

# Contribution Title<sup>\*</sup>

First Author<sup>1</sup>[0000–1111–2222–3333], Second Author<sup>2,3</sup>[1111–2222–3333–4444], and  
Third Author<sup>3</sup>[2222–3333–4444–5555]

<sup>1</sup> Princeton University, Princeton NJ 08544, USA

<sup>2</sup> Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany  
lncs@springer.com

<http://www.springer.com/gp/computer-science/lncs>

<sup>3</sup> ABC Institute, Rupert-Karls-University Heidelberg, Heidelberg, Germany  
{abc,lncs}@uni-heidelberg.de

**Abstract.** The abstract should briefly summarize the contents of the paper in 15–250 words.

**Keywords:** First keyword · Second keyword · Another keyword.

## 1 First Section

### 1.1 A Subsection Sample

Please note that the first paragraph of a section or subsection is not indented. The first paragraph that follows a table, figure, equation etc. does not need an indent, either.

Subsequent paragraphs, however, are indented.

**Sample Heading (Third Level)** Only two levels of headings should be numbered. Lower level headings remain unnumbered; they are formatted as run-in headings.

*Sample Heading (Fourth Level)* The contribution should contain no more than four levels of headings. Table ?? gives a summary of all heading levels.

## 2 Group decomposition based on 4D approximation

We select the better group among the points through the approximation hypervolume of each pair of points in the same layer. Exactly, we will select the pair of points which has the abstract zone with smaller hypervolume. Because the value of the hypervolume exactly is a interval, so the smaller hypervolume of a pair of points means the over-approximation of the hypervolume is smaller than under-approximation of others or the hypervolume has a smaller over-approximation

---

<sup>\*</sup> Supported by organization x.

than others which has overlapping with the interval of the hypervolume. At first, we assume that all of the hypervolume and volume mentioned following are convex and all of the situations discussed following under the  $k = 2$ .

We will introduce how to expand convex volume calculation formula from three dimension to four dimension, how to obtain the under- and over- approximation of four-dimensional hypervolume in order.

## 2.1 Expand convex volume calculation formula to four dimension

**Theorem 1.** *Making a three-dimensional prismatic cone as the base volume which volume is  $V$ , like the base area of a three-dimensional prismatic cone, integrating the volume of the three-dimensional prismatic cone along the fourth dimension, we will get a hypervolume of four-dimensional prismatic cone  $V^{(4)}$ , which equals to*

$$V^{(4)} = \frac{1}{4} V^{(3)} h^{(4)} \quad (1)$$

In equation 1,  $h^{(4)}$  is the value of the fourth dimension, like the 'height' in three-dimensional situation and  $V^{(3)}$  is the volume as the base of the four-dimensional prismatic cone. Before deriving 1, we will first review the derivation of volume for a three-dimensional prismatic cone.

*Proof.* The main idea is to consider a three-dimensional prismatic cone as a pile of cubes [1]. For example, in Fig. 1, ABCD is a convex polygon on the yOz plane and its area is  $S$ . The three-dimensional prismatic cone is consisted of several cubes with base areas similar to the base area ABCD. That means if we use a plane paralleling to ABCD to intercept the prismatic cone J-ABCD, we will get a intersection area  $A'B'C'D'$  which is similar to ABCD. As a result, there exists a similarity between the areas of  $A'B'C'D'$  (the area is defined as  $S'$ ) and  $S$ :

$$S' = \frac{S \cdot x^2}{h^2} \quad (2)$$

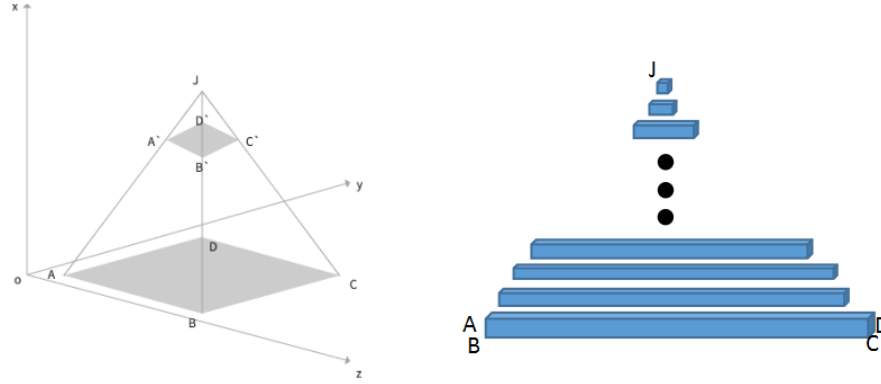
$x$  is the distance between  $A'B'C'D'$  and ABCD,  $h$  is the whole height of the three-dimensional prismatic cone J-ABCD.

