APPENDIX

PROOF OF PROPOSITION 2

Proof. Denote $\mathbf{z} = t\mathbf{x} + (1-t)\mathbf{y}$ and assume that at $\mathbf{z}$ the i-th function is largest, i.e. $F(\mathbf{z}) = A_i\mathbf{z} + b_i$. Then

$$F(\mathbf{z}) = t(A_i\mathbf{x} + b_i) + (1-t)(A_i\mathbf{y} + b_i) \leq tF(\mathbf{x}) + (1-t)F(\mathbf{y})$$

□

PROOF OF PROPOSITION 3

Proof. The proof is by induction. We need to prove two facts. First, applying a linear function to a vector of DCPAs produces another vector of DCPAs; second, that a maximum of two DCPAs is a DCPA.

Let $F - G$ be a vector of DCPAs, where F and G are vectors of n CPAs and A be an $m \times n$ matrix with real coefficients. Write $A = A_+ - A_-$ where both A_+ and A_- have non-negative entries. Then we have

$$A\,(F-G) = (A_+ - A_-)\,(F-G) = (A_+F + A_-G) - (A_-F + A_+G).$$

This proves the first fact.

The second fact is easy to see from $\max\{a,b\} + c = \max\{a+c, b+c\}$ and $\max\{a, \max\{b,c\}\} = \max\{a,b,c\}$. □

PROOF OF PROPOSITION 7

Proof. The proof follows the following steps.

1. Let $\mathbf{c} = (\mathbf{a}, b)$ and $H = (\mathbf{a} \mapsto \mathbf{x}^\intercal\mathbf{a}^\intercal + y$. Then both $\mathbf{c} \in H$ and $\mathcal{R}(H) \in \mathcal{R}(\mathbf{c})$ are equivalent to $b = \mathbf{x}^\intercal\mathbf{a} + y$.
2. k-dimensional dual plane F can be written as an intersection of $d-k$ dual hyperplanes $\mathcal{R}^{-1}(\mathbf{z}_0), \ldots, \mathcal{R}^{-1}(\mathbf{z}_{d-k})$. A dual point $\mathcal{R}^{-1}(f)$ belongs to F if and only if it is a dual of a real hyperplane f that contains the real points $\mathbf{z}_0, \ldots, \mathbf{z}_{d-k}$. Their affine span is the common plane we are looking for, and what we christen $\mathcal{R}(F)$.
   It is affinely spanned by $d-k+1$ points, so its dimension is at most $d-k$. If it was smaller, we could forget some $\mathbf{z}_i$, which means that F was an intersection of $d-k-1$ hyperplanes, and had dimension at least $k+1$.
3. F is contained in G if and only if for any hyperplane H we have
   $$G \subseteq H \Rightarrow F \subseteq H$$
   This happens precisely when for all points $\mathbf{z} = \mathcal{R}(H)$ we have
   $$\mathbf{z} \in \mathcal{R}(G) \Rightarrow \mathbf{z} \in \mathcal{R}(F)$$
   that is $\mathcal{R}(G) \subseteq \mathcal{R}(F)$.
4. Let $f : \mathbf{x} \mapsto \mathbf{a}^\intercal\mathbf{x} + b$. Then $p(\mathcal{R}^{-1}(f)) = \mathbf{a}$, which is perpendicular to surfaces $\mathbf{a}^\intercal\mathbf{x} = \text{const}$.
5. Let $\mathbf{c} = (\mathbf{a}, b)$ and $H : \mathbf{d} \mapsto \mathbf{x}^\intercal\mathbf{d} + y$. Then both $\mathbf{c} \succ H$ and $\mathcal{R}(\mathbf{c}) \succ \mathcal{R}(H)$ are equivalent to $b > \mathbf{x}^\intercal\mathbf{a} + y$.
6. Suppose $\mathbf{c} = (\mathbf{a}, b), \mathbf{c}' = (\mathbf{a}, b+\Delta)$, and denote $f : \mathbf{x} \mapsto \mathbf{a}^\intercal\mathbf{x} + b$. Then $\mathcal{R}(\mathbf{c}) = f, \mathcal{R}(\mathbf{c}') = f + \Delta$ – these functions differ by a constant, so specify parallel planes. The proof for $\mathcal{R}^{-1}$ is analogous.

□

PROOF OF PROPOSITION 9

Proof. Firstly, let us compare the planes dual to two points, $\mathbf{s}_1$ and $\mathbf{s}_2$, such that $\mathbf{s}_1$ lies directly above $\mathbf{s}_2$. This means that they differ only at the very last coordinate—let's say that $\mathbf{s}_1 = (\mathbf{a}_1, b_1)$ and $\mathbf{s}_2 = (\mathbf{a}_2, b_2)$ where $b_1 \geq b_2$. Then the dual planes $\mathcal{R}(\mathbf{s}_1)$ and $\mathcal{R}(\mathbf{s}_2)$ are precisely

$$\mathcal{R}(\mathbf{s}_1) = \{(\mathbf{x}, y_1) | y_1 = (\mathbf{a}_1)^\intercal\mathbf{x} + b_1\},$$

$$\mathcal{R}(\mathbf{s}_2) = \{(\mathbf{x}, y_2) | y_2 = (\mathbf{a}_2)^\intercal \mathbf{x} + b_2\},$$

and since $(\mathbf{a}_1)^\intercal \mathbf{x} + b_1 \geq (\mathbf{a}_2)^\intercal \mathbf{x} + b_2$ for all $\mathbf{x} \in \mathbb{R}^d$, the plane $\mathcal{R}(\mathbf{s}_1)$ lies above $\mathcal{R}(\mathbf{s}_2)$.

