

# Newton Polytopes of Two-Dimensional Hidden Markov Models

Jay Wierer and Nigel Boston

## CONTENTS

1. Introduction
  2. Properties of Newton Polytopes for 2D HMMs
  3. Bounding Polytopes for the Newton Polytope
  4. Newton Polytopes and Orbits
  5. Number of Vertices of Newton Polytopes
  6. Character-Recognition Algorithm
- References

---

In this paper, we show that the Newton polytope of an observation  $Y$  from a two-dimensional hidden Markov model (2D HMM) lies in a three-dimensional subspace of its ambient eight-dimensional space, whose vertices correspond to the most likely explanations (“hidden” states) for  $Y$  given the model. For each Newton polytope, there exists a set of “essential” vertices, which form a skeleton for the polytope. All observations in the same orbit (identical under translations, rotations, and transpositions) have the same Newton polytope. Our main conjecture is that the maximal number of vertices of any Newton polytope is of order  $n^2$ .

---

## 1. INTRODUCTION

### 1.1 Graphical Models

A *graphical model* [Murphy 98] is a graph in which nodes represent random variables and the (lack of) arcs represent conditional independence assumptions. In particular, a *directed* graphical model is one in which a node is independent of its ancestors given its parents. For example, a one-dimensional hidden Markov model (HMM) is a directed graphical model and is given by the following: An observation bit  $Y_i$  is dependent only on its corresponding hidden state  $X_i$ , that is,  $P(Y_i = y_i \mid Y_{i-1} = y_{i-1}, \dots, Y_1 = y_1, X_i = x_i, X_{i-1} = x_{i-1}, \dots, X_1 = x_1) = P(Y_i = y_i \mid X_i = x_i)$ , and the current state  $X_i$  is independent of the first  $i - 2$  states given the previous state  $X_{i-1}$ , that is,  $P(X_i = x_i \mid X_{i-1} = x_{i-1}, X_{i-2} = x_{i-2}, \dots, X_1 = x_1) = P(X_i = x_i \mid X_{i-1} = x_{i-1})$ . See Figure 1 for a graphical representation of the HMM.

### 1.2 Tropical Geometry and Newton Polytopes

When making an inference on a statistical (graphical) model, one typically seeks to maximize the probability  $f_Y = P(Y)$ , where  $Y = [Y_i]$  is an observation from the model. However, it is often simpler to look at  $L(Y) = -\log P(Y)$ , which is then minimized to find the best explanation for  $Y_1 \cdots Y_n$ . If  $f_Y$  can be expressed in

2000 AMS Subject Classification: Primary 52B12

Keywords: Newton polytopes, tropical geometry, hidden Markov models, graphical models

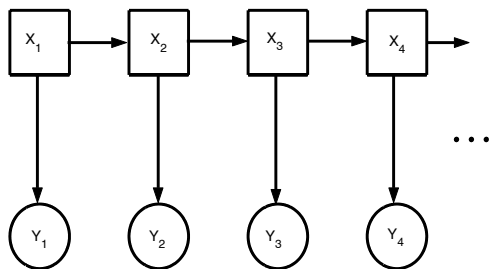


FIGURE 1. The hidden Markov model (HMM).

polynomial form, it can be *tropicalized* by replacing the operators  $(+, \times)$  with the operators  $(\min, +)$ . Minimizing the tropicalized map  $g$  of  $f$  is essentially equivalent to minimizing  $L(Y) = -\log(P(Y))$ .

**Definition 1.1.** Given a polynomial map  $f : \mathbf{R}^m \rightarrow \mathbf{R}$  that can be expressed as

$$f = \sum_{i=1}^q c_i x_1^{\nu_{i,1}} x_2^{\nu_{i,2}} \cdots x_m^{\nu_{i,m}},$$

the *Newton polytope*  $\text{NP}(f)$  of  $f$  is defined as the convex hull of the point set  $N = \{(\nu_{i,1}, \nu_{i,2}, \dots, \nu_{i,m}) \mid i = 1, \dots, q\}$ , i.e.,  $\text{NP}(f) = \text{conv}(N)$ .

An important property of the Newton polytope  $\text{NP}(f)$  is that the tropical morphism  $g$  of  $f$  is linear on each cone in the normal fan of  $\text{NP}(f)$  [Pachter and Sturmfels 04, Theorem 2]. Furthermore, each vertex of the Newton polytope indexes its corresponding normal cone in which  $g$  is minimized. This property will be useful to us in our character-recognition algorithm, which is currently under development.

### 1.3 Two-Dimensional Hidden Markov Model (2D HMM)

The two-dimensional hidden Markov model that we will use has the following properties:

1. Each hidden state  $X_{i,j}$  depends on its “past” only through its immediate left and immediate top neighboring states  $X_{(i-1)_n,j}$  and  $X_{i,(j-1)_n}$ , where  $(x)_n$  is defined to be  $x \pmod n$ .
2. Each observation bit  $Y_{i,j}$  depends only on its corresponding hidden state  $X_{i,j}$ .
3. All random variables are binary variables (having values 0 or 1).

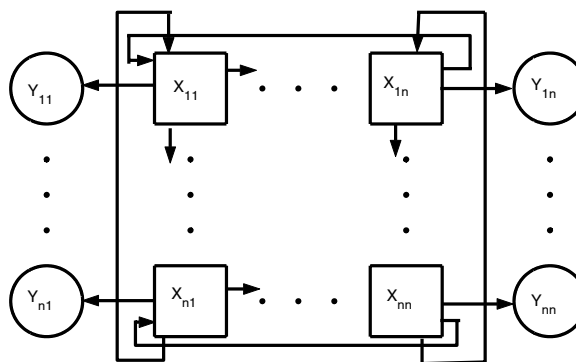


FIGURE 2. Graphical representation of an  $n \times n$  two-dimensional HMM.

4. We define  $P(X_{(i+1)_n,j} = l \mid X_{i,j} = k) = a_{k,l}$ ,  $P(X_{i,(j+1)_n} = l \mid X_{i,j} = k) = a_{k,l}$ , and  $P(Y_{i,j} = l \mid X_{i,j} = k) = b_{k,l}$ .

For a fixed observation  $Y = [y_{i,j}]$ , we have

$$\begin{aligned} f_Y &= P(Y) \\ &= \sum_{X \in X_n} \prod_{(i,j)} a_{x_{i,j}, x_{(i+1)_n,j}} a_{x_{i,j}, x_{i,(j+1)_n}} b_{x_{i,j}, y_{i,j}}, \end{aligned}$$

where  $X_n = \{0, 1\}^{n \times n}$ . The tropicalized map  $g_Y$  is then

$$g_Y = \min_{X \in X_n} \sum_{(i,j)} [s_{x_{i,j}, x_{(i+1)_n,j}} + s_{x_{i,j}, x_{i,(j+1)_n}} + t_{x_{i,j}, y_{i,j}}],$$

where  $s_{\cdot, \cdot} = -\log(a_{\cdot, \cdot})$  and  $t_{\cdot, \cdot} = -\log(b_{\cdot, \cdot})$ . Figure 2 gives a graphical representation of the 2D HMM.

