The Geometry of Niggli Reduction: BGAOL, Embedding Niggli Reduction and Analysis of Boundaries

Lawrence C. Andrews[a]* and Herbert J. Bernstein[b]

[a]*Micro Encoder Inc., 11533 NE 118th St., Kirkland, WA 98034 USA, and* [b]*Dowling College, 1300 William Floyd Pkwy, Shirley, NY 11967 USA.*

E-mail: andrewsl@ix.netcom.com

1. Supplementary Materials – The Monte Carlo Search for Boundary Polytopes

One can follow an entirely algebraic process to identify all boundary polytopes. In view of the complexity of the space it is helpful to use a computer to explore $\boldsymbol{N}$ especially to confirm information on the dimensionality of boundary polytopes and on the necessary boundary transformations.

In order to reduce the computational burden and deal with the open boundary polytopes, rather than randomly generating points in $\boldsymbol{G^6}$, we generate random short lines in $\boldsymbol{G^6}$ and look for cases in which the boundary of $\boldsymbol{N}$ must have been crossed because one end of the line is Niggli-reduced, while the other end is not. The process for the initial search for 5-D boundary polytopes is given in Table 1.

The search for the boundary polytopes resulting from the process in Table 1 produces transformations in the course of Niggli reduction (step 6). We sort the Niggli-reduced $\boldsymbol{G^6}$ vectors by the associated transformation matrix from step 6. A high population for a given matrix indicates a significant volume of $\boldsymbol{G^6}$ with access to the

associated boundary polytope, implying a 5-D boundary polytope (See Fig. 1 and Fig. 2). A lower population for a given matrix possibly implies a lower-dimensional boundary polytope, a common "edge" resulting from the intersection of multiple 5-D boundary polytopes. With enough probes, many of the lower dimensional boundary polytopes can be discovered, but only with difficulty because the information on those lower-dimensional boundary polytopes is swamped in a sea of data about the 5-D boundary polytopes.

Efficient discovery of all these intersections of multiple 5-D boundary polytopes requires the use of "projectors". A projector for a subspace $\boldsymbol{Q}$ in $\boldsymbol{G^6}$ is a symmetric matrix, P, that when applied to any point $\mathbf{g}$ in $\boldsymbol{G^6}$ is such that $P\mathbf{g}$ is the closest point in $\boldsymbol{Q}$ to $\mathbf{g}$. A projector has the nice property that $PP = P$, i.e. that it acts like the identity matrix on the subspace $\boldsymbol{Q}$.

The projector for a given boundary polytope can be discovered by examining the set of Niggli-reduced $\boldsymbol{G^6}$ vectors associated with the transform matrix for that boundary polytope. The examination can either be a simple inspection of the list of vectors or can be done algorithmically by use of a singular value decomposition (SVD) calculation (Beltrami, 1873; Jordan, 1874; Stewart, 1993). It should be noted that the projector projects onto the hyperplane associated with the polytope and may project a point from $\boldsymbol{G^6}$, or even from $\boldsymbol{N}$ itself, outside of $\boldsymbol{N}$.

Let us first derive projectors by simple inspection. A search for points near a boundary produces a list of vectors that can be examined for the conditions that are met. For instance, the trial vectors:

4.41605	**53.21164**	**53.3171**	-9.85206	-2.73956	-1.78806
4.95245	**106.2402**	**106.5968**	-72.3608	-0.26549	-4.79911
5.62821	**98.26772**	**98.36612**	24.37056	1.57819	1.85157

are seen to meet the condition $g_2 = g_3$.

The trial vectors:

4.85822	9.79018	40.14963	-3.6758	**-0.01092**	-2.18456
4.89205	21.22063	75.92303	6.01777	**0.05752**	2.78514
5.03365	26.32789	46.84058	-17.2646	**-0.01252**	-2.03757

are seen to meet the condition $g_5 = 0$.

In each case the necessary projector is obvious (see projectors P_2 and P_3, below). In other cases, such as the body-diagonal boundary polytope, or many of the lower-dimensional boundary polytopes involving multiple face-diagonals, simple inspection is challenging. Given a sufficient number of vectors, the projector may be recovered algorithmically from such a list of vectors, rather than by inspection. If a singular value decomposition is done on a list of vectors near to and covering a single 5-D boundary polytope, the vector corresponding to the smallest singular value is the unit normal to that boundary polytope, and the projector onto that polytope is simply the identity minus the projector onto the line of the unit normal. The projector onto the line of the unit normal is generated by forming a matrix, A, whose first row is the unit normal and whose remaining rows are set to zero. Then A^TA is the projector onto the line of that unit normal.

