
>16	41(97,107,121)	X+XyZ+Yz		4	38(52,155,211)	XyZ+Yz	
>16	45(75,89,101)	X+Yz		4	46(116,139,209)	Xy+Yz	
>16	54(147)	XZ+Y		4	72(237)	xy+yZ	
>16	57(99)	Xz+Y		5	5(95)	XZ	
>16	58(114,163,177)	xY+Xz		5	19(55)	xYZ+Y	
>16	62(118,131,145)	x+XYz+y		5	24(66,189,231)	xYZ+Xyz	
>16	74(88,173,229)	xyZ+Xz		5	36(219)	xYz+XyZ	
>16	78(92,141,197)	Xz+yZ		5	108(201)	xz+y	
>16	105	x+y+Z		5	132(222)	xy+yZ	
>16	110(124,137,193)	Xyz+y+z		6	23	xY+XZ+Yz	
>16	122(161)	x+XYz+z		6	50(179)	XYZ+Y	
>16	126(129)	xY+Xz+yZ		6	77	xy+XZ+yZ	
>16	130(144,190,246)	xz+YZ		6	178	xy+xZ+YZ	
>16	134(148,158,214)	X+XYZ+yz		6	232	xy+xz+yZ	
>16	146(182)	x+xyZ+YZ		8	44(100,203,217)	Xy+xYZ	
>16	152(188,194,230)	xYZ+yZ		8	73(109)	X+XyZ+yZ	
>16	156(198)	xZ+y		9	7(21,31,87)	X+Xyz	
>16	162(176,186,242)	Xyz+z		12	33(123)	xY+YZ	
1	170(240)	z		13	164(218)	XyZ+xz	
1	184(226)	xY+yZ		14	94(133)	x+XyZ+z	
1	204	y		14	104(233)	x+xYZ+yZ	

FIGURE 4

Summary of results for the elementary CAs (i.e. 1D, 2-color, neighborhood of 3 cells). In each half of the table, the first column gives the energy window size for the smallest nontrivial. A value of ∞ indicates that it is known no nontrivial can exist. A value of > 16 indicates that no nontrivial exists with energy window of size 16 or below, but it is unknown whether larger ones exist. The next column has the CA rule number (with isomorphs in parentheses). The next is the formula for the successor function, where cells have state 0 or 1, three consecutive cells are called x , y , z (with capitalized inverses, so $X=1-x$ etc.), and the formula modulo 2 gives the new state for y . Finally, the successor function is shown graphically, giving the new state as a function of the state in that cell and its immediate neighbors (shown at the top of the column).