For every cube as component of the J-ABCD, its volume defined as  $V_i$ ,  $i=1,2,\dots,n$ . According to the similarity theorem:

$$V_i = \lim_{n \rightarrow \infty} \frac{h}{n} \cdot \frac{S \cdot x_i^2}{h^2} \quad (3)$$

As mentioned before, a three-dimensional prismatic cone can be considered as a pile of cubes, so the whole volume of J-ABCD can be expressed as:

$$V = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{h}{n} \cdot \frac{S \cdot x_i^2}{h^2} \quad (4)$$



**Fig. 1.** The right picture of Fig. 1 means if we use a plane paralleling to ABCD to intercept the prismatic cone J-ABCD, we will get a intersection area  $A'B'C'D'$  which is similar to ABCD. The right one means that the three-dimensional prismatic cone is consisted of several cubes with base areas similar to the base area ABCD.

For  $\frac{h}{n} \in (0, 1]$  and  $x_i \in [0, h]$ , equation 4 and the conception of definite integration can be expressed as:

$$V = \int_0^h \frac{S \cdot x^2}{h^2} dx = \frac{1}{3}Sh \quad (5)$$

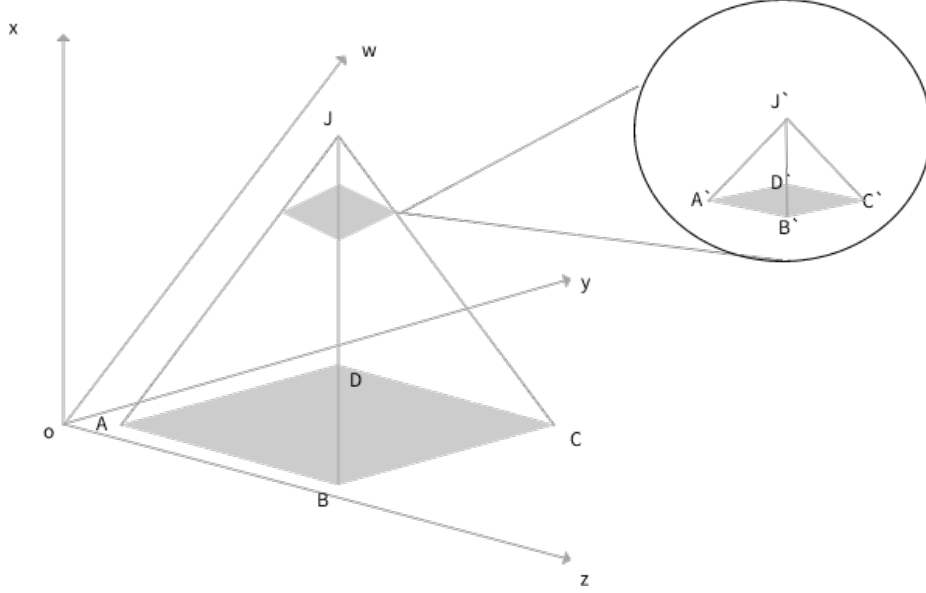
equation 5 is the general volume formula of a three-dimensional prismatic cone. Now we have reviewed the expansion from two-dimensional area to three-dimensional volume, it's time to expand from three-dimensional volume to four-dimensional prismatic cone. Comparing to the situation from two-dimensional area to three-dimensional volume, a four-dimensional prismatic cone as a pile of hypercubes (assume that there are  $n$  such hypercubes, marked as  $H_i$ ,  $i=1, 2, \dots, n$ ), and if we use a three-dimensional space paralleling to the base of the four-dimensional prismatic cone to intercept the four-dimensional hypervolume, the intersection volume (which is three-dimensional) is similar to the base three-dimensional volume [2], like Fig. 2.

If we mark the volume of base three-dimension volume as  $V^{(3)}$ , the value of the forth dimension as  $h^{(4)}$ , the intersection volumes as  $V_i^{(3)}$ ,  $i=1, 2, \dots, n$ , like the derivation of equation 2 according to the similarity theorem:

$$V_i^{(3)} = \frac{V^{(3)} \cdot x_i^{(4)3}}{h^{(4)3}} \quad (6)$$

In equation 6,  $x_i^{(4)}$  is the distance from the base volume to each of the hypercubes  $H_i$ . As same as the derivation above, the hypervolume of each four-dimensional hypercubes can be expressed as:

$$V_i^{(4)} = \lim_{n \rightarrow \infty} \frac{h^{(4)}}{n} \cdot \frac{V^{(3)} \cdot x_i^{(4)3}}{h^{(4)3}}, i = 1, 2, \dots, n \quad (7)$$



**Fig. 2.** The picture of Fig. 2 means if we use a three-dimensional space paralleling to the base of the four-dimensional prismatic cone to intercept the four-dimensional hypervolume JABCD, the intersection volume  $J' - A'B'C'D'$  (which is three-dimensional) is similar to the base three-dimensional volume

So the whole hypervolume defined as  $V_{(4)}$  can be expressed as:

$$V^{(4)} = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{h^{(4)}}{n} \cdot \frac{V^{(3)} \cdot x_i^{(4)3}}{h^{(4)3}} \quad (8)$$

Because of  $\frac{h^{(4)}}{n} \in (0, 1)$

$$V^{(4)} = \int_0^{h^{(4)}} \frac{V^{(3)} \cdot x^{(4)3}}{h^{(4)3}} dx^{(4)} = \frac{1}{4} V^{(3)} h^{(4)} \quad (9)$$

## 2.2 Obtain approximations of four-dimensional convex hypervolume

Because the state-of-the art algorithms for hypervolume calculation are time-consuming, the runtime is more than  $O(d^4)$  [4],  $d$  is the number of edges in a hypervolume, or preform far from pretty while facing the value of the hypervolume is a interval with perturbation in neuron net although they can get a approximation of the hypervolume more accurate than ours [5].