**Example 1.2. (Newton polytope for a 2D HMM.)** Let

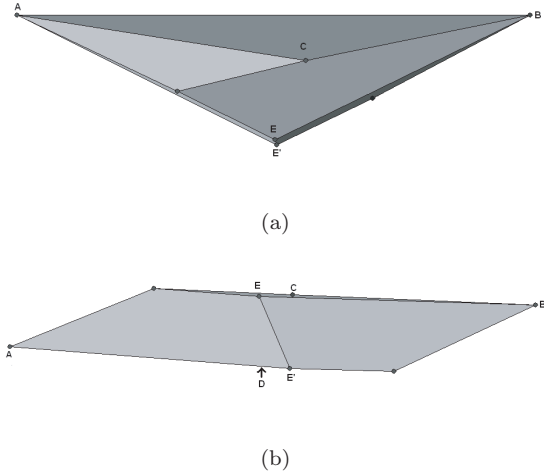
$$Y = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then  $\text{NP}(Y)$  has vertices

- $\mathbf{A} = (32, 0, 0, 0, 11, 5, 0, 0)$ ,
- $\mathbf{B} = (0, 0, 0, 32, 0, 0, 11, 5)$ ,
- $\mathbf{C} = (12, 10, 10, 0, 11, 0, 0, 5)$ ,
- $\mathbf{D} = (0, 10, 10, 12, 0, 5, 11, 0)$ ,
- $\mathbf{E} = (0, 16, 16, 0, 8, 0, 3, 5)$ ,
- $\mathbf{E}' = (0, 16, 16, 0, 3, 5, 8, 0)$ ,

and

$$(8, 6, 6, 12, 7, 0, 4, 5), \quad (12, 6, 6, 8, 4, 5, 7, 0).$$



**FIGURE 3.** NP( $Y$ ) for Example 1.2. (a) Top view; (b) Front view.

We will prove in the next section that these Newton polytopes are effectively three-dimensional; hence they can be visualized by any software package with 3D graphing capabilities, such as MATLAB or polymake. Using the first, second, and fifth entries of each vertex, we plot NP( $Y$ ) in Figure 3. Using terminology from the following section, we will call the all-zeros vertex **A**, the all-ones vertex **B**, the “self” vertex **C**, the “anti-self” vertex **D**, and the “checkerboard” or “near-checkerboard” vertex (or vertices) **E** (**E'**, **E''**, ...). The two remaining, unlabeled, vertices of NP( $Y$ ) are specific to  $Y$  and generally do not show up for every element in  $X_n$ .

## 2. PROPERTIES OF NEWTON POLYTOPES FOR 2D HMMS

**Definition 2.1.** Given any positive integer  $n$ , a candidate vertex  $u$  of a Newton polytope for a given hidden state grid  $X \in X_n$  and observation grid  $Y \in X_n$  is given by the powers of the monomial corresponding to  $X$  and  $Y$ ; that is, for the monomial  $a_{00}^{u_1} a_{01}^{u_2} a_{10}^{u_3} a_{11}^{u_4} b_{00}^{u_5} b_{01}^{u_6} b_{10}^{u_7} b_{11}^{u_8}$ , the associated candidate vertex is  $u = (u_1, \dots, u_8)$ .

**Lemma 2.2.** *The second and third coordinates (corresponding to the number of zero-to-one and one-to-zero state-to-state transitions, respectively) are always equal for any candidate vertex of any observation grid for any  $n = 1, 2, \dots$*

*Proof:* If there is no zero-to-one transition, then  $X$  must consist of all zeros or all ones. In either case there are

no zero-to-one or one-to-zero transitions. Otherwise, assume that there is at least one zero-to-one (or one-to-zero) transition for the hidden state grid  $X$  associated with a candidate vertex.

Find a row (or column) in  $X$  with a zero-to-one transition. Then  $x_{i,j} = 0$  and  $x_{i,(j+1) \pmod n} = 1$  for some  $i, j$ . Then there exists a  $k^* \leq n$  such that  $k^* = \arg \min_{k \geq 1} \{x_{i,(j+k) \pmod n} = 0\}$ .

Hence there is a one-to-zero transition corresponding to  $x_{i,(j+k^*-1) \pmod n} = 1$  and  $x_{i,(j+k^*) \pmod n} = 0$ . Thus each zero-to-one transition has a corresponding one-to-zero transition and vice versa.  $\square$

**Lemma 2.3.** *Given a positive integer  $n$ , for any candidate vertex  $u$  of any observation grid, we have  $u_1 + u_2 = 2(u_5 + u_6)$  and  $u_3 + u_4 = 2(u_7 + u_8)$ , where the  $u_i$  will be defined in the proof.*

*Proof:* For every state, there are two state-to-state transitions and one state-to-observation transition leaving it. Here  $u_1$  and  $u_2$  denote the number of state-to-state transitions leaving a zero state in the grid, and  $u_5$  and  $u_6$  denote the number of state-to-observation transitions leaving a zero state. Hence  $u_1 + u_2 = 2(u_5 + u_6)$ .

Likewise,  $u_3$  and  $u_4$  denote the number of state-to-state transitions leaving a one state in the grid. Finally,  $u_7$  and  $u_8$  denote the number of state-to-observation transitions from a one state. Hence  $u_3 + u_4 = 2(u_7 + u_8)$ .  $\square$

**Theorem 2.4.** *Given fixed positive integers  $n, k$ , the Newton polytope for any fixed observation grid  $Y \in X_n$  with  $k$  ones lies in a three-dimensional subspace of the ambient eight-dimensional space in which each vertex lies.*

*Proof:* Without loss of generality, we will assume that  $u_1, u_2$ , and  $u_5$  are the coordinates of each candidate vertex on which the remaining coordinates are linearly dependent. We have shown by the preceding lemmas that  $u_3$  and  $u_6$  are linearly dependent.

Furthermore, we know that  $u_1 + u_2 + u_3 + u_4 = 2n^2$ , the total number of state-to-state transitions, so  $u_4$  is linearly dependent. The observation  $Y$  has a fixed number  $k$  of ones, so we must have  $u_5 + u_7 = n^2 - k$  transitions to observation zeros and  $u_6 + u_8 = k$  transitions to observation ones; hence  $u_7$  and  $u_8$  are linearly dependent, the latter by Lemma 2.3.  $\square$

Not only are these Newton polytopes three-dimensional, they also have what we will call “essen-

tial” vertices, that is, vertices due to certain hidden state grids that occur in the Newton polytope for every observation. We first define the “checkerboard” and “near-checkerboard” grids.

**Definition 2.5.** For  $n$  even, the “checkerboard” grid  $X = [x_{i,j}]$  can be described by  $x_{i,j} = 1$  if  $i + j \equiv 0 \pmod{2}$  and  $x_{i,j} = 0$  otherwise.