Secondly, let us consider a point $\mathbf{s}$ in the dual space lying on a segment whose endpoints are $\mathbf{s}_1$ and $\mathbf{s}_2$. But then for some $p \in [0, 1]$ we have $\mathbf{s} = p \cdot \mathbf{s}_1 + (1 - p) \cdot \mathbf{s}_2$ and thus

$$(\mathbf{s})^\intercal \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} = p \cdot \left((\mathbf{s}_1)^\intercal \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \right) + (1 - p) \cdot \left((\mathbf{s}_2)^\intercal \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \right),$$

so, in particular,

$$(\mathbf{s})^\intercal \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \leq \max \left\{ (\mathbf{s}_1)^\intercal \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}, \quad (\mathbf{s}_2)^\intercal \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \right\}.$$

Thirdly, we want to piece the two together. For a point $\mathbf{s}_2$ lying below $\mathcal{U}(S)$, let us choose a point $\mathbf{s}_1 \in \mathcal{U}(S)$ lying exactly above $\mathbf{s}_2$. The plane defined by it lies above the one defined by $\mathbf{s}_2$ according to the first paragraph. Now we only need to show that points on $\mathcal{U}(S)$ define planes lying below the minimum, but this follows from the second paragraph and the fact that all points on a convex hull of a finite set of points can be generated by taking segments whose ends lie in the hull and adding all of the points of the segment to the hull. □

PROOF OF PROPOSITION 14

Proof. This is a straightforward consequence of the more elementary identities for scalar a, b: after reducing to upper hulls we have

$$(a + b)X = (aX) \oplus (bX) \tag{5.1}$$
$$a(X \oplus Y) = (aX) \oplus (aY) \tag{5.2}$$
$$(ab)X = a(bX) \tag{5.3}$$
$$a(X \cup Y) = (aX) \cup (aY) \tag{5.4}$$

Except 5.1, all of these hold even before taking the hull. To deal with this one, note that

$$\begin{aligned} (a + b)X &= \{ax + bx | x \in X\} \\ &\subseteq \{ax_1 + bx_2 | x_1, x_2 \in X\} = (aX) \oplus (bX) \end{aligned}$$

so we have $\mathcal{U}((a + b)X) \subseteq \mathcal{U}\big((aX) \oplus (bX)\big)$. To see the reverse inclusion, write

$$ax_1 + bx_2 = \tfrac{a}{a+b}(a + b)x_1 + \tfrac{b}{a+b}(a + b)x_2$$

which means that

$$(aX) \oplus (bX) \subseteq \mathcal{U}\big((a + b)X\big)$$

□

PROOF OF PROPOSITION 16

Proof. Firstly, let us note that

$$\begin{aligned} A_l\, F_{l-1} =& (A_l^+ - A_l^-) \left(\mathcal{R}(P_{l-1}) - \mathcal{R}(N_{l-1}) \right) \\ =& \left(A_l^+\, \mathcal{R}(P_{l-1}) + A_l^-\, \mathcal{R}(N_{l-1}) \right) \\ & - \left(A_l^-\, \mathcal{R}(P_{l-1}) + A_l^+\, \mathcal{R}(N_{l-1}) \right) \\ =& \mathcal{R}\big((A_l^+ \otimes P_{l-1}) \oplus (A_l^- \otimes N_{l-1}) \big) \\ & - \mathcal{R}\big((A_l^- \otimes P_{l-1}) \oplus (A_l^+ \otimes N_{l-1}) \big). \end{aligned}$$

Now, we use the fact that $\max\{x-y,0\} = \max\{x,y\} - y$ to get that for $N_l = (A_l^- \otimes P_{l-1}) \oplus (A_l^+ \otimes N_{l-1})$, we have

$$\begin{aligned}
&\sigma(A_l\, F_{l-1})\\
&= \max\{\mathcal{R}\big((A_l^+ \otimes P_{l-1}) \oplus (A_l^- \otimes N_{l-1})\big), \mathcal{R}(N_l)\} - \mathcal{R}(N_l)\\
&= \mathcal{R}\big((A_l^+ \otimes P_{l-1}) \oplus (A_l^- \otimes N_{l-1}) \cup N_l\big) - \mathcal{R}(N_l),
\end{aligned}$$

and thus, for $P_l = (A_l^+ \otimes P_{l-1}) \oplus (A_l^- \otimes N_{l-1}) \cup N_l$, we get

$$F_l = \sigma(A_l F_{l-1}) = \mathcal{R}(P_l) - \mathcal{R}(N_l).$$

□

PROOF OF PROPOSITION 18

Proof. k-cell of $\mathcal{T}(S)$ is the region defined by the system

$$\begin{aligned}
f_{i_0}(\mathbf{x}) &= \cdots = f_{i_{d-k}}(\mathbf{x})\\
f_{i_0}(\mathbf{x}) &\geq f_j(\mathbf{x}) \text{ for } j \neq i_0, \ldots, i_{d-k}
\end{aligned} \tag{5.5}$$

This can be written as

$$(\mathbf{x}, y) \in f_{i_0}, \ldots, f_{i_{d-k}} \qquad (\mathbf{x}, y) \succcurlyeq f_j$$