CA Basis	CABasis	CABasis
170 $f(x) = H(x, \mathbb{1})$	24 $f(x) = H(x, \mathbb{1}000\mathbb{1})$	33 $f(x) = H(x, 000\mathbb{1}000\mathbb{1}00\mathbb{1})$
184	$+H(x, \mathbb{1}00\mathbb{1}0)$	$-H(x, 000\mathbb{1}000\mathbb{1}00\mathbb{1})$
204	$+H(x, \mathbb{1}10\mathbb{1})$	$+2H(x, 000\mathbb{1}000\mathbb{1}1\mathbb{1})$
12 $f(x) = H(x, \mathbb{1}\mathbb{1}0)$	$+H(x, \mathbb{1}\mathbb{1}0\mathbb{1}0)$	$-H(x, 000\mathbb{1}000\mathbb{1}1\mathbb{1})$
14		$-H(x, 000\mathbb{1}000\mathbb{1}0\mathbb{1}0)$
15	36 $f(x) = H(x, 00\mathbb{1}00)$	$-H(x, 000\mathbb{1}000\mathbb{1}10\mathbb{1}0)$
34	$+H(x, \mathbb{1}\mathbb{1}0\mathbb{1}\mathbb{1})$	$+H(x, 000\mathbb{1}01000\mathbb{1})$
35		$+H(x, 000\mathbb{1}0100\mathbb{1}\mathbb{1})$
42		$+H(x, 000\mathbb{1}0100\mathbb{1}\mathbb{1})$
43	108 $f(x) = H(x, \mathbb{1}0\mathbb{1}00)$	$+2H(x, \mathbb{1}00\mathbb{1}0100000)$
51	$+H(x, \mathbb{1}0\mathbb{1}00)$	$-H(x, \mathbb{1}00\mathbb{1}0100000)$
140	$+H(x, \mathbb{1}0\mathbb{1}\mathbb{1})$	$+H(x, 000\mathbb{1}01000\mathbb{1}\mathbb{1})$
142	$+2H(x, \mathbb{1}100\mathbb{1})$	$+H(x, 000\mathbb{1}0101000)$
200 $f(x) = H(x, \mathbb{1}\mathbb{1}\mathbb{1})$	$+H(x, \mathbb{1}110\mathbb{1})$	$+H(x, 000\mathbb{1}010100\mathbb{1}0)$
2 $f(x) = H(x, \mathbb{1}\mathbb{1}00)$		$+H(x, 000\mathbb{1}01011\mathbb{1}\mathbb{1})$
3	132 $f(x) = H(x, 0\mathbb{1}0\mathbb{1}0)$	$+H(x, 000\mathbb{1}000\mathbb{1}0100)$
172		$+H(x, 0\mathbb{1}01000\mathbb{1}00)$
4 $f(x) = H(x, 0\mathbb{1}0)$		$+H(x, 0\mathbb{1}0101000\mathbb{1})$
	23 $f(x) = H(x, 00\mathbb{1}\mathbb{1}00)$	$+H(x, 0\mathbb{1}01100\mathbb{1}00)$
10 $f(x) = H(x, \mathbb{1}\mathbb{1}0)$	$+H(x, \mathbb{1}\mathbb{1}00\mathbb{1}\mathbb{1})$	$+3H(x, 001\mathbb{1}01000\mathbb{1})$
56 $f(x) = H(x, \mathbb{1}\mathbb{1}0)$		$-H(x, 001\mathbb{1}01000\mathbb{1})$
76 $+H(x, \mathbb{1}\mathbb{1}0)$	50 $f(x) = H(x, 0\mathbb{1}\mathbb{1}00\mathbb{1})$	$+H(x, 001\mathbb{1}01000\mathbb{1})$
	$+H(x, \mathbb{1}00\mathbb{1}\mathbb{1}0)$	$+3H(x, 001\mathbb{1}0100\mathbb{1}0)$
138 $f(x) = H(x, \mathbb{1}00)$	77 $f_1(x) = H(x, 0\mathbb{1}\mathbb{1}00\mathbb{1})$	$-H(x, 001\mathbb{1}0100\mathbb{1}0)$
$+H(x, \mathbb{1}\mathbb{1})$	$f_2(x) = H(x, \mathbb{1}00\mathbb{1}\mathbb{1}0)$	$+H(x, \mathbb{1}0000\mathbb{1}00\mathbb{1})$
		$+H(x, \mathbb{1}0000\mathbb{1}0\mathbb{1}\mathbb{1})$
1 $f(x) = H(x, \mathbb{1}000)$	232 $f_1(x) = H(x, 0\mathbb{1}\mathbb{1}000)$	$+2H(x, \mathbb{1}0000\mathbb{1}000\mathbb{1})$
	$-H(x, \mathbb{1}0\mathbb{1}00)$	$-H(x, \mathbb{1}0000\mathbb{1}000\mathbb{1}0)$
11 $f(x) = H(x, \mathbb{1}00\mathbb{1})$	$+H(x, \mathbb{1}100\mathbb{1})$	$+H(x, \mathbb{1}0000\mathbb{1}1000\mathbb{1}0)$
27 $+H(x, \mathbb{1}0\mathbb{1}\mathbb{1})$	$-H(x, \mathbb{1}1100\mathbb{1})$	$+2H(x, 000\mathbb{1}0100\mathbb{1}\mathbb{1})$
	$f_2(x) = H(x, \mathbb{1}\mathbb{1}00\mathbb{1}\mathbb{1})$	$-H(x, 000\mathbb{1}0100\mathbb{1}0)$