**Definition 1.** Assume that there is a fully connected net  $N$  with  $n$  pre-activation layers and there are  $n_s$  neurons(or points) in the  $s$  th pre-activation layer of  $N$ ,  $s \in \{1, 2, \dots, n\}$ . Point  $i$  and  $j$  mean the  $i$  th and  $j$  th points among the  $n_s$  points

in the  $s$  th pre-activation layer of  $N$ ,  $i, j \in \{1, 2, \dots, n_s\}, i \neq j$ . A point also means a tuple  $(x, y)$ ,  $x$  is the output of the point as the pre-activation value to ReLU layer,  $y$  means the result of  $\text{ReLU}(x)$ .

When calculating 2-relu,  $k = 2$ , assume that the pre-activation value and ReLU result of point  $i$  are  $x_1$  and  $y_1$ , similarly we have  $x_2$  and  $y_2$  of point  $j$ . And we start splitting quadrants from  $x_1$  dimension, as a result, the  $y_2$  will be the additional forth dimension.

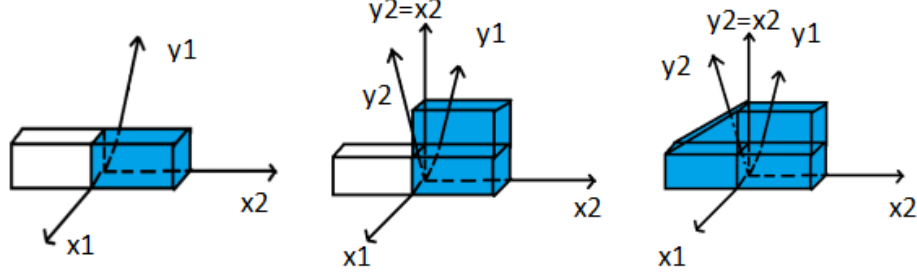
Given the set of three-dimensional constraints  $C^{(3)}$ ,  $c_{(i,j)}^{(3)} \in C^{(3)}$  is defined as three-dimensional constraints generated from point  $i$  and  $j$ . Given the set of four-dimensional constraints  $C^{(4)}$ ,  $c_{(i,j)}^{(4)} \in C^{(4)}$  is defined as four-dimensional constraints generated from point  $i$  and  $j$ . Define that  $c_{(i,j),\{x \leq 0\}}^{(3)}$  be the part of  $c_{(i,j)}^{(3)}$  satisfying the constrain  $x \leq 0$  and  $x \in \{x_1, x_2, y_1\}$ . Similarly,  $c_{(i,j),\{x \geq 0\}}^{(3)}$  be the part of  $c_{(i,j)}^{(3)}$  satisfying the constrain  $x \geq 0$ . In addition, define  $c_{(i,j),\{c_{(i,j)}^{(3)} \wedge c^{y_2}\}}^{(4)}$  as a  $c_{(i,j)}^{(4)}$  whose four-dimensional constraints are consisted of  $c_{(i,j)}^{(3)} \in C^{(3)}$  and a unitary linear constraint  $c^{y_2}$  about the forth dimension  $y_2$ , we can find that the constrains of  $c_{(i,j),\{c_{(i,j)}^{(3)} \wedge c^{y_2}\}}^{(4)}$  equals to the ones of  $c_{(i,j)}^{(3)} \wedge c^{y_2}$ . For example,  $c_{(i,j),\{c_{(i,j),\{x \leq 0\}}^{(3)} \wedge y_2 = 0\}}^{(4)}$  means the part of  $c_{(i,j)}^{(3)} \in C^{(3)}$  satisfying the constrain  $x \leq 0$  and a unitary linear constraint  $y_2 = 0$  about  $y_2$  make up a set of four-dimensional constraints  $c_{(i,j),\{c_{(i,j),\{x \leq 0\}}^{(3)} \wedge y_2 = 0\}}^{(4)} \in C^{(4)}$ .

**Definition 2.** According to Def. 1, we can define volume of the abstract zone of  $c_{(i,j)}^{(3)} \in C^{(3)}$  as  $V_{\{c_{(i,j)}^{(3)}\}}^{(3)}$ , and hypervolume of the abstract zone of  $c_{(i,j)}^{(4)} \in C^{(4)}$  as  $V_{\{c_{(i,j)}^{(4)}\}}^{(4)}$  whose upper and lower approximations are defined as  $V_{\{c_{(i,j)}^{(4)}\}}^{(4),upper}$  and  $V_{\{c_{(i,j)}^{(4)}\}}^{(4),lower}$  respectively.