In dual space this becomes

$$\mathcal{R}^{-1}((\mathbf{x}, y)) \ni \mathcal{R}^{-1}(f_{i_0}), \ldots, \mathcal{R}^{-1}(f_{i_{d-k}})$$

$$\mathcal{R}^{-1}((\mathbf{x}, y)) \succcurlyeq \mathcal{R}^{-1}(f_j)$$

Therefore, the duals of points of the k-cell are precisely the dual planes containing the $(d-k)$-cell on vertices $\mathcal{R}^{-1}(f_{i_0}), \ldots, \mathcal{R}^{-1}(f_{i_{d-k}})$ and tangent to the upper convex hull. □

PROOF OF PROPOSITION 19

Proof. The cell of $\mathcal{T}(P \cup N)$ is a boundary cell iff in the equation 5.5, we have both some function $f_i \in \mathcal{R}(P)$ and some function $g_j \in \mathcal{R}(N)$. This happens exactly when the dual cell has some vertex $\mathcal{R}^{-1}(f_i) \in P$ as well as some vertex $\mathcal{R}^{-1}(g_j) \in N$. □

PROOF OF PROPOSITION 20

Proof. Again, as before, we need to identify those linear pieces of $\max\{F, G\}$, which lie on the linear pieces of F and of G. However, this means identifying cells of $\mathcal{U}(P \cup N)$ which *contain* a cell of $\mathcal{U}(P)$ and a cell of $\mathcal{U}(N)$ (this is due to the duality reversing containment of hyperplanes; we mean set-wise containment here, not containment as subcells). □

PROOF OF PROPOSITION 22

Proof. A k-dimensional cell σ is the set of $\mathbf{x}$ satisfying the system

$$\begin{aligned}
f_{i_0}(\mathbf{x}) &= \cdots = f_{i_a}(\mathbf{x}) = s > f_{i'}(\mathbf{x})\\
g_{j_0}(\mathbf{x}) &= \cdots = g_{j_b}(\mathbf{x}) = t > g_{j'}(\mathbf{x})
\end{aligned}$$

Where $a + b = d - k$. This can be expressed as relations in the real space

$$\begin{aligned}
(\mathbf{x}, s) &\in f_{i_0}, \ldots, f_{i_a} \qquad (\mathbf{x}, s) \succ f_{i'}\\
(\mathbf{x}, t) &\in g_{j_0}, \ldots, g_{j_b} \qquad (\mathbf{x}, t) \succ g_{j'}
\end{aligned}$$

After passing to the dual space this becomes

$$\mathcal{R}^{-1}\big((\mathbf{x},s)\big) \ni \mathcal{R}^{-1}(f_{i_0}),\ldots,\mathcal{R}^{-1}(f_{i_a}) \tag{5.6}$$
$$\mathcal{R}^{-1}\big((\mathbf{x},s)\big) \succ \mathcal{R}^{-1}(f_{i'}) \tag{5.7}$$
$$\mathcal{R}^{-1}\big((\mathbf{x},t)\big) \ni \mathcal{R}^{-1}(g_{j_0}),\ldots,\mathcal{R}^{-1}(g_{j_b}) \tag{5.8}$$
$$\mathcal{R}^{-1}\big((\mathbf{x},t)\big) \succ \mathcal{R}^{-1}(g_{j'}) \tag{5.9}$$

We know that $\mathcal{R}^{-1}\big((\mathbf{x},s)\big)$ and $\mathcal{R}^{-1}\big((\mathbf{x},t)\big)$ are a pair of parallel hyperplanes; the former is tangent to $\mathcal{U}(P)$ (5.7) and contains its a-cell (5.6), while the latter is tangent to $\mathcal{U}(N)$ (5.9) and contains its b-cell (5.8).

View these hyperplanes as subsets of $\mathbb{R}^{d+1}$ and consider their Minkowski sum $\mathcal{R}^{-1}\big((\mathbf{x},s)\big) \oplus \mathcal{R}^{-1}\big((\mathbf{x},t)\big)$. It is straightforward to verify that it equals the hyperplane $\mathcal{R}^{-1}\big((\mathbf{x},s+t)\big)$. Since the relation $\succ$ of lying above is preserved by translations, we have

$$\mathcal{R}^{-1}\big((\mathbf{x},s+t)\big) = \mathcal{R}^{-1}\big((\mathbf{x},s)\big) \oplus \mathcal{R}^{-1}\big((\mathbf{x},t)\big) \succcurlyeq \mathcal{R}^{-1}(f_i) + \mathcal{R}^{-1}(g_j) \qquad \text{for all } \mathcal{R}^{-1}(f_i) \in P, \mathcal{R}^{-1}(g_j) \in N$$

This means that the plane $\mathcal{R}^{-1}\big((\mathbf{x},s+t)\big)$ is tangent to $\mathcal{U}(P \oplus N)$. Also, it contains the $(a+b=d-k)$-cell σ' on vertices

$$\big\{\mathcal{R}^{-1}(f_{i_\alpha}) + \mathcal{R}^{-1}(g_{j_\beta}) \mid 0 \le \alpha \le a, 0 \le \beta \le b\big\} \tag{5.10}$$

Conversely, suppose a hyperplane H is tangent to $\mathcal{U}(P \oplus Q)$ and contains the $(d-k)$-cell σ'on the vertices from equation 5.10. Let $\mathbf{x} = p(H)$ be the vector of linear coefficients of H. If we had $f_{i'}(\mathbf{x}) > f_{i_\alpha}(\mathbf{x})$ for any $i' \notin \{i_0,\ldots,i_a\} \ni i_\alpha$, then the point $\mathcal{R}^{-1}(f_{i'}) + \mathcal{R}^{-1}(g_{j_0})$ would lie above H, which is impossible. Therefore we must have

$$f_{i_0}(\mathbf{x}) = \cdots = f_{i_a}(\mathbf{x}) > f_{i'}(\mathbf{x}) \tag{5.11}$$

and a similar set of conditions involving g's. This means that $x = p(H)$ lies in the real cell σ.

