A REDUCING EMPIRICAL AND CERTIFIED ERRORS THROUGH PHYSICS-INFORMED ADVERSARIAL TRAINING

The goal of reducing the solution errors obtained by PINNs has been the research focus of several previous works Kim et al. (2021); Krishnapriyan et al. (2021); Shekarpaz et al. (2022). To observe the effects of one of these different training schemes on the verified correctness certification of PINNs, we consider Physics-informed Adversarial Training (PIAT) (Shekarpaz et al., 2022). The procedure consists in replacing the residual loss term from Raissi et al. (2019b) with an adversarial version inspired by Madry et al. (2017). While this procedure leads to improvements in the example PINNs from Shekarpaz et al. (2022) and using our own implementation in Burgers' equation, we were unable to stably train Schrödinger's equation using PIAT. Since Schrödinger's equation is not considered in Shekarpaz et al. (2022), we only show PIAT results for Burgers' equation.

We solve the inner optimization problem using 5 PGD steps (Madry et al., 2017), and for $\epsilon = 0.05$ and a step size of 1.25ϵ. To improve convergence, we warm start PIAT training using a standard training solution after 6,000 L-BFGS iterations. The results in Table 2 show that as expected PIAT improves both empirical and certified residual bounds.

Table 2: *PIAT on Burgers' equation*: Monte Carlo sampled maximum values (10^6 samples in 0.21s) and upper bounds computed using ∂-CROWN with N_b branchings for (1) initial conditions ($t = 0$, $x \in \mathcal{D}$, $N_b = 5k$), (2) boundary conditions ($t \in [0, T]$, $x = -1 \vee x = 1$, $N_b = 5k$), and (3) residual norm ($t \in [0, T]$, $x \in \mathcal{D}$, $N_b = 125k$), for a PINN trained using PIAT from Shekarpaz et al. (2022).

			MC - max	∂-CROWN - u_b (time [s])
PIAT Burgers (Shekarpaz et al., 2022)	(1)	$\lvert u_\theta(0, x) - u_0(x)\rvert^2$	$7.40 \cdot 10^{-6}$	$8.18 \cdot 10^{-6}$ (90.9)
	(2)	$\lvert u_\theta(t, -1)\rvert^2$	$2.31 \cdot 10^{-7}$	$3.32 \cdot 10^{-7}$ (49.4)
		$\lvert u_\theta(t, 1)\rvert^2$	$8.41 \cdot 10^{-8}$	$1.39 \cdot 10^{-7}$ (48.5)
	(3)	$\lvert f_\theta(\mathbf{x})\rvert^2$	$3.60 \cdot 10^{-3}$	$2.39 \cdot 10^{-2}$ (2.8×10^5)

Certification convergence in PIAT vs. standard training The regularization provided by adversarial training often leads to verification algorithms converging faster to tighter lower and upper bounds. We investigate whether this is the case with ∂-CROWN's greedy branching strategy by comparing the *relative convergence* (*i.e.*, the deviation between the upper bound and the empirical maximum, $|f_\theta|^U - \max_{\mathcal{D}'} |f_\theta|$) for the first $125k$ splits of PINNs trained in the standard and PIAT cases. The results presented in Figure 5 show that adversarial training leads to quicker convergence, requiring a lower number of branches to reach the same error when compared to standard. This suggests that our method, while already efficient, would benefit from smarter training strategies that lead to lower residual errors.

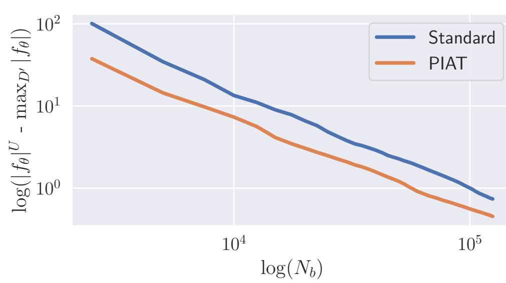


Figure 5: *Certification Convergence*: log-log plot of the relative convergence of ∂-CROWN certification for a standard trained PINN (in blue) and PIAT (in orange).

B ∂-CROWN VS. IBP (GOWAL ET AL., 2018; MIRMAN ET AL., 2018)

To the best of our knowledge, ∂-CROWN is the first framework designed to bound the errors of PINNs. However, for the sake of completeness of analysis we extend Interval Bound Propagation (IBP) (Gowal et al., 2018; Mirman et al., 2018) – known for its simplicity and trading-off bound tightness for speed – to this setting. Table 3 presents the performance of IBP and ∂-CROWN on the initial, boundary and residual errors for a fixed runtime limit in Burgers' equation. This is a fair comparison which takes into account the runtime/tightness trade-off of the two methods. We observe that ∂-CROWN is significantly more efficient than IBP, achieving bounds that are $165 - 1,566\times$

Table 3: *∂-CROWN vs. IBP*: comparison of the ∂-CROWN and Interval Bound Propagation (IBP) bounds for the initial, boundary and residual errors in Burgers' equation for fixed runtime limits.