Now, we have defined some symbols, it is time to get the formula of the four-dimensional approximations. For example, assume that one of the three-dimensional abstract zones generated by PRIMA is a cube, shown by the left picture of Fig. 3. Then we add the forth dimension  $y_2$  to expand the three-dimensional abstract zone to a four-dimensional one, the result shown by the middle of Fig. 3. Finally, the blue part of the right one is the abstract of

$\text{conv}(c_{(i,j),\{c_{(i,j),\{x_2 \geq 0\}}^{(3)} \wedge y_2 = x_2\}}^{(4)}, c_{(i,j),\{c_{(i,j),\{x_2 \leq 0\}}^{(3)} \wedge y_2 = 0\}}^{(4)})$  which is the final four-dimensional convex hull  $c_{(i,j)}^{(4)}$  obtained by PRIMA.

According to the properties of the four-dimensional abstract zone generated by PRIMA, we use  $\frac{1}{4}V_{\{c_{(i,j),\{x_2 \geq 0\}}^{(3)}\}}^{(3)}u_2$  as the under-approximation of hypervolume and  $\frac{1}{4}V_{\{c_{(i,j)}^{(3)}\}}^{(3)}u_2$  as the over-approximation of hypervolume, of which  $u_4$  is the upper bound of  $y_2$ . Here is a example in Fig. 4 about the under- and over-



**Fig. 3.** We assume that the left picture is one of the three-dimensional abstract zones  $c_{(i,j)}^{(3)}$  generated by PRIMA, the shape of the abstract zone is a cube, its blue part is the abstract zone of  $c_{(i,j),\{x_2 \geq 0\}}^{(3)}$ . The blue part of the middle one is the abstract zone of  $c_{(i,j),\{c_{(i,j),\{x_2 \geq 0\}}^{(3)} \wedge y_2 = x_2\}}^{(4)}$ . the blue part of the right one is the abstract of  $\text{conv}(c_{(i,j),\{c_{(i,j),\{x_2 \geq 0\}}^{(3)} \wedge y_2 = x_2\}}^{(4)}, c_{(i,j),\{c_{(i,j),\{x_2 \leq 0\}}^{(3)} \wedge y_2 = 0\}}^{(4)})$  which is the final four-dimensional convex hull  $c_{(i,j)}^{(4)}$  obtained by PRIMA.

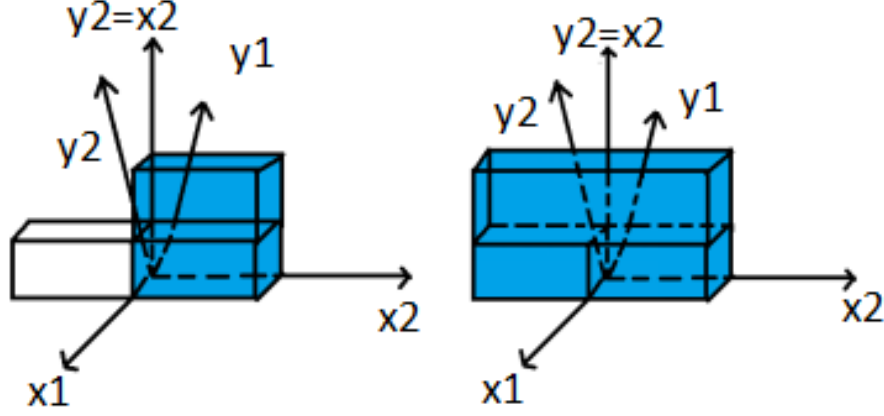
approximation based on Fig. 3 and we will give the proof about approximation in the following part.

Before giving the proof about approximation, give Theorem 2 first as the front theorem for the proof of approximation.

**Theorem 2.** *According to the geometric significance and properties of determinant [3], the volume of a three-dimensional convex polyhedra can be expressed as a sum of volumes of some three-dimensional prismatic cones, and the hypervolume of a four-dimensional convex polyhedron can be expressed as a sum of hypervolumes of some four-dimensional prismatic cones. The hypervolume of a four-dimensional convex polyhedron can be expressed by a sum of volumes of some three-dimensional prismatic cones.*

*Proof.* Assume that the hypervolume of a four-dimensional convex polyhedron is  $V^{(4)}$ , we have known that  $V^{(4)}$  can be expressed as a sum of hypervolumes of several four-dimensional prismatic cones marked as  $V_1^{(4)}, V_2^{(4)}, \dots, V_n^{(4)}$ , the base of each of the  $V_i^{(4)}$ ,  $i = 1, 2, \dots, n$  is marked as  $V_{base_i}^{(3)}$ ,  $i = 1, 2, \dots, n$ . We have known that  $V_{base_i}^{(3)}$  can be expressed as a sum of volumes of several three-dimensional prismatic cones marked as  $V_{i,1}^{(3)}, V_{i,2}^{(3)}, \dots, V_{i,n_i}^{(3)}$ ,  $n_i$  is the number of three-dimensional prismatic cones consisting  $V_{base_i}^{(3)}$ , with the help of Theorem. 1, we can get the relationships between  $V^{(4)}$  and  $V_{i,j}^{(3)}$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, n_i$ ,  $h_i^{(4)}$  is the forth dimension 'height' of each of  $V_i^{(4)}$ :

$$V^{(4)} = \sum_{i=1}^n V_i^{(4)} = \sum_{i=1}^n \frac{1}{4} h_i^{(4)} V_{base_i}^{(3)} = \sum_{i=1}^n \frac{1}{4} h_i^{(4)} \sum_{j=1}^{n_i} V_{i,j}^{(3)} \quad (10)$$