These functions are mutually inverse, and hence provide a bijection between real points of σ and dual tangent hyperplanes containing σ'.

Since every point of the real space belongs to a unique cell, and every dual hyperplane tangent to $\mathcal{U}(P \oplus N)$ intersects it in a unique cell, the assignment $\sigma \leftrightarrow \sigma'$ is bijective. ☐

Remark 1. Sign of the function on the cell (equivalently, the class to which the region belongs) depends on which of $\mathcal{R}^{-1}\big((\mathbf{x},s)\big), \mathcal{R}^{-1}\big((\mathbf{x},t)\big)$ lies above the other.

NUMERICAL EXPERIMENTS DETAILS

The neural networks are initialized by the default Uniform distribution[1]. For all ReLU neural networks, the optimization is done by stochastic gradient descent with learning rate= 0.1, momentum= 0.9 and weight decay= 0.001 (if not specified otherwise).

2D spiral The synthetic spiral data is from the two-dimensional distribution $P = (\rho \sin\theta + 0.04, \rho\cos\theta)$ where $\rho = (\theta/4\pi)^{4/5} + \epsilon$ with selected θ from $(0, 4\pi]$ and $\epsilon \sim \text{unif}([-0.03, 0.03])$. We draw 300 positive and 300 negative training samples from $-P$ and P, respectively, with a random seed fixed for every run. Both the Gaussian noise injection strength and the adversarial training strength are set at 0.01.

2D Gaussian mixture There are 3×3 mixing components, each is an isotropic Gaussian with standard deviation $\sigma = 0.1$. The means are grid points from $\{-1, 0, 1\} \times \{-1, 0, 1\}$. The mixing weight is equal for all components. Both the Gaussian noise injection strength and the adversarial training strength are set at 0.1.

Below we show some training trends for CE, Noisy and Adv in the Gaussian mixture case. It is worth noting that all trend plots in this work, including Figure 4 and 5 are smoothed with moving averages.

[1]The default weight initialization in `torch.nn.linear` is uniform on $[-\sqrt{1/N}, \sqrt{1/N}]$ where N is the width.

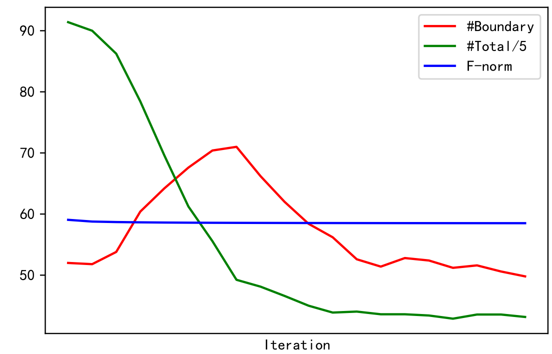


Figure 6: CE training trends of #Boundary (red), #Total (green), F-norm (red) vs. iteration in the Gaussian mixture case.

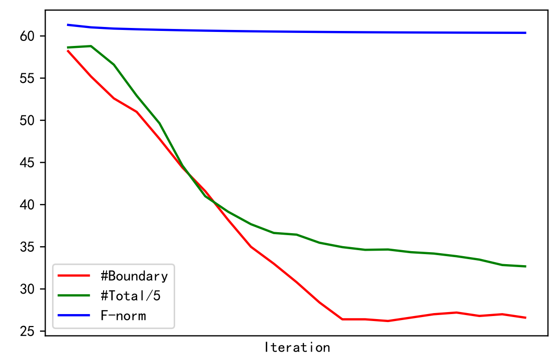


Figure 7: Noisy training trends of #Boundary (red), #Total (green), F-norm (red) vs. iteration in the Gaussian mixture case.

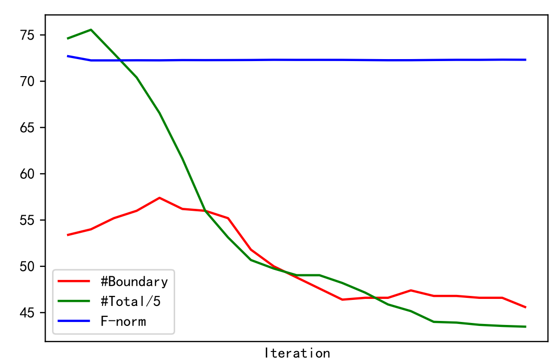


Figure 8: Adv training trends of #Boundary (red), #Total (green), F-norm (red) vs. iteration in the Gaussian mixture case.

AN EXAMPLE OF COMPUTATION WITH PROPOSITION 16

First we should note that in a standard ReLU network the transition functions are any affine functions but we can introduce a 'dummmy dimension' to realise these as *linear* functions.