**Remark 2.6.** By symmetry there exist two versions of the checkerboard grid, which yields the maximum number of zero-to-one (also one-to-zero) state-to-state transitions ( $n^2$ ) when  $n$  is even.

**Definition 2.7.** For  $n$  odd, the “near-checkerboard” grids  $X = [x_{i,j}]$  and  $Z = [z_{i,j}]$  can be described by  $x_{i,j} = 1$  if  $(j - i) \pmod{n} \equiv k$  for  $k = 0, 2, \dots, n - 3$  and  $x_{i,j} = 0$  otherwise, and  $z_{i,j} = 1 - x_{i,j} \pmod{2}$ .

**Remark 2.8.** The near-checkerboard grids for  $n$  odd yield the maximum number of zero-to-one (one-to-zero) state-to-state transitions ( $n^2 - n$ ). For a given row (or column) of a hidden state grid, when  $n$  is odd, the maximum number of zero-to-one input transitions is  $\frac{n-1}{2}$ , which is achieved by each row (and column) of  $X$  and  $Z$ .

Additionally, the near-checkerboard grids have the maximum number of zero-to-zero (one-to-one) state-to-state transitions ( $2n$ ) for all hidden state grids with the maximum number of zero-to-one (one-to-zero) state-to-state transitions. By symmetry there are  $2n$  shifted versions of the near-checkerboard grids.

**Proposition 2.9.** *The candidate vertices of the following hidden state grids are always vertices of any Newton polytope for a given observation  $Y$ :*

- A. *The all-zeros grid.*
- B. *The all-ones grid.*
- C. *The “self” grid  $Y$ .*
- D. *The “anti-self” grid  $W = [1 - y_{i,j}]$ .*
- E. *For  $n$  even, the checkerboard grid or its shifted version; for  $n$  odd, the near-checkerboard grids (or their shifted versions).*

We will refer to these as the “essential” vertices.

*Proof:* The all-zeros (all-ones) grid alone yields the maximal number of zero-to-zero (one-to-one) state-to-state

transitions  $2n^2$ , whence the first two vertices on our list. Also, the self (anti-self) grid causes the maximal number of zero-to-zero (zero-to-one) state-to-observation transitions  $n^2 - k$  along with the maximal number of one-to-one (one-to-zero) state-to-observation transitions  $k$ , whence the third and fourth vertices.

Since for  $n$  even and  $n$  odd, respectively, the checkerboard and near-checkerboard grids cause the maximal number of zero-to-one (one-to-zero) state-to-state transitions, there will be vertices due to the checkerboard or near-checkerboard grids for every Newton polytope.  $\square$

### 3. BOUNDING POLYTOPES FOR THE NEWTON POLYTOPE

For  $n \geq 6$  we have encountered difficulty computing the Newton polytopes for observations  $Y \in X_n$  (cf. Section 5).

It may be possible to use an approximation of the Newton polytope in our character-recognition algorithm, which is currently under development. Certainly there exists a lower-bound polytope  $P_{\text{lower}}$  such that  $P_{\text{lower}} \subseteq \text{NP}(Y)$ , which provides a first approximation of  $\text{NP}(Y)$ .

A good choice for a lower-bound polytope  $P_{\text{lower}}$  is the convex hull of the essential vertices for  $Y$  described in the previous section. Since each of the essential vertices is also a vertex of  $\text{NP}(Y)$ , then we must have  $P_{\text{lower}} \subseteq \text{NP}(Y)$ .

**Lemma 3.1.** *The polytope given by the convex hull of the essential vertices for a given observation  $Y$ ,  $P_{\text{lower}}$ , is a subset of the Newton polytope of  $Y$ ,  $\text{NP}(Y)$ .*

*Proof:* Since the collection of essential vertices is a subset of the vertices of  $\text{NP}(Y)$ , any convex combination of the essential vertices must also be in  $\text{NP}(Y)$ ; hence  $P_{\text{lower}} \subseteq \text{NP}(Y)$ .  $\square$

In addition, there exist many possibilities for an upper-bound polytope  $P_{\text{upper}}$ . One possibility for  $P_{\text{upper}}$  is the convex hull of the union of the Newton polytopes of all observations in  $X_n$  with the same number  $k$  of ones as  $Y$ . This polytope  $P_{\text{upper}}$  has the form given by the following theorem. For brevity we consider only  $1 \leq k < n$ .

**Theorem 3.2.** *Let  $m_n(k) = k - 1$  be the maximal number of one-to-one state-to-state transitions in an  $n \times n$  grid with  $1 \leq k < n$  ones. Let  $X_n^{(k)} = \{X \in X_n \mid X \text{ has } k \text{ ones}\}$ . The upper-bound polytope  $P_{\text{upper}} = \text{conv}(\bigcup_{X \in X_n^{(k)}} \text{NP}(X))$  is equal to  $P$ , which has vertices given in Table 1 and is graphically depicted in Figure 4.*

Grid	$u_1$	$u_2 = u_3$	$u_4$	$u_5$	$u_6$	$u_7$	$u_8$
1. All-zeros	$2n^2$	0	0	$n^2 - k$	$k$	0	0
2. All-ones	0	0	$2n^2$	0	0	$n^2 - k$	$k$
3. Maximum self	$2n^2 - 4k + m_n(k)$	$2k - m_n(k)$	$m_n(k)$	$n^2 - k$	0	0	$k$
4. Minimum self	$2n^2 - 4k$	$2k$	0	$n^2 - k$	0	0	$k$
5. Maximum anti-self	$m_n(k)$	$2k - m_n(k)$	$2n^2 - 4k + m_n(k)$	0	$k$	$n^2 - k$	0
6. Minimum anti-self	0	$2k$	$2n^2 - 4k$	0	$k$	$n^2 - k$	0
7a. CB ( $n$ even)	0	$n^2$	0	$\frac{n^2}{2}$	0	$\frac{n^2}{2} - k$	$k$
8a. CB ( $n$ even)	0	$n^2$	0	$\frac{n^2}{2} - k$	$k$	$\frac{n^2}{2}$	0
7b. NCB ( $n$ odd)	$2n$	$n^2 - n$	0	$\frac{n^2+n}{2} - k$	$k$	$\frac{n^2-n}{2}$	0
8b. NCB ( $n$ odd)	$2n$	$n^2 - n$	0	$\frac{n^2+n}{2}$	0	$\frac{n^2-n}{2} - k$	$k$
9b. NCB ( $n$ odd)	0	$n^2 - n$	$2n$	$\frac{n^2-n}{2}$	0	$\frac{n^2+n}{2} - k$	$k$
10b. NCB ( $n$ odd)	0	$n^2 - n$	$2n$	$\frac{n^2-n}{2} - k$	$k$	$\frac{n^2+n}{2}$	0

TABLE 1. Vertices of the upper bound polytope  $P$

*Proof:* First we will show that  $P \subseteq P_{\text{upper}}$ . If we show that each vertex  $v$  of  $P$  is contained in  $\text{NP}(X)$  for some  $X \in X_n^{(k)}$ , then certainly  $v \in \bigcap_{X \in X_n^{(k)}} \text{NP}(X)$  and hence  $v \in P_{\text{upper}}$ . Then  $P \subseteq P_{\text{upper}}$ , since  $P_{\text{upper}}$  is also convex.