**Fig. 4.** In the left part of this picture, we directly use the  $V^{(3)}_{\{c^{(3)}_{(i,j),\{x_2 \geq 0\}}\}}$  as the base, the upper bound of  $x_2$ ,  $u_2$ , as the four-dimensional 'height' to structure the under-approximation  $V^{(4)}_{\{c^{(3)}_{(i,j),\{x_2 \geq 0\}} \wedge y_2 = x_2\}}$ . In the right part of this picture, we use the  $V^{(3)}_{\{c^{(3)}_{(i,j)}\}}$  as the base, the upper bound of  $x_2$ ,  $u_2$ , as the four-dimensional 'height' to structure the over-approximation  $V^{(4)}_{\{c^{(3)}_{(i,j)} \wedge y_2 = x_2\}}$

According to Theorem.1, we are able to find relationships between three-dimensional convex hull and four-dimensional convex hull in order to obtain the under- and over- approximation mentioned before.

*Proof.* For two selected points  $i$  and  $j$ , their precise constraints formed by PRIMA is marked as  $c^{(4)}_{(i,j),precise} \in C^{(4)}$ , so with the help of Def. 1 and Def. 2, the hypervolume of the abstract zone from  $c^{(4)}_{(i,j),precise}$  can be marked as  $V^{(4)}_{\{c^{(4)}_{(i,j),precise}\}}$ . There has a equation according to the process of PRIMA or k-ReLU (assume that split about  $x_1$  firstly so that the  $y_2$  will be the forth dimension):

$$c^{(4)}_{(i,j),precise} = conv(c^{(4)}_{(i,j),\{c^{(3)}_{(i,j),\{x_2 \leq 0\}} \wedge y_2 = 0\}}, c^{(4)}_{(i,j),\{c^{(3)}_{(i,j),\{x_2 \geq 0\}} \wedge y_2 = x_2\}}) \quad (11)$$

Previously there has inequality relationships:

$$V^{(4)}_{\{c^{(4)}_{(i,j),precise}\}} \geq V^{(4)}_{\{c^{(3)}_{(i,j),\{x_2 \geq 0\}} \wedge y_2 = x_2\}} \quad (12)$$

Otherwise, it will have contradictions with equation 11. Because of equation 12, we select  $V^{(4)}_{\{c^{(3)}_{(i,j),\{x_2 \geq 0\}} \wedge y_2 = x_2\}}$  as the  $V^{(4)}_{\{c^{(4)}_{(i,j)}\},lower}$ :

$$V_{\{c_{(i,j),\{x_2 \geq 0\}}^{(3)} \wedge y_2 = x_2\}}^{(4)} = V_{\{c_{(i,j)}^{(4)}\}, lower}^{(4)} = \frac{1}{4} V_{\{c_{(i,j),\{x_2 \geq 0\}}^{(3)}\}}^{(3)} h_{(4)} \quad (13)$$

In equation 13,  $h_{(4)}$  is the maximum value of the forth dimation  $y_2$ , because the constrains  $y_2 = x_2$  and  $x_2 \in [l_2, u_2]$ , with the help of Theorem. 1 and Theorem. 2, the finall equation for  $V_{\{c_{(i,j)}^{(4)}\}, lower}^{(4)}$  is:

$$V_{\{c_{(i,j)}^{(4)}\}, lower}^{(4)} = \frac{1}{4} V_{\{c_{(i,j),\{x_2 \geq 0\}}^{(3)}\}}^{(3)} u_2 \quad (14)$$

*Proof.* At First, mark  $c_{(i,j),\{c_{(i,j),\{x_2 \leq 0\}}^{(3)} \wedge y_2 = 0\}}^{(4)}$  as  $con_1$ ,  $c_{(i,j),\{c_{(i,j),\{x_2 \geq 0\}}^{(3)} \wedge y_2 = x_2\}}^{(4)}$  as  $con_2$  for convenience. As mentioned before, the precise four-dimensional constrains is  $c_{(i,j), precise}^{(4)}$  in equation 11. Because of  $x_2 \in [l_2, u_2]$ , we have:

$$con_1 = c_{(i,j),\{c_{(i,j),\{x_2 \leq 0\}}^{(3)} \wedge y_2 = 0\}}^{(4)} \subseteq c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)} \quad (15)$$

$$con_2 = c_{(i,j),\{c_{(i,j),\{x_2 \geq 0\}}^{(3)} \wedge y_2 = x_2\}}^{(4)} \subseteq c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)} \quad (16)$$