We will consider a very simple network with two-dimensional input, one hidden layer with three neurons, and the following transition matrices (with the dummy dimension included). For illustrative purposes we assume that ReLU is applied also at the last layer.

$$A_1 = \begin{bmatrix} 1 & -0.5 & 4 \\ -2 & 1 & 0 \\ 3 & 3 & -1 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.5 & -1 & -0.5 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The input function $F_0 = (x, y, 1)$ (where the last coordinate is a dummy) is decomposed into $\mathcal{R}(P_0) - \mathcal{R}(N_0)$ with

$$P_0 = \begin{pmatrix} \{(1,0,0)\} \\ \{(0,1,0)\} \\ \{(0,0,1)\} \end{pmatrix}, \quad N_0 = \begin{pmatrix} \{(0,0,0)\} \\ \{(0,0,0)\} \\ \{(0,0,0)\} \end{pmatrix}.$$

To compute P_1 and N_1, we need to decompose the matrix A_1 into its positive and negative parts A_1^+ and A_1^-.

$$\begin{aligned}
N_1 &= (A_1^+ \otimes N_0) \oplus (A_1^- \otimes P_0) \\
&= \left(\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 3 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \otimes \begin{pmatrix} \{(0,0,0)\} \\ \{(0,0,0)\} \\ \{(0,0,0)\} \end{pmatrix} \right) \oplus \left(\begin{bmatrix} 0 & 0.5 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \otimes \begin{pmatrix} \{(1,0,0)\} \\ \{(0,1,0)\} \\ \{(0,0,1)\} \end{pmatrix} \right) \\
&= \begin{pmatrix} 1\{(0,0,0)\} \oplus 0\{(0,0,0)\} \oplus 4\{(0,0,0)\} \\ 0\{(0,0,0)\} \oplus 1\{(0,0,0)\} \oplus 0\{(0,0,0)\} \\ 3\{(0,0,0)\} \oplus 3\{(0,0,0)\} \oplus 0\{(0,0,0)\} \\ 0\{(0,0,0)\} \oplus 0\{(0,0,0)\} \oplus 1\{(0,0,0)\} \end{pmatrix} \oplus \begin{pmatrix} 0\{(1,0,0)\} \oplus 0.5\{(0,1,0)\} \oplus 0\{(0,0,1)\} \\ 2\{(1,0,0)\} \oplus 0\{(0,1,0)\} \oplus 0\{(0,0,1)\} \\ 0\{(1,0,0)\} \oplus 0\{(0,1,0)\} \oplus 1\{(0,0,1)\} \\ 0\{(1,0,0)\} \oplus 0\{(0,1,0)\} \oplus 0\{(0,0,1)\} \end{pmatrix} \\
&= \begin{pmatrix} \{(0,0,0)\} \\ \{(0,0,0)\} \\ \{(0,0,0)\} \\ \{(0,0,0)\} \end{pmatrix} \oplus \begin{pmatrix} \{(0,0.5,0)\} \\ \{(2,0,0)\} \\ \{(0,0,1)\} \\ \{(0,0,0)\} \end{pmatrix} = \begin{pmatrix} \{(0,0.5,0)\} \\ \{(2,0,0)\} \\ \{(0,0,1)\} \\ \{(0,0,0)\} \end{pmatrix} \\
P_1 &= (A_1^+ \otimes P_0) \oplus (A_1^- \otimes N_0) \cup N_1 \\
&= \begin{pmatrix} \{(1,0,4)\} \\ \{(0,1,0)\} \\ \{(3,3,0)\} \\ \{(0,0,1)\} \end{pmatrix} \oplus \begin{pmatrix} \{(0,0,0)\} \\ \{(0,0,0)\} \\ \{(0,0,0)\} \\ \{(0,0,0)\} \end{pmatrix} \cup \begin{pmatrix} \{(0,0.5,0)\} \\ \{(2,0,0)\} \\ \{(0,0,1)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&= \begin{pmatrix} \{(1,0,4)\} \\ \{(0,1,0)\} \\ \{(3,3,1)\} \\ \{(0,0,1)\} \end{pmatrix} \cup \begin{pmatrix} \{(0,0.5,0)\} \\ \{(2,0,0)\} \\ \{(0,0,1)\} \\ \{(0,0,0)\} \end{pmatrix} = \begin{pmatrix} \{(1,0,4),(0,0.5,0)\} \\ \{(0,1,0),(2,0,0)\} \\ \{(3,3,1),(0,0,1)\} \\ \{(0,0,0),(0,0,1)\} \end{pmatrix} =_{\mathcal{U}^*} \begin{pmatrix} \{(1,0,4),(0,0.5,0)\} \\ \{(0,1,0),(2,0,0)\} \\ \{(3,3,1),(0,0,1)\} \\ \{(0,0,1)\} \end{pmatrix}
\end{aligned}$$

The last operation is reducing to the upper hull vertices and it doesn't change the dual function $\mathcal{R}(P_1)$.