In the general case of a lower-dimensional boundary polytope, the projector onto the boundary polytope is the identity minus the projector Q onto the hyperplane $\boldsymbol{\Omega}$ orthogonal to the boundary polytope. In that case a singular value decomposition on the list of vectors near to and covering the lower-dimensional boundary polytope will have small singular values for the vectors spanning $\boldsymbol{\Omega}$. The projector Q onto $\boldsymbol{\Omega}$ is generated by forming a matrix A with initial rows consisting of the vectors corresponding to the small singular values and the remaining rows set to zero. Then A^TA is the projector onto $\boldsymbol{\Omega}$ and $I - A^TA$ is the projector onto the boundary polytope.

Once we have found the projectors for the 5-D boundary polytopes, the projectors for their intersections may be generated from their products. The product of two different projectors is not, itself, likely to be a projector, but repeated squaring of that

product rapidly converges to the correct projector (Andrews, 1976). Given these projectors, the process for the remaining boundary polytopes is given in Table 2. This process is similar to the initial process, but three new steps have been added: In step 0 the projector for a particular boundary polytope is read in. In step 2A the vector confirmed by step 2 is projected onto the hyperplane containing the boundary polytope, and step 2B verifies that the projected vector is still a valid unit cell. When doing a random perturbation of a vector already projected onto the hyperplane containing a boundary polytope while looking for intersections with that boundary polytope, it is sufficient to use a random vector projected onto that the hyperplane containing the boundary polytope. However, that risks missing some very low dimension cases. That search produced 215 distinct boundary polytopes. The search was then redone, stepping back into the Niggli cone after each projection, and using a full 6-D spherical random search to ensure catching any nearby boundary polytopes. That search added one more distinct boundary polytope, for a total of 216 boundary polytopes.

The dimension of a boundary polytope can be determined by computing the number of eigenvalues equal to 1 of the projector onto the hyperplane containing the boundary polytope, and having distinct boundary projectors is a sufficient condition for two boundary polytopes to be distinct. However, distinct projectors is a necessary, but not sufficient, condition for two boundary polytopes to be distinct, because crossing one bounding hyperplane in two different places can require two different transformation matrices to reduce the result. The reduction transformation matrices themselves are needed to disambiguate cases with the identical bounding hyperplane projector.

In the course of the Monte Carlo investigation of the boundary polytopes resulting from combinations of 5-D boundary polytopes, we track the relative populations as an indicator of a consistent assignment of dimension. The candidates for assignment to a particular dimension are sorted by population, and the list is cut off on a precipitous

drop of more than an order of magnitude. For the remaining populations, the mean population μ and estimated standard deviation σ are computed. For each population τ, a Z-score, $(\tau - \mu)/\sigma$, is computed (see http://en.wikipedia.org/wiki/Standard_score). Z-scores of less than -1 suggest the possible need for further study of the boundary polytopes involved. A combination of the population analysis resulting from the Monte Carlo search, algebraic analysis of constraints and computation of the eigenvalues of projectors allows the identification of the dimensions of the boundary polytopes, helps to identify distinct combinations of intersections of 5-D boundary polytopes that represent the same lower dimensional boundary polytopes and confirms the critical initial identification of the 15 5-D boundary polytopes identified purely by the Monte Carlo search.

We limit our consideration of valid boundary polytopes to those avoiding the mathematically interesting but crystallographically impossible cases of zero length cell edges. Combinations of boundary polytopes without a valid intersection or with an intersection that would force any of $g_{\{1,2,3\}}$ to zero or that did not have neighboring Niggli-reduced probe points are eliminated. 574 combinations of 1 through 8 intersecting 5-D boundary polytopes were not degenerate. Many combinations represent equivalent boundary polytopes. There are 216 distinct boundary polytopes. There are 15 5-D boundary polytopes of the full $\boldsymbol{G^6}$ Niggli cone, 53 4-D boundary polytopes resulting from intersections of pairs of the 15 5-D boundary polytopes, 79 3-D boundary polytopes resulting from the 2-fold and higher intersections of the 15 5-D boundary polytopes, 55 2-D boundary polytopes resulting from 2-fold and higher intersections of the 15 5-D boundary polytopes, 14 1-D boundary polytopes resulting from 3-fold and higher intersections of the 15 5-D boundary polytopes. The ability of the intersection of only 2 5-D boundary polytopes to produce 4-D, 3-D or 2-D boundary polytopes results from the additional constraints imposed by Niggli reduction.