Next, we will use rebuttal method to prove that  $conv(con_1, con_2) \subseteq c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)}$ :

if  $conv(con_1, con_2)$  is not belongs to  $c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)}$ , exsits a point  $\alpha$ . Because  $conv(con_1, conv2)$  is a convex hull, so there exsits  $\beta_1, \beta_2 \in \{con_1, con_2\}$  and  $\lambda_0 \geq 0$  s.t.

$$\lambda_0 \beta_1 + (1 - \lambda_0) \beta_2 = \alpha \quad (17)$$

Because  $\beta_1, \beta_2 \in \{con_1, con_2\}$  and equation 16 and equation /refD2, there exsits  $\gamma_1, \gamma_2, \theta_1, \theta_2 \in conv(con_1, con_2) \subseteq c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)}$  and  $\lambda_1, \lambda_2 \in (0, 1)$  s.t.

$$\beta_1 = \lambda_1 \gamma_1 + (1 - \lambda_1) \gamma_2 \quad (18)$$

$$\beta_2 = \lambda_2 \theta_1 + (1 - \lambda_2) \theta_2 \quad (19)$$

Because  $\gamma_1, \gamma_2, \theta_1, \theta_2 \in c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)}$  and  $c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)}$  is a convex hull formed by PRIMA or k-ReLU,  $\beta_1, \beta_2 \in c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)}$  as the same as  $\alpha \in c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)}$ . Here is a contradiction.

Therefore, we have used rebuttal method to prove that  $conv(con_1, con_2) \subseteq c_{(i,j),\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)}$ .



$$\text{As a result, } V_{\{conv(con_1, con_2)\}}^{(4)} \leq V_{\left\{c_{(i,j)}^{(4)}, \{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}\right\}}^{(4)} = V_{\{c_{(i,j)}^{(3)} \wedge y_2 = x_2\}}^{(4)}$$

Finally with the help of Theorem. 1 and Theorem. 2, we get the over-approximation of  $V_{\{c_{(i,j)}^{(4)}\}, upper}^{(4)}$ :

$$V_{\{c_{(i,j)}^{(4)}\}, upper}^{(4)} = \frac{1}{4} V_{\{c_{(i,j)}^{(3)}\}}^{(3)} u_2 \quad (20)$$

Then, if we can find a method to get the volume of the three-dimensional abstract zone generated from point  $i$  and  $j$ , we will get a better group of points to make k-ReLU for reducing time without losing a lot of accuracy.

### 2.3 Obtain the volume of three-dimensional volume

At first, because we only discuss the situation that  $k = 2$ , so it is not quiet difficult for us to use enumeration method to get the whole situations of points which will form the three-dimensional convex hull according to the process of PRIMA or k-ReLU. Then, make the points we get in the first step as the input of Melkman's Algorithm [6] to get the vertex-representation of the convex hull. Finally, we use three-order determinant [3], [7] to calculate the volume of the three-dimensional abstract zone about  $i$  and  $j$ . Through equation 14 and equation 20, we can know that the four-dimensional approximation is in direct proportion to their three-dimensional approximation, so we only need to compare the three-dimensional. Algorithm. 1 shows the process of our algorithm.

## 3 Experiments

In this section, we will evaluate the effectiveness of Our selection strategy and show that it improves over state-of-the-art verifiers on a range of challenging perturbations on ReLU-based networks. Further, we will also give the consparison among the robust radius of verifiers with variours selecting strategies on MNIST networks under  $k = 2$ .

### 3.1 Experimental setup

The neural network certification benchmarks for fully connected networks were worked on a 8 cores 2.90GHz Intel(R) Core(TM) i7-10700 CPU with 12GB of main memory and use Gurobi 9.5.1 for solving LP problems [8].

### 3.2 Image Classification with ReLU activation

For our experiments, we use the same setup as PRIMA in ERAN Toolbox [9], before multi-verificaiton, we use DEEPOLY to do backpropagation once to make the three-dimensional octahedral inputs required by PRIMA. All constrains generated from PRIMA are added to the LP encoding of the network, after all layers

---

**Algorithm 1:** Our algorithm

---

**Input:** the neuron net information  $nn$ , the number of present pre-activation layer  $layerdepth$ , the lower and upper bound of the present pre-activation layer  $lb, ub$ , the lower and upper bound of the previous pre-activation layer  $lbi, ubi$