The vertices due to the all-zeros and all-ones grids (1, 2) are the same for all  $X \in X_n^{(k)}$ ; hence they are in  $P_{\text{upper}}$ . The “extremal” self vertices (that is, the maximal and minimal self grids 3 and 4) are due to the grids in  $X_n^{(k)}$  with maximal and minimal ( $m_n(k)$  and 0, respectively) one-to-one state-to-state transitions. The former grid can be expressed by placing the  $k$  ones starting at  $x_{1,1}$  and proceeding from left to right and top to bottom until all the ones are exhausted, then placing zeros in

the remainder of the locations. The latter grid can be achieved for  $n$  even by placing the  $k$  ones in a subset of the locations that contain ones in the checkerboard grid (or near-checkerboard grid with fewer ones) and placing zeros in the remainder of the locations. The “extremal” anti-self vertices (5 and 6) are due to the ones complement of the extremal self grids.

For  $n$  even, the vertices due to the checkerboards (CB) 7a and 7b can be achieved by the observation equal to the minimal self grid 4 from above and the ones complement of the minimal self grid. For  $n$  odd, the near-checkerboard (NCB) vertices 8b and 10b are achieved with the observation equal to the minimal self grid 4 from above and the ones complement of the minimal self grid. The near-checkerboard vertex 9b is realized by the observation equal to the grid where the  $k$  ones are placed in locations of ones in the near-checkerboard grid with  $\frac{n^2+n}{2}$  ones and the remainder of the locations are filled with zeros. The remaining near-checkerboard vertex 7b is accomplished by taking the ones complement of the previous observation. Hence we have shown that  $P \subseteq P_{\text{upper}}$ .

Now we show that  $P_{\text{upper}} \subseteq P$  by considering the facets of  $P$  and proving that each candidate vertex  $u$  of each Newton polytope  $\text{NP}(X), X \in X_n^{(k)}$  is contained in  $P$ .

*Facet 1.* The first facet is formed by vertices 1, 2, and 3 and is given by  $(n^2 - k)(2k - m_n(k))u_1 + (n^2 - k)(4k -$

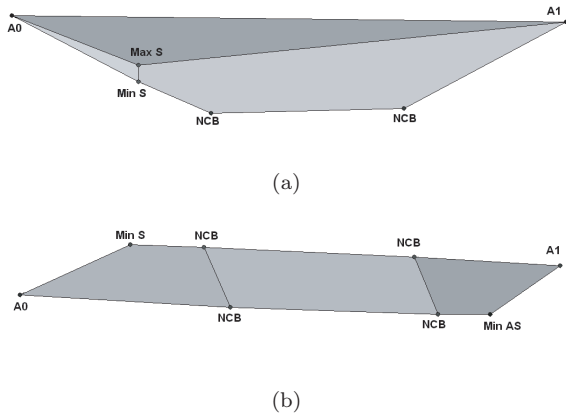


FIGURE 4. Upper-bound polytope  $P_{\text{upper}}$  for  $n = 3$  and  $k = 2$ .(a) Top view; (b) Front view.

$m_n(k)u_2 \geq 2n^2(2k - m_n(k))u_5$ . By setting  $k'$  to be the number of ones in the state grid causing  $u$ , we can rearrange the previous equation to be

$$I_1 = k(n^2 - k)u_2 + n^2(k + 1)u_6 - (n^2 - k')k(k + 1) \geq 0.$$

Since  $u_4 \leq m_n(k')$  for any state grid with  $k'$  ones, we must have  $u_2 \geq 2k' - m_n(k') \geq 2$  for  $0 < k' < n^2$  (we ignore the  $k' = 0, n^2$  cases; the all-zeros and all-ones grids cause vertices of  $P$ ).

We consider two cases:  $k \geq k'$ , which implies  $u_6 \geq k - k'$ ; and  $k' > k$ , which implies  $u_6 \geq 0$ . For the first case,

$$\begin{aligned} I_1 &\geq k(n^2 - k)(2k' - m_n(k')) + n^2(k + 1)(k - k') \\ &\quad - (n^2 - k')k(k + 1) \\ &= k(n^2 - k)(k' + 1) + n^2(k + 1)(k - k') \\ &\quad - (n^2 - k')k(k + 1) \\ &= (n^2 - k)(k - k') \geq 0. \end{aligned}$$

So  $I_1$  holds for  $k \geq k'$ . Next, we consider  $k' > k$ :

$$\begin{aligned} I_1 &\geq k(n^2 - k)(2k' - m_n(k')) - (n^2 - k')k(k + 1) \\ &= \begin{cases} k[(n^2 - k)(k' + 1) - (n^2 - k')(k + 1)] \\ \quad \text{for } k \leq k' \leq n - 1, \\ k[(n^2 - k)n - (n^2 - k')(k + 1)] \\ \quad \text{for } n \leq k' = p'n \leq (n - 1)n, \\ k[(n^2 - k)(n + 1) - (n^2 - k')(k + 1)] \\ \quad \text{for } n < k' = p'n + q' < n^2 - n, q' \neq 0, \\ k[(n^2 - k)(n^2 + 1 - k') - (n^2 - k')(k + 1)] \\ \quad \text{for } n^2 - n + 1 \leq k' = n^2 - q' \leq n^2 - 1, \end{cases} \\ &\geq \begin{cases} k(n^2 + 1)(k' - k) \geq 0, \\ k[(n^2 - k)n - (n^2 - n)(k + 1)] \\ \quad = kn[n^2 - (k + 1)n + 1] > 0, \\ k[(n^2 - k)(n + 1) - (n^2 - (n + 1))(k + 1)] \\ \quad = k[n^3 - kn^2 + n + 1] > 0, \\ k[(n^2 - k)(1 + q') - q'(k + 1)] \\ \quad \geq k[2n^2 - 3k - 1] > 0. \end{cases} \end{aligned}$$

Hence  $I_1$  also holds for  $k' > k$ .