	Runtime limit [s]	∂-CROWN u_b (N_b)	IBP u_b (N_b)
$\lvert u_\theta(0,x)\rvert^2$	150	2.63×10^{-6} (10^4)	4.12×10^{-3} (10^5)
$\lvert u_\theta(t,-1)\rvert^2$	100	6.63×10^{-7} (10^4)	1.23×10^{-5} (10^5)
$\lvert u_\theta(t,1)\rvert^2$	100	9.39×10^{-7} (10^4)	5.69×10^{-5} (10^5)
$\lvert f_\theta(x,t)\rvert^2$	10^4	13.0 (1.3×10^5)	2.78×10^3 (5×10^6)

Table 4: *Failure identification using residual bounds*: empirical analysis of the connection between the residual bounds obtained by ∂-CROWN and the maximum solution error computed with respect to a numerical solver, u, over a sampled dataset $\mathcal{D}'$. The range of the solution values over the samples in $\mathcal{D}'$ are included for ease of comparison.

	Residual ∂-CROWN u_b	Max solution error ($\max_{\mathcal{D}'} \lvert u_\theta - u\rvert$)	Solution range ($\min / \max_{\mathcal{D}'} u_\theta$)
Burgers	1.80×10^{-2}	3.78×10^{-3}	$[-1, 1]$
Schrödinger	7.67×10^{-4}	7.05×10^{-5}	$[1.82 \times 10^{-4}, 15.98]$
Allen-Cahn	10.76	0.86	$[-1, 1]$
Diffusion-Sorption	21.09	0.99	$[0, 1]$

tighter than the baseline given the same total runtime – this is despite the fact that IBP is able to branch more in total.

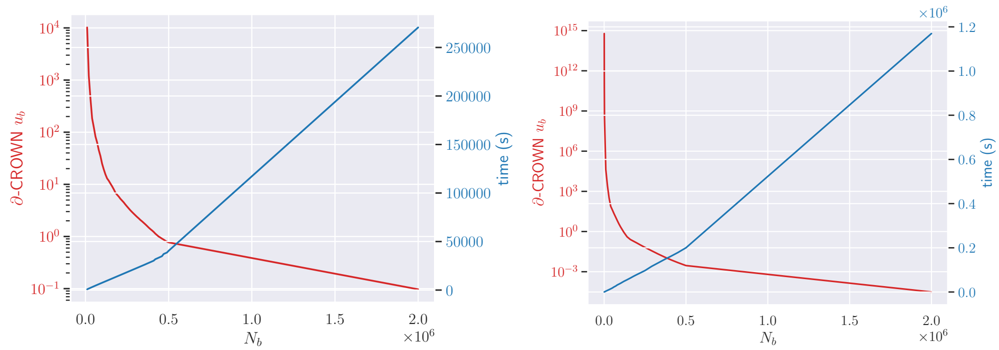
C ∂-CROWN FOR FAILURE IDENTIFICATION

In Section 6.2 we establish the empirical correlation between residual and solution errors for PINNs at different training stages (Figure 3). While comparing PINN errors for different PDEs is not easy due to residual scaling factors, note from Table 1 that the errors obtained for Burgers' and Schrödinger's equations are orders of magnitude lower than the ones for the Allen-Cahn and Diffusion-Sorption equations. Even with different residual tolerances, this would suggest the maximum solution error of the latter, harder to train PINNs should be higher.

Table 4 presents the residual bounds obtained using ∂-CROWN as well as the maximum solution error with respect to a numerical solver for each of the four PINNs studied, which empirically reinforces that correlation. E.g., Burgers' equation has a maximum solution error of 3.78×10^{-3}, which is significantly lower than the trained Allen-Cahn PINN at 0.86, as expected from the residual bounds of 1.80×10^{-2} and 10.76, respectively. This contextualizes the results of Table 1 and showcases our framework can identify weaker models.

D ABLATION ON N_b

We use $N_b = 2M$ for all the PINNs evaluated in this paper. A high number of branchings is required to obtain the tight bounds presented in Table 1. To justify that need, we have added plots of the variation of the obtained residual bound for Burgers' and Schrödinger' equations in Figure 6. Generally for both these PINNs we only get closer than one order of magnitude from the empirical estimates (considering the empirical MC sampled errors from Table 1) by using around $2M$ branches.