**Output:** a list record selected points pairs  $kactargs$  and points haven't pair  $restpoints$

```

1 split0  $\leftarrow$  unkownact(lb,ub);
2 w  $\leftarrow$  nn.weight[layerdepth];
3 b  $\leftarrow$  nn.bias[layerdepth];
4 record = -np.ones((4,len(split0)));
5 for i in range(len(split0)-1) do
6     for j in range(i+1,len(split0)) do
7          $V_{lower}, V_{upper} \leftarrow$  3Dcalculation(w,b,lb,ub,lbi,ubi,i,j);
8         if record[0][i] == -1 then
9             if  $V_{upper}$  and  $V_{lower}$  then
10                 | record[i]  $\leftarrow$   $V_{lower}, V_{upper}, split0[j], False$ 
11             end
12         end
13         if record[0][j] == -1 then
14             if  $V_{upper}$  and  $V_{lower}$  then
15                 | record[j]  $\leftarrow$   $V_{lower}, V_{upper}, split0[i], False$ 
16             end
17         end
18         if  $V_{upper} \leq record[1][i]$  and  $V_{lower} \geq record[0][i]$  then
19             | record[i]  $\leftarrow$   $V_{lower}, V_{upper}, split0[j]$ 
20         end
21         if  $V_{upper} \leq record[1][j]$  and  $V_{lower} \geq record[0][j]$  then
22             | record[j]  $\leftarrow$   $V_{lower}, V_{upper}, split0[i]$ 
23         end
24         if  $V_{upper} \leq record[0][i]$  then
25             | record[i]  $\leftarrow$   $V_{lower}, V_{upper}, split0[j], True$ 
26         end
27         if  $V_{upper} \leq record[0][j]$  then
28             | record[j]  $\leftarrow$   $V_{lower}, V_{upper}, split0[i], True$ 
29         end
30     end
31 end
32 restpoints.append[i for i in split0 if record[3][i] == False];
33 kactargs  $\leftarrow$  getpair(record,split0);
34 return kactargs,restpoints;
```

---

are processed, an LP solver will prove the property. *PRIMA-h* is the PRIMA using heuristic selecting strategy in ERAN Toolbox. *PRIMA-all* is the PRIMA considering all of the  $C_n^2$  constrains while verification,  $n$  is the number of the points with lower bound  $l \leq 0$  and upper bound  $u \geq 0$  in one pre-activation layer. *PRIMA-o* is the PRIMA using our selecting strategy mentioned in Sec. 2. Our experiment process on MNIST  $10 \times 80$  network.

| Netwr<br>ok             | $\epsilon$ | PRIMA-h |         |     | PRIMA-all |         |     | PRIMA-o |         |     |
|-------------------------|------------|---------|---------|-----|-----------|---------|-----|---------|---------|-----|
| MNIST<br>$10 \times 80$ |            | Ver     | T(s)    | Acc | Ver       | T(s)    | Acc | Ver     | T(s)    | Acc |
|                         | 0.026      | 43753   | 3243.3  | 94  | 54034     | 3311.4  | 94  | 4226    | 3160.2  | 94  |
|                         | 0.05       | 137331  | 4447.4  | 90  | 168460    | 4551.5  | 91  | 10544   | 4183.3  | 90  |
|                         | 0.1        | 865177  | 9703.8  | 47  | 1049824   | 10240.6 | 49  | 125906  | 8233.3  | 49  |
|                         | 0.1125     | 1225209 | 11970.9 | 24  | 1510080   | 12729.5 | 30  | 97636   | 10133.1 | 29  |

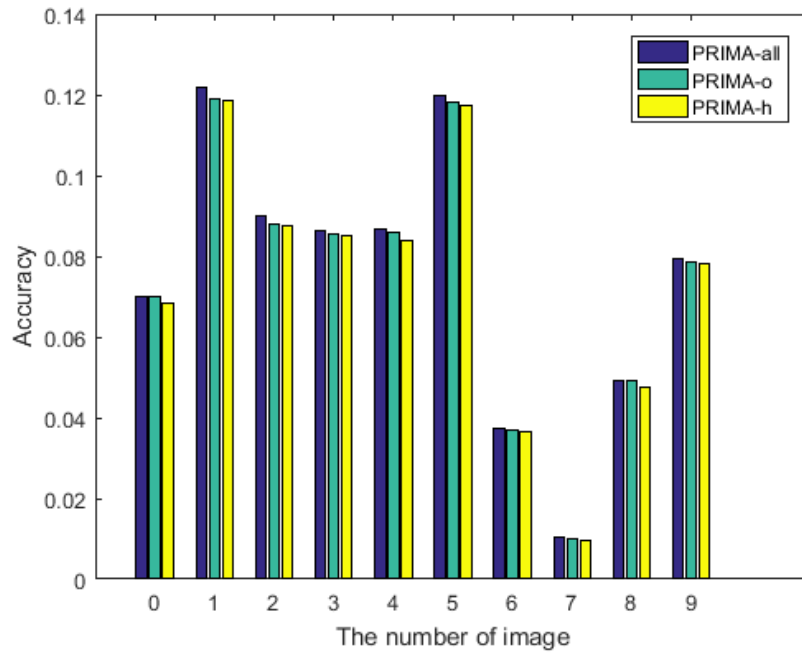
**Fig. 5.** In the this table, *Ver* is the number of constrains generated by the strategy. *T* is the whole time for the program of the strategy. *Acc.* is the number of verified images among 100 pictures in MNIST test set.

According to Fig. 5, on the one hand, we can find that the number of verified images obtained by PRIMA-o is always more than the one obtained by PRIMA-h, and the execution time of ours is always less than the one of PRIMA-h. The reason is that the heuristic selection strategy of PRIMA-h ignore some points while doing multi-verification in order to save time, the ignored part bring the loss of accuracy. On the other hand, because PRIMA-all consider all of the possible groups to generate constrains, the constrains generated from our strategy are included in it. This experiment prove that our strategy can save a lot of computer sources with losing a little of accuracy through comparing with PRIMA-all. For example, when  $\epsilon = 0.1125$ , PRIMA-o work about 20.3% faster than PRIMA-all with losing about 3.3% verified pictures on the  $10 \times 80$  MNIST network.