*Facet 2.* The second facet is formed by vertices 1, 2, and 5, and can be given by the following inequality:

$$I_2 = k(k + 1)(n^2 - k') + k(n^2 - k)u_2 - n^2(k + 1)u_6 \geq 0,$$

where again  $k'$  is the number of ones in the state grid that cause  $u$  for some observation  $Y \in X_n^{(k)}$ . Because

$u_6 \leq \min(k, n^2 - k')$ , there are two cases to consider. First, we consider  $k' \leq n^2 - k$ , which implies  $u_6 \leq k$ :

$$\begin{aligned} I_2 &\geq k(k + 1)(n^2 - k') + k(n^2 - k)(2k' - m_n(k')) \\ &\quad - n^2(k + 1)k, \\ \frac{I_2}{k} &\geq (n^2 - k)(2k' - m_n(k')) - k'(k + 1) \\ &= \begin{cases} (n^2 - k)(k' + 1) - k'(k + 1) \\ \quad \text{for } 1 \leq k' \leq n - 1, \\ (n^2 - k)n - k'(k + 1) \\ \quad \text{for } n \leq k' = p'n \leq (n - 1)n \\ (n^2 - k)(n + 1) - k'(k + 1) \\ \quad \text{for } n < k' = p'n + q' < (n - 1)n, q' \neq 0, \\ (n^2 - k)(n^2 - k' + 1) - k'(k + 1) \\ \quad \text{for } n^2 - n - 1 \leq k' = n^2 - q' \leq n^2 - k, \end{cases} \\ &\geq \begin{cases} 2(n^2 - k) - (k + 1) > 0, \\ n[n^2 - kn - (n - 1)] > 0, \\ n^3 - (k + 1)n^2 + n + k + 1 > n + k + 1 > 0, \\ (n^2 - k)(1 + q') - (n^2 - q')(k + 1) \\ \quad = (q' - k)(n^2 + 1) \geq 0. \end{cases} \end{aligned}$$

So  $I_2$  holds for  $k' \leq n^2 - k$ . Next we consider  $n^2 - k < k' = n^2 - q' < n^2$ , so  $u_6 \leq n^2 - k'$ :

$$\begin{aligned} I_2 &\geq k(k + 1)(n^2 - k') + k(n^2 - k)(2k' - m_n(k')) \\ &\quad - n^2(k + 1)(n^2 - k'), \\ &= (n^2 - k)[k(2k' - m_n(k)) - (k + 1)(n^2 - k')], \\ &= (n^2 - k)[k(q' + 1) - (k + 1)q'], \\ &= (n^2 - k)(k - q') > 0. \end{aligned}$$

Furthermore,  $I_2$  holds for  $k' > n^2 - k$ .

*Facet 3:* Another facet of  $P$  is formed by vertices 1, 3, and 4. The inequality corresponding to this facet is described by  $u_7 \geq 0$ , which is true for all candidate vertices  $u$ .

*Facet 4:* Similar to facet 3, this facet is formed by vertices 2, 5, and 6, which has inequality given by  $u_5 \geq 0$ , which again is satisfied by all candidate vertices  $u$ .

*Facet 5:* Vertices 1, 4, 7a, 8a (or 1, 4, 7b, 8b for  $n$  odd) form this facet, which has inequality given by  $u_4 \geq 0$ . All candidate vertices  $u$  satisfy this condition.

*Facet 6:* Similarly, vertices 2, 6, 7a, 8a (or 2, 6, 9b, 10b for  $n$  odd) form this facet, whose inequality is described by  $u_1 \geq 0$ . All  $u$  satisfy this.

*Facet 7:* Vertices 1, 5, 6, 8a (or 1, 5, 6, 7b, 10b for  $n$  odd) form the facet with inequality given by  $u_8 \geq 0$ , which is satisfied by all  $u$ .

*Facet 8:* Similarly, vertices 2, 3, 4, 7a (or 2, 3, 4, 8b, 9b for  $n$  odd) form the facet with inequality  $u_6 \geq 0$ , which is also satisfied by all  $u$ .



*Facet 9* ( $n$  odd): The facet with inequality  $u_2 \leq n^2 - n$  is formed by the near-checkerboard vertices 7b, 8b, 9b, and 10b, which holds for all candidate vertices  $u$ .

Thus we have shown that  $P_{\text{upper}} \subseteq P$ , and hence  $P = P_{\text{upper}}$ .  $\square$

With the above lemma and theorem, we can determine the vertices of the Newton polytope for a few special cases for all  $n$ . The following corollaries list the vertices for the all-zeros observation grid and the vertices for the grid with a single one and the rest zeros, considering  $n$  even and odd separately.

**Corollary 3.3.** *For  $n$  even, the vertices of the Newton polytope for the  $n \times n$  all-zeros observation grid  $Y$  are given by*

$$\begin{aligned} v_1 &= [2n^2, 0, 0, 0, n^2, 0, 0, 0], \\ v_2 &= \left[0, n^2, n^2, 0, \frac{n^2}{2}, 0, \frac{n^2}{2}, 0\right], \\ v_3 &= [0, 0, 0, 2n^2, 0, 0, n^2, 0]. \end{aligned}$$

*Proof:* From Proposition 2.9,  $v_1$  is the all-zeros (and self) vertex,  $v_2$  the checkerboard vertex, and  $v_3$  the all-ones (and anti-self) vertex. Since this polytope is two-dimensional (because  $u_6 = u_8 = 0$  for all candidate vertices), we need only show that any candidate vertex  $u$  satisfies the inequalities given by the three edges:

**Edge 1:** The inequality due to the edge between  $v_1$  and  $v_3$  is given by  $u_2 \geq 0$ , which is satisfied by all  $u$ .

**Edge 2:** The edge between  $v_1$  and  $v_2$  gives the inequality  $u_4 \geq 0$ , which holds for all  $u$ .

**Edge 3:** The inequality from the edge between  $v_2$  and  $v_3$  is  $u_1 \geq 0$ , which also holds for all  $u$ .

Hence the lower-bound polytope  $P_{\text{lower}}$  is equal to  $\text{NP}(Y)$ .  $\square$

**Corollary 3.4.** *For  $n \geq 3$  odd, the vertices of the Newton polytope for the  $n \times n$  all-zeros observation grid  $Y$  are given by*

$$\begin{aligned} v_1 &= [2n^2, 0, 0, 0, n^2, 0, 0, 0], \\ v_2 &= \left[2n, n^2 - n, n^2 - n, 0, \frac{n^2 + n}{2}, 0, \frac{n^2 - n}{2}, 0\right], \\ v_3 &= \left[0, n^2 - n, n^2 - n, 2n, \frac{n^2 - n}{2}, 0, \frac{n^2 + n}{2}, 0\right], \\ v_4 &= [0, 0, 0, 2n^2, 0, 0, n^2, 0]. \end{aligned}$$

*Proof:* By the aforementioned proposition,  $v_1$  is the all-zeros (and self) vertex,  $v_4$  the all-ones (and anti-self) vertex, and  $v_2$  and  $v_3$  are vertices due to the near-checkerboard input grids. Again, as for  $n$  even, this polytope is two-dimensional, and we need only show that any candidate vertex  $u$  satisfies the inequalities given by the four edges:

**Edge 1:** The inequality due to the edge between  $v_1$  and  $v_4$  is  $u_2 \geq 0$ , which is satisfied by all  $u$ .

**Edge 2:** The edge between  $v_1$  and  $v_2$  gives the inequality  $u_4 \geq 0$ , which holds for all  $u$ .

**Edge 3:** The edge between  $v_3$  and  $v_4$  gives the inequality  $u_1 \geq 0$ , which is true for all  $u$ .

**Edge 4:** The inequality for the edge between  $v_2$  and  $v_3$  is  $u_2 \leq n^2 - n$ , which holds for all  $u$  with  $n$  odd.