We repeat this calculation for the next layer.

$$
\begin{aligned}
N_2 &= (A_2^+ \otimes N_1) \oplus (A_2^- \otimes P_1) \\
&= \left(\begin{bmatrix} 0.5 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \otimes \begin{pmatrix} \{(0,0.5,0)\} \\ \{(2,0,0)\} \\ \{(0,0,1)\} \\ \{(0,0,0)\} \end{pmatrix} \right) \oplus \left(\begin{bmatrix} 0 & 1 & 0.5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \otimes \begin{pmatrix} \{(1,0,4),(0,0.5,0)\} \\ \{(0,1,0),(2,0,0)\} \\ \{(3,3,1),(0,0,1)\} \\ \{(0,0,1)\} \end{pmatrix} \right) \\
&= \begin{pmatrix} 0.5\{(0,0.5,0)\} \\ \{(0,0,0)\} \end{pmatrix} \oplus \begin{pmatrix} 1\{(0,1,0),(2,0,0)\} \oplus 0.5\{(3,3,1),(0,0,1)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&= \begin{pmatrix} \{(0,0.25,0)\} \\ \{(0,0,0)\} \end{pmatrix} \oplus \begin{pmatrix} \{(0,1,0),(2,0,0)\} \oplus \{(1.5,1.5,0.5),(0,0,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&= \begin{pmatrix} \{(0,0.25,0)\} \\ \{(0,0,0)\} \end{pmatrix} \oplus \begin{pmatrix} \{(1.5,2.5,0.5),(0,1,0.5),(3.5,1.5,0.5),(2,0,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&= \begin{pmatrix} \{(1.5,2.75,0.5),(0,1.25,0.5),(3.5,1.75,0.5),(2,0.25,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
P_2 &= (A_2^+ \otimes P_1) \oplus (A_2^- \otimes N_1) \cup N_2 \\
&= \left(\begin{bmatrix} 0.5 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \otimes \begin{pmatrix} \{(1,0,4),(0,0.5,0)\} \\ \{(0,1,0),(2,0,0)\} \\ \{(3,3,1),(0,0,1)\} \\ \{(0,0,1)\} \end{pmatrix} \right) \oplus \left(\begin{bmatrix} 0 & 1 & 0.5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \otimes \begin{pmatrix} \{(0,0.5,0)\} \\ \{(2,0,0)\} \\ \{(0,0,1)\} \\ \{(0,0,0)\} \end{pmatrix} \right) \\
&\quad \cup \begin{pmatrix} \{(1.5,2.75,0.5),(0,1.25,0.5),(3.5,1.75,0.5),(2,0.25,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&= \begin{pmatrix} 0.5\{(1,0,4),(0,0.5,0)\} \oplus 2\{(0,0,1)\} \\ 1\{(0,0,1)\} \end{pmatrix} \oplus \begin{pmatrix} 1\{(2,0,0)\} \oplus 0.5\{(0,0,1)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&\quad \cup \begin{pmatrix} \{(1.5,2.75,0.5),(0,1.25,0.5),(3.5,1.75,0.5),(2,0.25,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&= \begin{pmatrix} \{(0.5,0,2),(0,0.25,0)\} \oplus \{(0,0,2)\} \\ \{(0,0,1)\} \end{pmatrix} \oplus \begin{pmatrix} \{(2,0,0)\} \oplus \{(0,0,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&\quad \cup \begin{pmatrix} \{(1.5,2.75,0.5),(0,1.25,0.5),(3.5,1.75,0.5),(2,0.25,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&= \begin{pmatrix} \{(0.5,0,4),(0,0.25,2)\} \\ \{(0,0,1)\} \end{pmatrix} \oplus \begin{pmatrix} \{(2,0,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&\quad \cup \begin{pmatrix} \{(1.5,2.75,0.5),(0,1.25,0.5),(3.5,1.75,0.5),(2,0.25,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&= \begin{pmatrix} \{(2.5,0,4.5),(2,0.25,2.5)\} \\ \{(0,0,1)\} \end{pmatrix} \\
&\quad \cup \begin{pmatrix} \{(1.5,2.75,0.5),(0,1.25,0.5),(3.5,1.75,0.5),(2,0.25,0.5)\} \\ \{(0,0,0)\} \end{pmatrix} \\
&= \begin{pmatrix} \{(2.5,0,4.5),(2,0.25,2.5),(1.5,2.75,0.5),(0,1.25,0.5),(3.5,1.75,0.5),(2,0.25,0.5)\} \\ \{(0,0,1),(0,0,0)\} \end{pmatrix}
\end{aligned}
$$

Now, to reduce the result to the upper hull vertices, we can note that

$$
\frac{1}{5}(0,1.25,0.5) + \frac{4}{5}(2.5,0,4.5) = (0,0.25,0.1) + (2,0,3.6) = (2,0.25,3.7) \succ (2,0.25,2.5),(2,0.25,0.5),
$$

so the two points of the right hand side can be dropped without changing the upper hull. This gives

$$
P_2 =_{\mathcal{U}^*} \begin{pmatrix} \{(2.5,0,4.5),(1.5,2.75,0.5),(0,1.25,0.5),(3.5,1.75,0.5)\} \\ \{(0,0,1)\} \end{pmatrix}.
$$

Finally, let's recover the representation as a DCPA function.

$$
\begin{aligned}
F_2(x,y) &= \left(\mathcal{R}(P_2) - \mathcal{R}(N_2)\right)(x,y) \\
&= \max\{1.5x + 2.75y + 0.5, 1.25y + 0.5, 3.5x + 1.75y + 0.5, 2.5x + 4.5\} \\
&- \max\{1.5x + 2.75y + 0.5, 1.25y + 0.5, 3.5x + 1.75y + 0.5, 2x + 0.25y + 0.5\}
\end{aligned}
$$

EXAMPLES OF APPLICATION OF PROPOSITIONS 20 AND COROLLARY 21

One-dimensional example Consider

$$
\begin{aligned}
&f_1(x) = -\tfrac{1}{2}x - \tfrac{3}{2} \qquad && f_2(x) = \tfrac{1}{2}x + \tfrac{1}{2} \qquad && f_3(x) = 2x + 1 \\
&g_1(x) = 0 && g_2(x) = 2x && g_3(x) = 3x - 1
\end{aligned}
$$

The DCPA function $F(x) = \max\{f_1(x), f_2(x), f_3(x)\} - \max\{g_1(x), g_2(x), g_3(x)\}$ is plotted in figure 9(a). It has 5 affine regions and 3 zeros.