29 $f(x) = H(x, \mathbb{1}\mathbb{1}00)$	44 $f(x) = H(x, \mathbb{1}000\mathbb{1}\mathbb{1}\mathbb{1})$	$-H(x, 000\mathbb{1}01000\mathbb{1}\mathbb{1})$
$+H(x, \mathbb{1}\mathbb{1}00)$	$+H(x, \mathbb{1}0\mathbb{1}00\mathbb{1}\mathbb{1})$	$+H(x, \mathbb{1}01001000\mathbb{1}\mathbb{1})$
$+H(x, \mathbb{1}0\mathbb{1}\mathbb{1})$	$+H(x, \mathbb{1}010\mathbb{1}0\mathbb{1})$	$+H(x, \mathbb{1}00001000\mathbb{1}\mathbb{1})$
	$+H(x, \mathbb{1}01\mathbb{1}0\mathbb{1})$	$+H(x, \mathbb{1}00001000\mathbb{1}\mathbb{1})$
38 $f(x) = H(x, \mathbb{1}\mathbb{1}00\mathbb{1})$	$+H(x, \mathbb{1}\mathbb{1}00\mathbb{1}\mathbb{1})$	$+2H(x, \mathbb{1}1000000\mathbb{1}\mathbb{1})$
46 $+H(x, \mathbb{1}\mathbb{1}0\mathbb{1})$	73 $f(x) = H(x, 0\mathbb{1}\mathbb{1}00\mathbb{1}\mathbb{1}0)$	$-H(x, \mathbb{1}1000000\mathbb{1}\mathbb{1})$
		$+2H(x, \mathbb{1}1010000\mathbb{1}\mathbb{1})$
72 $f(x) = H(x, 0\mathbb{1}\mathbb{1}0)$	7 $f(x) = H(x, 000\mathbb{1}\mathbb{1}000\mathbb{1})$	$-H(x, \mathbb{1}1010000\mathbb{1}\mathbb{1})$
	$+H(x, 000\mathbb{1}\mathbb{1}00\mathbb{1}0)$	$+2H(x, \mathbb{1}101000\mathbb{1}\mathbb{1})$
5 $f_1(x) = H(x, 0\mathbb{1}0\mathbb{1}\mathbb{1})$	$+H(x, 000\mathbb{1}\mathbb{1}000\mathbb{1}\mathbb{1})$	$-H(x, \mathbb{1}1011000\mathbb{1}\mathbb{1})$
$+2H(x, \mathbb{1}0\mathbb{1}0\mathbb{1})$	$+H(x, \mathbb{1}1000\mathbb{1}\mathbb{1}\mathbb{1})$	$+2H(x, \mathbb{1}101100\mathbb{1}\mathbb{1})$
$-H(x, \mathbb{1}0\mathbb{1}0\mathbb{1})$	$+H(x, \mathbb{1}1000\mathbb{1}\mathbb{1}\mathbb{1})$	$-H(x, \mathbb{1}101100\mathbb{1}\mathbb{1})$
$+H(x, \mathbb{1}0\mathbb{1}0\mathbb{1})$	$+H(x, \mathbb{1}1000\mathbb{1}\mathbb{1}\mathbb{1})$	$+2H(x, \mathbb{1}101100\mathbb{1}\mathbb{1})$
$f_2(x) = H(x, \mathbb{1}\mathbb{1}00\mathbb{1}0)$		$-H(x, \mathbb{1}101100\mathbb{1}\mathbb{1})$
19 $f(x) = H(x, \mathbb{1}\mathbb{1}000)$		$+2H(x, \mathbb{1}101100\mathbb{1}\mathbb{1})$
$+H(x, \mathbb{1}\mathbb{1}00\mathbb{1})$		$-H(x, \mathbb{1}101100\mathbb{1}\mathbb{1})$
		$-H(x, \mathbb{1}101100\mathbb{1}\mathbb{1})$
		164 $f(x) = H(x, 00\mathbb{1}\mathbb{1}00\mathbb{1}00\mathbb{1}00\mathbb{1})$
		$+H(x, 0\mathbb{1}\mathbb{1}0\mathbb{1}0\mathbb{1}\mathbb{1}0\mathbb{1})$
		94 $f(x) = H(x, 0\mathbb{1}\mathbb{1}00\mathbb{1}0\mathbb{1}\mathbb{1}0\mathbb{1}0\mathbb{1})$
		$+H(x, \mathbb{1}0\mathbb{1}00\mathbb{1}0\mathbb{1}\mathbb{1}0\mathbb{1})$
		$+H(x, \mathbb{1}0\mathbb{1}110\mathbb{1}00\mathbb{1}0\mathbb{1})$
		$+H(x, \mathbb{1}0\mathbb{1}110\mathbb{1}00\mathbb{1}0\mathbb{1})$
		104 $f_1(x) = H(x, 00\mathbb{1}\mathbb{1}00\mathbb{1}0\mathbb{1}0\mathbb{1}0\mathbb{1})$
		$+H(x, 00\mathbb{1}\mathbb{1}00\mathbb{1}0\mathbb{1}0\mathbb{1}0\mathbb{1})$
		$+H(x, 00\mathbb{1}\mathbb{1}00\mathbb{1}00\mathbb{1}0\mathbb{1})$
		$+H(x, 00\mathbb{1}\mathbb{1}00\mathbb{1}00\mathbb{1}0\mathbb{1})$
		$+H(x, 00\mathbb{1}\mathbb{1}00\mathbb{1}00\mathbb{1}0\mathbb{1})$
		$+H(x, 00\mathbb{1}\mathbb{1}00\mathbb{1}00\mathbb{1}0\mathbb{1})$
		$f_2(x) = H(x, 00\mathbb{1}\mathbb{1}00\mathbb{1}00\mathbb{1}0\mathbb{1})$