Hence, as with  $n$  even, the lower-bound polytope  $P_{\text{lower}}$  equals  $\text{NP}(Y)$ .  $\square$

**Corollary 3.5.** *For  $n$  even, the vertices of the Newton polytope for the  $n \times n$  observation grid  $Y$  with a single bit equal to one are given by*

$$\begin{aligned} v_1 &= [2n^2, 0, 0, 0, n^2 - 1, 1, 0, 0], \\ v_2 &= [2n^2 - 4, 2, 2, 0, n^2 - 1, 0, 0, 1], \\ v_3 &= \left[0, n^2, n^2, 0, \frac{n^2}{2} - 1, 1, \frac{n^2}{2}, 0\right], \\ v_4 &= \left[0, n^2, n^2, 0, \frac{n^2}{2}, 0, \frac{n^2}{2} - 1, 1\right], \\ v_5 &= [0, 2, 2, 2n^2 - 4, 0, 1, n^2 - 1, 0], \\ v_6 &= [0, 0, 0, 2n^2, 0, 0, n^2 - 1, 1]. \end{aligned}$$

*Proof:* Notice that  $v_1$  is the all-zeros vertex,  $v_2$  the self vertex,  $v_3$  and  $v_4$  the near-checkerboard vertices,  $v_5$  the anti-self, and  $v_6$  the all-ones vertex; hence the above polytope is  $P_{\text{lower}}$ . For  $k = 1$ , the minimal- and maximal-self (hence also minimal- and maximal-anti-self) vertices become one vertex, since  $m_n(1) = 0$ . Then  $P_{\text{upper}} = P_{\text{lower}}$ , which implies  $\text{NP}(Y) = P_{\text{lower}} = P_{\text{upper}}$ , which has the above vertices.  $\square$

**Corollary 3.6.** *For  $n \geq 3$  odd, the vertices of the Newton polytope for the  $n \times n$  observation grid  $Y$  with a single*

bit equal to one are given by

$$\begin{aligned} v_1 &= [2n^2, 0, 0, 0, n^2 - 1, 1, 0, 0], \\ v_2 &= [2n^2 - 4, 2, 2, 0, n^2 - 1, 0, 0, 1], \\ v_3 &= \left[ 2n, n^2 - n, n^2 - n, 0, \frac{n^2 + n - 2}{2}, 1, \frac{n^2 - n}{2}, 0 \right], \\ v_4 &= \left[ 2n, n^2 - n, n^2 - n, 0, \frac{n^2 + n}{2}, 0, \frac{n^2 - n - 2}{2}, 1 \right], \\ v_5 &= \left[ 0, n^2 - n, n^2 - n, 2n, \frac{n^2 - n - 2}{2}, 1, \frac{n^2 + n}{2}, 0 \right], \\ v_6 &= \left[ 0, n^2 - n, n^2 - n, 2n, \frac{n^2 - n}{2}, 0, \frac{n^2 + n - 2}{2}, 1 \right], \\ v_7 &= [0, 2, 2, 2n^2 - 4, 0, 1, n^2 - 1, 0], \\ v_8 &= [0, 0, 0, 2n^2, 0, 0, n^2 - 1, 1]. \end{aligned}$$

*Proof:* Notice that  $v_1$  is the all-zeros vertex,  $v_2$  the self vertex,  $v_3, v_4, v_5,$  and  $v_6$  the near-checkerboard vertices,  $v_7$  the anti-self, and  $v_8$  the all-ones vertex. As in the previous corollary,  $\text{NP}(Y) = P_{\text{lower}} = P_{\text{upper}}$ , which has the above vertices.  $\square$

#### 4. NEWTON POLYTOPES AND ORBITS

Let  $X \in X_n = \{0, 1\}^{n \times n}$  be a grid. Let  $G = D_{2n} \wr C_2$  be the group of rotations, reflections, and translations acting on  $X_n$ . Let  $\text{orbit}(X) = \{Z \in X_n : g(X) = Z \text{ for some } g \in G\}$ .

The existence of orbits in our model may be problematic, since the two-dimensional images of the characters 6 and 9, or b, d, and p, for example, may lie in the same orbit and, by the following theorem, have the same Newton polytope. We first prove a useful lemma.

**Lemma 4.1.** *Let  $Y = g(X)$  for some  $g \in G$ . Let  $Z$  cause point  $v_Z$  in the Newton polytope of  $X, \text{NP}(X)$ . Then  $W = g(Z)$  causes  $v_Z$  in  $\text{NP}(Y)$ .*

*Proof:* Suppose  $Z$  has  $n_1$  zero-to-zero,  $n_2 = n_3$  zero-to-one (one-to-zero), and  $n_4$  one-to-one state-to-state transitions. Because  $g$  does not change the neighbors of a bit in the grid,  $W$  will have the same number of state-to-state transitions  $n_1, n_2, n_3, n_4$ . Now suppose  $Z$  and  $X$  have  $n_5$  zero-to-zero state-to-observation transitions,  $n_6$  zero-to-one,  $n_7$  one-to-zero, and  $n_8$  one-to-one. Each hidden state bit  $z_{i,j}$  has corresponding observation bit  $x_{i,j}$ . Likewise, each hidden state bit  $w_{k,l} = g(z_{i,j})$  of  $W$  has corresponding observation bit  $y_{k,l} = g(x_{i,j})$  of  $Y$ . Hence the number of zero-to-zero state-to-observation transitions  $n_5$  has not changed. Likewise,  $n_6, n_7, n_8$  have not

changed. So  $Z$  causes  $v_Z = (n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8)$  in  $\text{NP}(X)$ , which is also caused by  $W = g(Z)$  in  $\text{NP}(Y)$ , where  $Y = g(X)$ .  $\square$

**Theorem 4.2.** *The Newton polytope of an observation grid  $X \in X_n$  is the same as for all  $Y \in \text{orbit}(X)$ .*

*Proof:* Let  $g(X) = Y$  for some  $g \in G$ . Suppose the vertices  $v_1, v_2, \dots, v_M$  of  $\text{NP}(X)$  are caused by hidden state grids  $Z_1, Z_2, \dots, Z_M$ , where  $M$  is the number of vertices of  $\text{NP}(X)$ . Notice that the hidden state grids  $g(Z_1), g(Z_2), \dots, g(Z_M)$  will cause the same (candidate) vertices of  $\text{NP}(Y)$  by the above lemma.

It now remains to show that no other hidden state grid can cause any additional vertices or cancel any previous vertices of  $\text{NP}(Y)$ . Let  $V = \{v_1, v_2, \dots, v_M\}$ . Suppose  $W$  can cause a new vertex  $v_W \notin V$  of  $\text{NP}(Y)$ . However, by the above lemma,  $g^{-1}(W)$  will cause the same point  $v_W$  in  $\text{NP}(X)$ , which has vertex set  $V$ . Since  $v_W$  is not a vertex of  $\text{NP}(X)$ , it cannot be a vertex of  $\text{NP}(Y)$ .  $\square$

We now consider the number of orbits to determine how much calculation is necessary to determine every possible Newton polytope for a given  $n$ .