(a) (b)

Figure 6: *Ablation on* N_b: comparison of the residual error bounds ($|f_\theta|^2$) and runtime performance of our framework, ∂-CROWN on (a) Burgers' equation and (b) Schrödinger's equation.

E Proofs of partial derivative computations

E.1 Proof of Lemma 1: computing $\partial_{\mathbf{x}_i} u_\theta$

Let us now derive $\partial_{\mathbf{x}_i} u_\theta(\mathbf{x})$ for a given $i \in \{1, ..., n_0\}$. Starting backwards from the last layer and applying the chain rule we obtain:

$$\partial_{\mathbf{x}_i} u_\theta(\mathbf{x}) = \frac{\partial y^{(L)}}{\partial z^{(L-1)}} \cdot \frac{\partial z^{(L-1)}}{\partial z^{(L-2)}} \cdot ... \cdot \frac{\partial z^{(1)}}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial x_i}$$

Given that $\partial_{\mathbf{x}_i} x = \mathbf{e}_i$ and $\frac{\partial y^{(L)}}{\partial z^{(L-1)}} = \mathbf{W}^{(L)}$, all that's left to compute to obtain the full expression is $\frac{\partial z^{(k)}}{\partial z^{(k-1)}}$, $k \in \{L-1, ..., 1\}$. Note that, for simplicity of the expressions, $z^{(0)} = \mathbf{x}$. For every element $j \in \{1, ..., d_k\}$ of $z^{(k)}$ denoted by $z_j^{(k)}$, we have:

$$\frac{\partial z_j^{(k)}}{\partial z^{(k-1)}} = \sigma' \left(\mathbf{W}_{[j,:]}^{(k)} z^{(k-1)} + \mathbf{b}_j^{(k)} \right) \mathbf{W}_{[j,:]}^{(k)}$$

where $\mathbf{W}_{j,:}^{(k)}$ denotes the j-th row of $\mathbf{W}^{(k)}$, and $\mathbf{b}_j^{(k)}$ the j-th element of $\mathbf{b}$. Thus, the final expression can be obtained by stacking the columns of the previous expression to obtain the full Jacobian:

$$\frac{\partial z^{(k)}}{\partial z^{(k-1)}} = \text{diag} \left[\sigma' \left(\mathbf{W}^{(k)} z^{(k-1)} + \mathbf{b}^{(k)} \right) \right] \cdot \mathbf{W}^{(k)}$$

This concludes the proof.

E.2 Proof of Lemma 2: computing $\partial_{\mathbf{x}_i^2} u_\theta$

Given the result obtained in Appendix E.1, let us now derive $\partial_{\mathbf{x}_i^2} u_\theta(\mathbf{x})$ for a given $i \in \{1, ..., d_0\}$. Starting backwards from the last layer of $\partial_{\mathbf{x}_i} u_\theta$ and applying the chain rule we obtain:

$$\partial_{\mathbf{x}_i^2} u_\theta = \frac{\partial}{\partial x_i} \left(\frac{\partial y^{(L)}}{\partial z^{(L-1)}} \cdot \frac{\partial z^{(L-1)}}{\partial z^{(L-2)}} \cdot ... \cdot \frac{\partial z^{(1)}}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial x_i} \right) = \mathbf{W}^{(L)} \partial_{\mathbf{x}_i^2} z^{(L-1)}$$

Now the same can be applied to $\partial_{\mathbf{x}_i^2} z^{(L-1)}$, and in general to $\partial_{\mathbf{x}_i^2} z^{(k)}$ to obtain:

$$\partial_{\mathbf{x}_i^2} z^{(k)} = \frac{\partial}{\partial x_i} \left(\frac{\partial z^{(k)}}{\partial z^{(k-1)}} \partial_{\mathbf{x}_i} z^{(k-1)} \right) = \frac{\partial^2 z^{(k)}}{\partial x_i \partial z^{(k-1)}} \partial_{\mathbf{x}_i} z^{(k-1)} + \frac{\partial z^{(k)}}{\partial z^{(k-1)}} \partial_{\mathbf{x}_i^2} z^{(k-1)},$$

forming a recursion which can be taken until the first layer of $\partial_{\mathbf{x}_i} u_\theta$, *i.e.*,:

$$\partial_{\mathbf{x}_i^2} z^{(1)} = \frac{\partial}{\partial x_i} \left(\frac{\partial z^{(1)}}{\partial \mathbf{x}} \cdot \mathbf{e}_i \right) = \frac{\partial^2 z^{(1)}}{\partial x_i \partial \mathbf{x}} \cdot \mathbf{e}_i.$$