For the 5-D boundary polytopes, three special cases arise in which the Niggli reduction conditions divide a single polytope into two sub-polytopes, the "flat boundary intersections". In each of those three cases, instead of two boundary polytopes meeting at an angle, two boundary polytopes meet edge on. These three cases are $g_4 = g_2$, $g_5 = g_1$, and $g_6 = g_1$. In each case the division in the polytope is based on the equality of the other two members of $g_{\{4,5,6\}}$. For example, when $g_4 = g_2$, the division is along $g_5 = g_6$. Therefore, when multiplying projectors, instead of multiplying the projector for $g_4 = g_2$ by itself when the two half-polytopes are required, the second projector is replaced by the projector for the division, *i.e.* in the example given instead of multiplying the projector for $g_4 = g_2$ by itself, it is multiplied by the projector for $g_5 = g_6$.

Table 1. *Initial process to locate 5-D boundary polytopes*

1. Generate a random vector in $\boldsymbol{G^6}$.
2. Confirm that the vector represents a proper unit cell (for instance, the sum of the interaxial angles must be less than 360 degrees).
3. Niggli-reduce the vector.
4. Randomly perturb the vector resulting from step 3.
5. Confirm that the perturbed vector represents a proper unit cell.
6. Niggli-reduce the vector from step 5, accumulating the total transformation matrix from the vector of step 3 to the new reduced vector.
7. If the transformation is the unit matrix, then the perturbed vector was not near the boundary of $\boldsymbol{N}$. Discard the trial. Otherwise, proceed.
8. If the transformation matrix has been discovered before, increment a counter for its occurrences. Otherwise, add it to the list of discovered transformations.
9. Repeat many times, starting at step 1.

Table 2. *Process to locate lower-dimensional boundary polytopes.*

0. **Read in the projector for the desired boundary to be probed.**
1. Generate a random vector in $\boldsymbol{G^6}$.
2. Confirm that the vector represents a proper unit cell (for instance, the sum of the interaxial angles must be less than 360 degrees).

2A. **Project the vector confirmed in step 2 onto the boundary.**

2B. **Confirm that the projected vector is still a valid unit cell.**

3. Niggli-reduce the vector.
4. Randomly perturb the vector resulting from step 3.
5. Confirm that the perturbed vector represents a proper unit cell.
6. Niggli-reduce the vector from step 5, accumulating the total transformation matrix from the vector of step 3 to the new reduced vector.
7. If the transformation is the unit matrix, then the perturbed vector was not near the boundary of $\boldsymbol{N}$. Discard the trial. Otherwise, proceed.
8. If the transformation matrix has been discovered before, increment a counter for its occurrences. Otherwise, add it to the list of discovered transformations.
9. Repeat many times, starting at step 1.

Table 3. *The 53 4-D boundary polytopes. The boundary conditions use the symbol notation of Table 1 in the main text for the 5-D boundary polytopes. The bounding 4-D polytopes of each of the 15 5-D boundary polytopes can be read off from this table by going down the appropriate column to the main diagonal and then across. For example, the edges of the 3 polytope are 13, 23, 34, 35, 3A, 3B, 3D and 3E.*

$g_2 = g_3$	$g_4 = 0$	$g_5 = 0$	$g_6 = 0$	a-face-diagonals			b-face-diagonals			c-face-diagonals			body
12	13	14	15	16	17	18	19	1A	1B	1C	1D	1E	1F
	23	24	25	26	27	28	29	2A	2B	2C	2D	2E	2F
		34	35					3A	3B		3D	3E	
			45		47	48				4C		4E	
				56		58	59		5B				
					67		69						
										7C			
													8F
								9A		9C			
											AD		
												BE	BF
											CD		
													EF

Table 4. *The 79 3-D boundary polytopes. In some cases the most natural presentation of a given 3-D polytope is a 4-fold intersection. In each of those cases an equivalent 3-fold intersection is given in parentheses immediately below the 4-fold.*

$g_4=0$	$g_5=0$	$g_6=0$	a-face-diagonals			b-face-diagonals			c-face-diagonals			body
123	124	125	126	127	128	129	12A	12B	12C	12D	12E	12F
	134	135					13A	13B		13D	13E	
		145		147	148				14C		14E	
			156		158	159		15B				
				167		169						
									17C			
												18F
							19A					
										1AD		
												1BF
										1CD		
												1EF
	234	235					23A	23B		23D	23E	
		245		247	248				24C		24E	
			256		258	259		25B				
				267		269						
									27C			
												28F
							29A		29C			
										2AD		
											2BE	2BF
										2CD		
												2EF
		345							34CD		34E	
									(34C)			
						359A		35B				
						(359)						
										3AD		
											3BE	
				4567								
				(456)								
					458							
									47C		48EF	
											(48E)	
						569						
								58BF				
								(58B)				
						679C						
						(679)						
									9ACD			
									(9AC)			
												BEF