**Theorem 4.3.** *The number of distinct orbits in  $X_n$  when  $n$  is an odd prime is*

$$\frac{1}{8n^2} \left[ 2^{n^2} + (n^2 - 1)2^n + 4n2^{(n^2+n)/2} + 4(n^2 - n)2^{(n+1)/2} + n^22^{(n^2+1)/2} + 2n^22^{(n^2+3)/4} \right].$$

**Theorem 4.4.** *The number of distinct orbits in  $X_n$  when  $n = 2^m$ , where  $m$  is a positive integer, is*

$$\begin{aligned} &\frac{1}{8 \cdot 2^{2m}} \left[ 2^{2^{2m}} + 3 \sum_{i=1}^m \left( 2^{2^{i-1} + 2^{2m-i}} \right) \right. \\ &\quad + 2^{m+2^{m+1}+2^{2m-1}-2^m} + 3 \cdot 2^{m+2^{2m-1}} \\ &\quad + 2 \sum_{i=1}^{m-1} \left( 2^{2^{m-i} + 2^{m-1+i}} \right) + 2^{2m+2^{2m-1}} \\ &\quad + 3 \cdot 2^{2m-2+2^{2m-1}} + 2^{m+1+2^m+2^{2m-1}-2^{m-1}} \\ &\quad + 2 \sum_{i=1}^m \left( 2^{m-1+i+2^{2m-i-1}} \right) \\ &\quad \left. + 2^{2m+2+2^{2m-2}} + 2^{2m+2^{2m-2}} \right]. \end{aligned}$$

*Sketch of proof:* The idea for this proof comes from [Biggs 89, Sections 20.1 and 20.4]. The number of distinct orbits



of  $X_n$  is given by

$$\frac{1}{|G|} \sum_{g \in G} \zeta_g(x_1, x_2, \dots, x_{n^2}) \Big|_{x_1=2, x_2=2, \dots, x_{n^2}=2},$$

where  $\zeta_g(x_1, x_2, \dots, x_{n^2}) = \prod_{i=1}^{n^2} x_i^{\alpha_i}$  satisfies the following conditions:

1.  $\alpha_i$  is the number of  $i$ -cycles in  $g$ .
2.  $\sum_{i=1}^{n^2} \alpha_i = n^2$ .

This completes the proof sketch. □

**Corollary 4.5.** *The number of orbits for  $X_2$  is 6; for  $X_3$ , it is 26; for  $X_4$ , 805; and for  $X_5$ , 172 112.*

**Remark 4.6.** For  $n \leq 4$ , since the number of hidden state grids ( $2^{n^2}$ ) and the number of orbits are reasonably small, we have calculated the Newton polytope for every observation grid in  $X_n$ ,  $n = 2, 3, 4$ . The simple algorithm we use is to calculate every possible candidate vertex by considering every hidden state grid and then to perform a convex-hull operation to determine the Newton polytope.

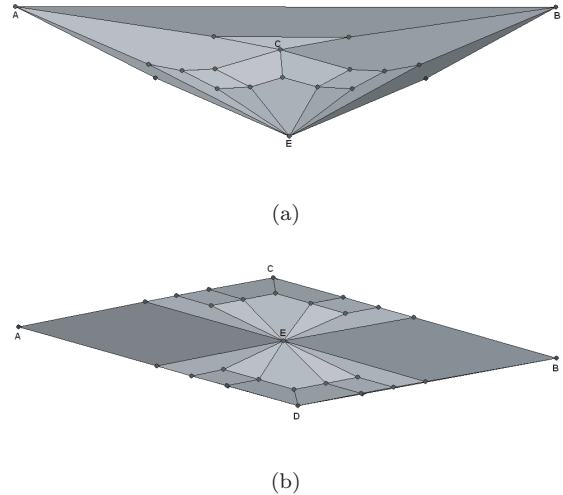
### 5. NUMBER OF VERTICES OF NEWTON POLYTOPES

A major concern about Newton polytopes is how many vertices they have as the dimension  $n$  of the underlying grid increases and how to compute those vertices as  $n$  increases.

**Conjecture 5.1.** *The maximum number of vertices of a Newton polytope  $NP(X)$  for  $X \in X_n$  is of order  $n^2$ .*

**Remark 5.2.** An exhaustive search using the computer software packages MATLAB and LRS over all the orbits of  $X_2$ ,  $X_3$ , and  $X_4$  yields 9, 20, and 31, respectively, as the maximum number of vertices for a Newton polytope, suggesting that the upper bound should be approximately  $2n^2$ .

An example of a Newton polytope for a  $4 \times 4$  observation grid with the maximal number (31) of vertices is plotted using `polymake` in Figure 5. An interesting observation is that this same polytope arises for all three orbits in  $X_4$ , whose polytope has 31 vertices. For viewing purposes, we consider the line segment from the all-zeros vertex (**A**) to the all-ones vertex (**B**) to be the “spine.” We consider the checkerboard vertex (**E**) to be the “nose.” (If  $n$  is odd, there will be vertices due to



**FIGURE 5.** Newton polytope for a  $4 \times 4$  observation with 31 vertices. (a) Top view; (b) Front View.

the near-checkerboard grids, which form the “snubbed nose.”) The self (**C**) and anti-self (**D**) vertices will be considered the top and bottom vertices, respectively, of the Newton polytope.

Using suffixes of candidate vertices, we show that the upper bound on the number of vertices is no greater than  $O(n^4)$ .

**Definition 5.3.** A candidate vertex

$$u = (u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8)$$

has *prefix*  $(u_1, u_2, u_3, u_4)$ , denoting the number of zero-to-zero, zero-to-one, one-to-zero, and one-to-one state-to-state transitions, respectively, and *suffix*  $(u_5, u_6, u_7, u_8)$ , denoting the number of zero-to-zero, zero-to-one, one-to-zero, and one-to-one state-to-observation transitions, respectively.

The entries of the suffix of  $u$  satisfy  $u_5 + u_7 = n^2 - k$  and  $u_6 + u_8 = k$ , where  $k$  is the number of ones in the observation grid  $Y$ . Suppose a hidden state grid  $X$  has  $i$  ones. Then also  $u_5 + u_6 = n^2 - i$  and  $u_7 + u_8 = i$ . Without loss of generality, assume  $k, i \leq \frac{n^2}{2}$ . Then the number of possible suffixes for  $X$  and  $Y$  is  $\min(k, i) + 1$ . Considering all possible hidden state grids  $X \in X_n$ , there are no more than

$$2 \sum_{i=0}^{\lfloor \frac{n^2}{2} \rfloor} [\min(k, i) + 1]$$

possible distinct suffixes for  $Y$  in  $NP(Y)$ .