It is represented by dual points as

$$
\begin{aligned}
\max\{f_1, f_2, f_3\} &= \mathcal{R}(P), \qquad & P &= \left\{ \begin{pmatrix} -\frac{1}{2} \\ -\frac{3}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\} \\
\max\{g_1, g_2, g_3\} &= \mathcal{R}(N), & N &= \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right\}
\end{aligned}
$$

Their upper convex hull $\mathcal{U}(P \cup N)$ is shown on figure 9(b). As predicted by proposition 20, the zero set of F is in bijection with 1-cells of $\mathcal{U}(P \cup N)$ which join a point of P with a point of N. This bijection is shown explicitly in table 9(c). The x-coordinates of zeros of F are given by negative slopes of these 1-cells.

The hull of the Minkowski sum $P \oplus N$ is shown in figure 9(d). In agreement with corollary 21, there are 5 vertices on $\mathcal{U}(P \cup N)$. The explicit bijections between the vertices of $\mathcal{U}(P \cup N)$ and affine regions of F, and between tangents at each vertex and points of the corresponding linear region, is given in the table 9(e).

Two-dimensional example Take

$$
\begin{aligned}
&f_1 = -x + y + 4 \qquad && f_2 = x + y - 2 \qquad && f_3 = -2x - y - 1 \\
&g_1 = 0 && g_2 = 2x - y + 2 && g_3 = -x + 2y + 2
\end{aligned}
$$

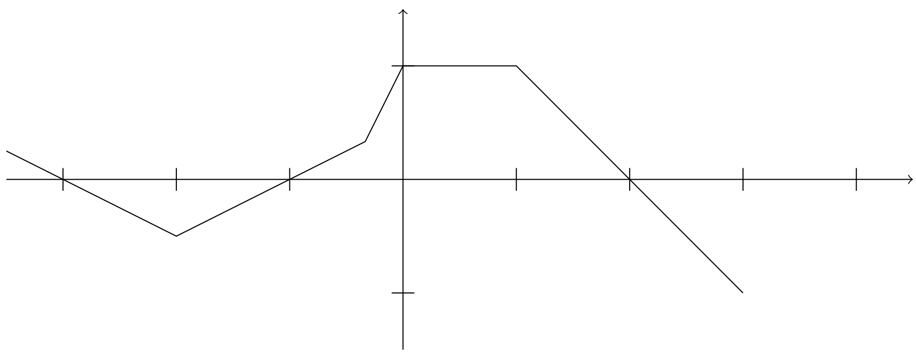
which correspond to dual points

$$
P = \left\{ \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix} \right\}, \qquad N = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix} \right\}
$$

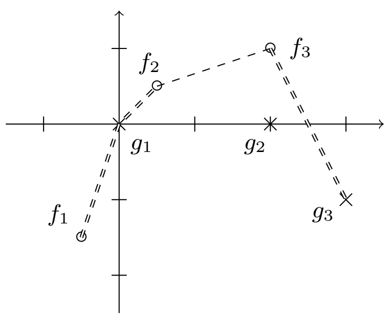
The function $F = \max\{f_1, f_2, f_3\} - \max\{g_1, g_2, g_3\}$ is shown on figure 10(a). There are 7 affine regions and 6 boundary pieces.

The configuration of dual points $P \cup N$ is shown on figure 10(b). The upper convex hull $\mathcal{U}(P \cup N)$ contains 4 faces, 8 edges and 5 vertices. As predicted by proposition 20, edges joining a point of P with a point of N correspond precisely to those affine regions of F which contain a boundary piece. Explicitly, these are $f_1 - g_2, f_1 - g_3, f_2 - g_2, f_2 - g_3, f_3 - g_2, f_3 - g_3$.

The Minkowski sum $P \oplus N$ is shown in figure 10(c). In agreement with corollary 21, 7 of the vertices lie on the upper convex hull. Explicitly, the functions $f_1 - g_1$ and $f_1 - g_2$ are the only ones which do not have a nonempty affine region, and the points $\mathcal{R}^{-1}(f_1) + \mathcal{R}^{-1}(g_1)$ and $\mathcal{R}^{-1}(f_2) + \mathcal{R}^{-1}(g_2)$ are the only ones which lie fully below the upper convex hull.



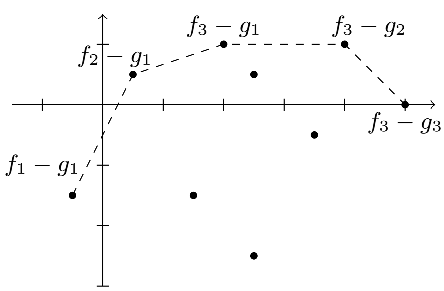
(a) Plot of F.