Table 5. *The 55 2-D boundary polytopes*

$g_5=0$	$g_6=0$	a-face-diagonals	b-face-diagonals	c-face-diagonals	body
1234	1235 1245	1247 1248 1256 1258 1267	123A 123B 1259 125B 1269 129A	123D 123E 124C 124E 127C 12AD 12CD	128F 12BF 12EF
	1345		134CD (134C) 1359A 135B (1359)	134E 13AD 13BEF (13BE)	
		14567 1458 (1456)		147C 148EF (148E)	
			1569 158BF (158B)		
				1679ACD (6D, 7A)	
	2345		2359A 235B (2359)	234CD 234E (234C) 23AD 23BE	
		24567 2458 (2456)		247C 248EF (248E)	
			2569 258BF (258B)		
			2679C (2679)		
			29ACD (29AC)		
					2BEF

Fig. 1. Illustration of Monte Carlo search for boundary polytopes in $\boldsymbol{G^6}$. The hatched line represents a boundary polytope, *e.g.* $g_2 = g_4$. The unhatched side of the boundary (***N***) consists of Niggli-reduced cells. In this case the hatched side is divided into two regions, ***NN1*** (*e.g.* $g_5 \geqslant g_6$) and ***NN2*** (*e.g.* $g_5 < g_6$) for which different reduction matrices are required to get a Niggli-reduced cell, *e.g.* M_6 for ***NN1*** and M_7 for ***NN2***. "***NN***" stands for not Niggli-reduced. $V1$ is a randomly generated probe point in region ***N*** for which a random short line reaches across the boundary to reach region ***NN1***, so the starting point is associated with M_6. $V2$, $V4$ and $V5$ are randomly generated probe points for which the random short line remains in the Niggli-reduced region ***N***. Therefore $V2$, $V4$ and $V5$ are discarded. $V3$ is a randomly generated probe point in region ***N*** for which a random short line reaches across the boundary to reach region ***NN2***, so the starting point is associated with M_7. Because of this difference in reduction matrices, the boundary polytope is treated as consisting of two distinct boundary polytopes, in this instance cases 6 and 7.

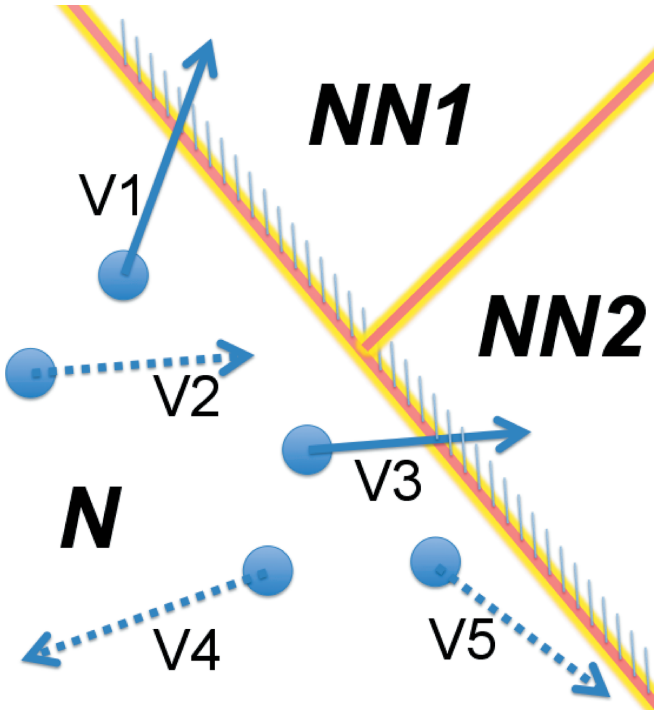

Fig. 2. Counts of points found near various boundary polytopes in 100 million trials, organized in declining order of counts, showing the most populated 23 of the 92 boundary polytopes found in this run. This is a run with no filtering for any particular boundary with the counts shown on a logarithmic scale. Note the precipitous drop of nearly 2 orders of magnitude after the first 15 boundary polytopes. This drop confirms that those 15 boundary polytopes are the 5-D boundary polytopes and that there is a vanishingly small probability of there being any other 5-D boundary polytopes of the Niggli-reduced cells.