**Lemma 5.4.** *The Newton polytope of  $Y$  can have at most two vertices with the same suffix  $s$ .*

*Proof:* Suppose  $v_1 = [2n^2 - 4i + \nu_1, 2i - \nu_1, 2i - \nu_1, \nu_1, s]$ ,  $v_2 = [2n^2 - 4i + \nu_2, 2i - \nu_2, 2i - \nu_2, \nu_2, s]$ , and  $v_3 = [2n^2 - 4i + \nu_3, 2i - \nu_3, 2i - \nu_3, \nu_3, s]$  are vertices of  $\text{NP}(Y)$ , where  $\nu_1 < \nu_2 < \nu_3$ . Now let  $1 > \alpha = \frac{\nu_3 - \nu_2}{\nu_3 - \nu_1} > 0$ , and form  $v = \alpha v_1 + (1 - \alpha)v_3 = v_2$ . So  $v_2$  is a convex combination of  $v_1$  and  $v_3$ ; hence it cannot be a vertex of  $\text{NP}(Y)$ .  $\square$

The maximum number of possible suffixes for any observation grid  $Y \in X_n$  occurs when  $k = \lfloor \frac{n^2}{2} \rfloor$ , which gives  $(\lfloor \frac{n^2}{2} \rfloor + 1)(\lfloor \frac{n^2}{2} \rfloor + 2)$  possible suffixes. From this we can see that the maximum number of vertices in  $\text{NP}(Y)$  can be no more than  $2(\lfloor \frac{n^2}{2} \rfloor + 1)(\lfloor \frac{n^2}{2} \rfloor + 2)$ , which is of order  $n^4$ .

Suffixes have proved useful in computing the Newton polytopes of observations with  $n > 4$ . Since the number of possible hidden state grids  $2^{n^2}$  is at least  $2^{25} = 33\,554\,432$ , it is computationally expensive to consider every hidden state grid in order to find approximately  $2n^2$  vertices, especially since we know some vertices that arise every time. Instead of considering every hidden state grid, we propose the following method for computing Newton polytopes:

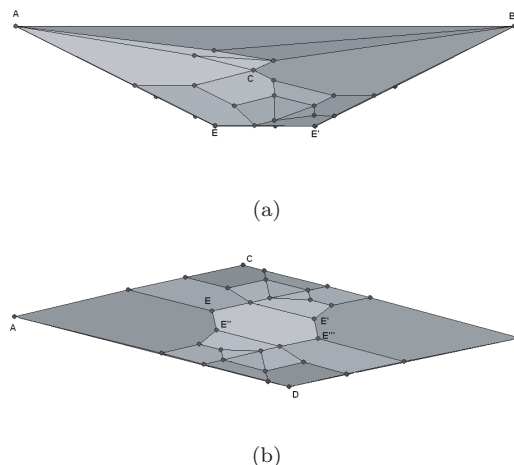
1. Compute the “essential” vertices of the Newton polytope.
2. Form a list of possible suffixes, whose length is of order  $n^4$ .
3. For each suffix, determine which prefixes could not cause vertices of the polytope by using convex-hull methods, and discard them.
4. If any prefixes remain from step 3, determine whether any are possible.

In determining whether prefixes are possible given a suffix, we invoke a combinatorial search, which, for large enough values of  $u_6$  and  $u_7$ , may unfortunately slow down computation considerably.

**Example 5.5.** For the  $5 \times 5$  observation grid

$$Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix},$$

we can use the suffix method to calculate the Newton polytope, which has 38 vertices and is plotted in Figure 6.



**FIGURE 6.** Newton polytope for a  $5 \times 5$  observation with 38 vertices. (a) Top view; (b) Front view.

In attempting to prove that the bound on the maximal number of vertices for the Newton polytope is of order  $n^2$ , we find that a bound closer to the desired bound results from a theorem of Andrews [Andrews 63], via [Pachter and Sturmfels 04, Theorem 7], which states that

$$\# \text{ vertices}(\text{NP}(x)) \leq \text{constant} \cdot E^{d(d-1)/(d+1)},$$

where  $E = 3n^2$  is the number of edges in the graphical model and  $d$  is the number of parameters in the model. Because we have shown that the effective dimension of the Newton polytope is three, the bound on the number of vertices is  $c(3n^2)^{3/2}$ , which is of order  $n^3$ .

## 6. CHARACTER-RECOGNITION ALGORITHM

We use an approach similar to that in [Meriardo et al. 00], wherein we train the 2D HMM by determining the optimal set of parameters  $\lambda^{(k)}$  for each training image  $O^{(k)}$ ,  $k \in A$ ,  $A$  a collection of characters (numbers, letters, symbols). After tropicalizing the probability map  $P(O^{(k)} | \lambda)$ , we find that the optimal parameters  $\gamma^{(k)} = -\log(\lambda^{(k)})$  satisfy

$$\gamma^{(k)} = \arg \min_{\gamma_v \in \Gamma_V} \langle v, \gamma_v \rangle,$$

where  $\gamma_v$  minimizes the tropicalized probability map in the normal cone of  $v$ ,  $\Gamma_V$  is the set of  $\gamma_v$  for each  $v \in V$ , and  $V$  is the vertex set of  $\text{NP}(O^{(k)})$ . Hence the number of computations in training the 2D HMM is of order  $n^2$ , according to our conjecture.

Similarly, for each testing image  $I$ , we search for the optimal index  $k^* \in A$ , where

$$k^* = \arg \min_{k \in A} \left\{ \min_{v \in V} \langle v, \gamma^{(k)} \rangle \right\},$$

where  $V$  is the vertex set of  $\text{NP}(I)$ ; hence the testing algorithm should also have  $O(n^2)$  computations for each image  $I$ .

## REFERENCES

- [Andrews 63] George E. Andrews. “A Lower Bound for the Volume of Strictly Convex Bodies with Many Boundary Lattice Points.” *Transactions of the American Mathematical Society* 106:2 (1963), 270–279.
- [Biggs 89] Norman L. Biggs. *Discrete Mathematics*. Oxford: Oxford University Press, 1989.
- [Merialdo et al. 00] Bernard Merialdo, Stéphane Marchand-Maillet, and Benoit Huet. “Approximate Viterbi Decoding for 2D-Hidden Markov Models.” *IEEE International Conference on Acoustics, Speech, and Signal Processing 2000* 6 (2000), 2147–2150.
- [Murphy 98] Kevin Murphy. “A Brief Introduction to Graphical Models and Bayesian Networks” Available online (<http://www.cs.ubc.ca/~murphyk/Bayes/bnintro.html>), 1998.
- [Pachter and Sturmfels 04] Lior Pachter and Bernd Sturmfels. “Tropical Geometry of Statistical Models.” *Proceedings of the National Academy of Sciences* 101 (2004), 16132–16137.

Jay Wierer, Department of Electrical and Computer Engineering, University of Wisconsin, Madison, Madison, WI 53706 (jdwierer@wisc.edu)

Nigel Boston, Department of Mathematics, University of Wisconsin, Madison, Madison, WI 53706 (boston@math.wisc.edu)

Received April 21, 2006; accepted September 9, 2006.