(b) Points of P (marked $\circ$) and N (marked $\times$) in the dual space. Dashed lines are the upper convex hull $\mathcal{U}(P \cup N)$. Double lines join a point of P with a point of N.

vertices	1-cell	zero of F
f_1, g_1	$3x$ over $[-\frac{1}{2}, 0]$	-3
f_2, g_2	x over $[0, \frac{1}{2}]$	-1
f_3, g_3	$-2x + 4$ over $[2, 3]$	2

(c) Correspondence between 1-cells of $U(P \cup N)$ and zeros of F. We represent 1-cell as a graph of a linear function over an interval.

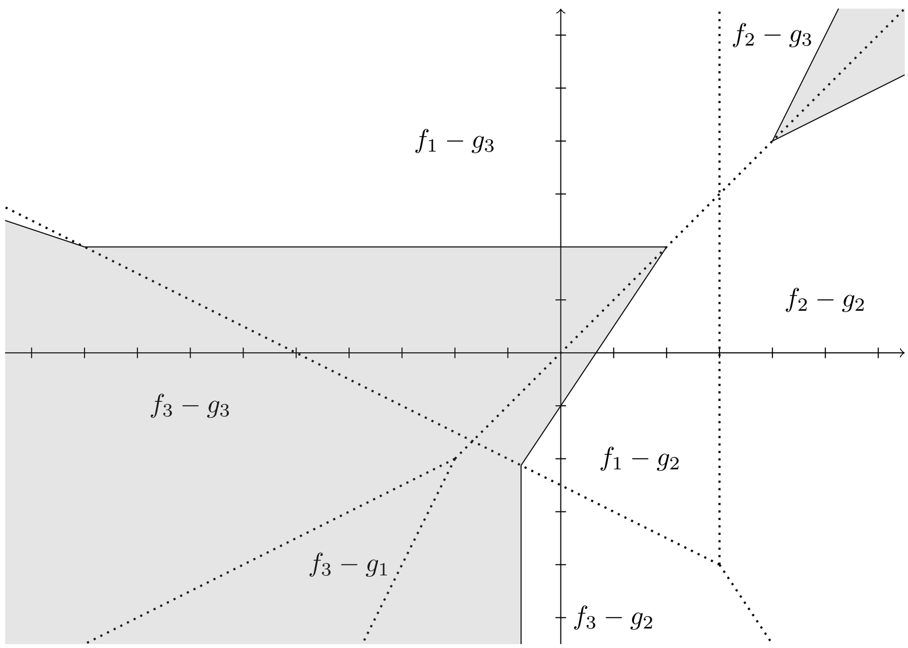


(d) Minkowski sum $P \oplus N$ in the dual space. Dashed lines represent the upper convex hull.

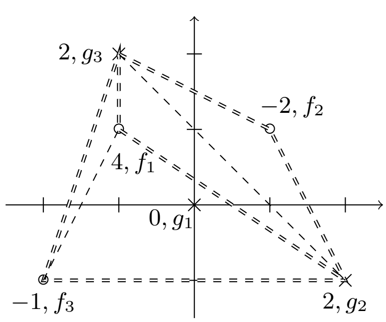
vertex	tangents	affine region
$f_1 - g_1$	$\{tx + \frac{t-3}{2}\}_{t \in [2,\infty)}$	$(-\infty, -2]$
$f_2 - g_1$	$\{tx + \frac{1-t}{2}\}_{t \in [\frac{1}{2}, 2]}$	$[-2, -\frac{1}{2}]$
$f_3 - g_1$	$\{tx + 1 - 2t\}_{t \in [0, \frac{1}{2}]}$	$[-\frac{1}{2}, 0]$
$f_3 - g_2$	$\{tx + 1 - 4t\}_{t \in [-1, 0]}$	$[0, 1]$
$f_3 - g_3$	$\{tx - 5t\}_{t \in (-\infty, -1]}$	$[1, \infty)$

(e) Tangents to $U(P \cup N)$ and the corresponding affine regions. We represent each tangent line as a graph of a linear function. They are parameterised by their slope t.

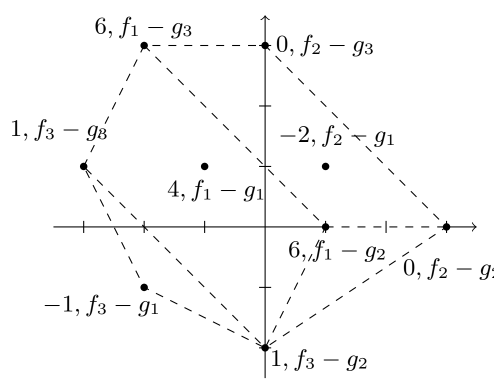
Figure 9: Illustration of the results on a one-dimensional example.



(a) The function F on the xy plane. Dotted lines mark the boundaries of affine regions. The annotation $f_i - g_j$ means that $F = f_i - g_j$ on the corresponding affine region. Solid lines indicate the zero set, and the shaded region contains arguments for which F is positive.



(b) Positions of $P \cup N$ in the dual space. The first number indicates the z-coordinate. Dashed lines are projections of edges (1-cells) of the upper convex hull. The point $\mathcal{R}^{-1}(g_1)$ is fully below the hull.



(c) Projection of the faces of $\mathcal{U}(P \oplus N)$ on the xy-plane. The first number indicates the z-coordinate; $f_i - g_j$ is a shorthand for the dual point $\mathcal{R}^{-1}(f_i) + \mathcal{R}^{-1}(g_j)$. Two points lie fully below the hull.

Figure 10: Illustration of the results on a two-dimensional example.