### 3.3 Improvement for precision on high-dimensional networks

In this experiment, we take the first 10 images from test set of MNIST as input. Respectively use PRIMA-all, PRIMA-o, and PRIMA-h to calculate the robust radius of these 10 images on 10times 80 MNIST network in order to compare the difference on accuracy among these three strategies, under  $k = 2$ . The method of calculating robust radius is dichotomy [10] and the iterations is 15, accurating to the extreme position, the result shown by Fig. 6.

In the histogram in Fig 6, the blue part means the robust radius obtained by PRIMA-all, the cyan one means the robust radius obtained by PRIMA-o,



**Fig. 6.** Comparison of verified robust radius among PRIMA-all, PRIMA-h, and PRIMA-o.

the yellow one means the robust radius obtained by PRIMA-h. PRIMA-all cost XXXs to get only about 0.0017 advance than PRIMA-h, however, PRIMA-o just spending XXXs will have about 0.0015 advance than PRIMA-h. According to the experiment, PRIMA-o fills the accuracy gap between PRIMA-all and PRIMA-h and cost less time than others.

## 4 Conclusion

Among the several of NN verification methods, considering the relationships and constraints among points in the same pre-activation layer is a new and promising direction. It provides a novel and effective way to overcome the *Delta* barrier produced by just considering the constraints of one point. In this paper, according to some experimental facts from other work and mathematical theories, we propose a new method and thinking to select which pairs of points is better to generate multi-neuron approximations. The experimental result shows that our method can economize a lot of time and computer resources with losing a little of accuracy under  $k = 2$ .

It is worth noticing that our method can be expanded to higher dimensional situation to solve the situation when  $k \geq 2$  and can be used under sigmoid or tanh instead of ReLU, and these will be the research directions of our team. On the other hand, the most interesting thing is that, although it has work that uses three-dimensional convex hull to do the symbolic propagation, it doesn't expand to higher dimension. It is possible to use the approximations mentioned in our method to achieve multi-neuron propagation with high accuracy and lower time cost in the future.

## References

1. G. P. Collins : The shapes of space, Sci. Amer., 291 (2004), pp. 94–103. SCAMAC 0036-8733
2. Dimensions:a walk through mathematics, [http://www.dimensions-math.org/Dim\\_ZH\\_si.htm](http://www.dimensions-math.org/Dim_ZH_si.htm) Last accessed in June 2022
3. Zhaokui He, Sujuan Zhang, Zhongxi Sun, The Geometric Significance of Determinant and Calculation of Polyhedron Volume. COLLEGE MATHEMATICS **37**(3), 105–109 (2021)
4. Karl Bringmann, Tobias Friedrich, Approximating the volume of unions and intersections of high-dimensional geometric objects. Computational Geometry **43**, 601–610 (2010)
5. LARRY T. COOK, P. NONG COOK, KYO RAK LEE, SOLOMON BATNITZKY, BERT Y. S. WONG, STEVEN L. FRITZ, JONATHAN OPHIR, SAMUEL J. DWYER III, LAWRENCE R. BIGONGIARI, AND ARCH W. TEMPLETON, An Algorithm for Volume Estimation Based on Polyhedral Approximation, IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING **27**(9), 493–500 (1980)
6. Qhull. "The Geometry Center Home Page." <http://www.qhull.org/> Last accessed in June 2022

7. J.B.Lasserre, An Analytical Expression and an Algorithm for the Volume of a Convex Polyhedron in  $R^n$ , JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS **39**(3), 363–377 (1983)
8. Gurobi Optimization, LLC, "Gurobi optimizer reference manual." <http://www.gurobi.com> Last accessed in July 2022
9. SRI Lab, ETH Zurich, "eran." <https://github.com/eth-sri/eran> Last accessed in July 2022
10. Guy Katz, Clark Barrett, David Dill, Kyle Julian and Mykel Kochenderfer, Reluplex: An Efficient SMT Solver for Verifying Deep Neural Networks, CAV (2017)

*Proof.* Proofs, examples, and remarks have the initial word in italics, while the following text appears in normal font.

For citations of references, we prefer the use of square brackets and consecutive numbers. Citations using labels or the author/year convention are also acceptable. The following bibliography provides a sample reference list with entries for journal articles [11], an LNCS chapter [12], a book [13], proceedings without editors [14], and a homepage [15]. Multiple citations are grouped [11,12,13], [11,13,?,15].

11. Author, F.: Article title. Journal **2**(5), 99–110 (2016)
12. Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016). <https://doi.org/10.1007/1234567890>
13. Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, Location (1999) pp. 94–103. SCAMAC 0036-8733
14. Author, A.-B.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010)
15. LNCS Homepage, <http://www.springer.com/lncs>. Last accessed 4 Oct 2017