FIGURE 5
 1D Basis functions. For each CA, this lists the lowest-order nontrivial conserved functions. The given functions, combined with the trivials, constitute a basis set for the space of all conserved functions for that CA. Of the 88 isomorphic classes for elementary CAs, the 47 listed are all that have nontrivials of size 16 or less.

xSSSSSSSB, and xDSSSSSSB have nontrivial conserved functions. All CA's of the form xSSSSSSSx have a nontrivial as well, unless exactly one of the x's is 'S'.

Finally, our results show that all known nontrivials for life-like CAs correspond to energy windows for which the width and the height differ by no more than one. Whether this holds true for all larger nontrivials on life-like CA's remains an open question.

5 THE GAME OF LIFE

Because of the special significance of Conway's Game of Life (rule B3/S23, rule vector DDSBDDDDD), we have examined it for nontrivial conserved energy functions up to order 13. We verified by exhaustive search that none exist up to that size.

6 CONCLUSIONS AND FUTURE WORK

Table 2 and Figures 2 through 4 represent a complete list of basis sets for all known nontrivial conserved functions for elementary and life-like cellular automata up to isomorphism. We have discussed some of the patterns we have observed.

We previously defined the notion of core nontrivials [2], recognizing that cellular automata could exhibit different nontrivials of higher orders that are not simple extensions of lower ones. We have yet to apply this idea to the life-like automata shown here. Thus the functions we report are only the first core nontrivials found for each CA. So for each elementary or life-like CA, we give the complete basis set for all nontrivials of the minimum order for which nontrivials exist.

Number-conserving 1D cellular automata [3] are automata with transition rules that conserve the sum of the states in a neighborhood. That sum is one kind of conserved energy function as defined in Definition 1.10, where the function is simply the sum of all terms in the window. Our work therefore includes number-conserving as a special case. The theory described here applies to all cellular automata with finite states and arbitrary dimensionality. The results for 2D automata are all new.

Continuing improvements in computing power and further refinements of our code should enable us to identify nontrivials at increasingly higher orders. The existence of nontrivials for $m \times n$ energy windows with $|m - n| > 1$

remains an open question. Higher dimensional CAs, non-totalistic CAs, and k-colored CAs could also be explored.

As yet, an elegant, unifying description of cellular automata relating their decision rules and a given energy window to a nontrivial conserved function remains elusive. While the general problem of whether a CA has any nontrivials of any size is undecidable [5], we have mapped out the space for lower orders and life-like CAs well enough to suggest some ideas for a more elegant classification scheme than the present ad hoc one we are currently forced to adopt. Such a scheme may in fact exist, or it may remain forever elusive, an fundamentally complex property inherent in the nature of computational automata. We hope further work may yet resolve this question.

7 ACKNOWLEDGMENTS

The authors are grateful for the support of the Air Force Academy Center for Cyberspace Research, and to the reviewers for their helpful comments.

REFERENCES

- [1] L. Baird. (2008). The cellular automata $*0*0*000$ have no nontrivial conserved functions. Technical Report USAFA-TR-2008-ACCR-04, U. S. Air Force Academy, Academy Center for Cyberspace Research.
- [2] L. Baird and B. Fagin. (2008). Conservation functions for 1-d automata: Efficient algorithms, new results, and a partial taxonomy. *Journal of Cellular Automata*, 3(4):271–288.
- [3] N. Boccara and H. Fuks. (2002). Number-conserving cellular automaton rules. *Fundam. Inform.*, 52(1-3):1–13.
- [4] B. Fagin and L. Baird. (2007). New higher-order conservation functions for 1-d cellular automata. In *Proceedings of the IEEE Symposium on Artificial Life*.
- [5] E. Formenti, J. Kari, and S. Taati. (2008). The most general conservation law for a cellular automaton. In *Proceedings of the 3rd international conference on Computer science: theory and applications*, CSR'08, pages 194–203, Berlin, Heidelberg. Springer-Verlag.
- [6] T. Hattori and S Takesue. (1991). Additive conserved quantities in discrete-time lattice dynamical systems. *Physica D*, 49:295–322.
- [7] Shinji Takesue, (2008). personal communication.