With the computation of $\partial_{\mathbf{x}_i} u_\theta$, both $\partial_{\mathbf{x}_i} z^{(k-1)}$ and $\frac{\partial z^{(k)}}{\partial z^{(k-1)}}$ are known. As such, the only missing pieces in the general recursion is the computation of $\frac{\partial^2 z^{(k)}}{\partial x_i \partial z^{(k-1)}}$. Recall from the previous section that $\frac{\partial z^{(k)}}{\partial z^{(k-1)}} = \text{diag}\left[\sigma'\left(\mathbf{W}^{(k)} z^{(k-1)} + \mathbf{b}^{(k)}\right)\right] \mathbf{W}^{(k)}$. As such:

$$\frac{\partial^2 z^{(k)}}{\partial x_i \partial z^{(k-1)}} = \frac{\partial}{\partial x_i}\left(\text{diag}\left[\sigma'\left(\mathbf{W}^{(k)} z^{(k-1)} + \mathbf{b}^{(k)}\right)\right] \mathbf{W}^{(k)}\right).$$

Following the element-wise reasoning from above, we have that:

$$\begin{aligned}\frac{\partial^2 z_j^{(k)}}{\partial x_i \partial z^{(k-1)}} &= \sigma''\left(\mathbf{W}_{j,:}^{(k)} z^{(k-1)} + \mathbf{b}_j^{(k)}\right) \frac{\partial}{\partial x_i}\left(\mathbf{W}_{j,:}^{(k)} z^{(k-1)} + \mathbf{b}_j^{(k)}\right) \mathbf{W}_{j,:}^{(k)} \\ &= \sigma''\left(\mathbf{W}_{j,:}^{(k)} z^{(k-1)} + \mathbf{b}_j^{(k)}\right)\left(\mathbf{W}_{j,:}^{(k)} \frac{\partial z^{(k-1)}}{\partial x_i}\right) \mathbf{W}_{j,:}^{(k)}\end{aligned}$$

Stacking as in the previous case, we obtain:

$$\frac{\partial^2 z^{(k)}}{\partial x_i \partial z^{(k-1)}} = \text{diag}\left[\sigma''\left(\mathbf{W}^{(k)} z^{(k-1)} + \mathbf{b}^{(k)}\right)\left(\mathbf{W}^{(k)} \partial_{\mathbf{x}_i} z^{(k-1)}\right)\right] \mathbf{W}^{(k)},$$

completing the derivation of $\partial_{\mathbf{x}_i^2} u_\theta(\mathbf{x})$.

E.3 Theorem 1: Formal Statement and Proof

Theorem 1 (∂-CROWN: linear lower and upper bounding $\partial_{\mathbf{x}_i} u_\theta$)**.** *For every $j \in \{1, \ldots, d_L\}$ there exist two functions $\partial_{\mathbf{x}_i} u_{\theta,j}^U$ and $\partial_{\mathbf{x}_i} u_{\theta,j}^L$ such that, $\forall \mathbf{x} \in \mathcal{C}$ it holds that $\partial_{\mathbf{x}_i} u_{\theta,j}^L \leq \partial_{\mathbf{x}_i} u_{\theta,j} \leq \partial_{\mathbf{x}_i} u_{\theta,j}^U$, with:*

$$\begin{aligned}\partial_{\mathbf{x}_i} u_{\theta,j}^U &= \phi_{0,j,i}^{(1),U} + \sum_{r=1}^{d_0} \phi_{1,j,r}^{(1),U} \mathbf{x} + \phi_{2,j,r}^{(1),U} \\ \partial_{\mathbf{x}_i} u_{\theta,j}^L &= \phi_{0,j,i}^{(1),L} + \sum_{r=1}^{d_0} \phi_{1,j,r}^{(1),L} \mathbf{x} + \phi_{2,j,r}^{(1),L}\end{aligned}$$

where for $p \in \{0, 1, 2\}$, $\phi_{p,j,r}^{(1),U}$ and $\phi_{p,j,r}^{(1),L}$ are functions of $\mathbf{W}^{(k)}$, $y^{(k),L}$, $y^{(k),U}$, $\mathbf{A}^{(k),L}$, $\mathbf{A}^{(k),U}$ $\mathbf{a}^{(k),L}$, and $\mathbf{a}^{(k),U}$, and can be computed using a recursive closed-form expression in $\mathcal{O}(L)$ time.

Proof: Assume that through the computation of the previous bounds on u_θ, the pre-activation layer outputs of u_θ, $y^{(k)}$, are lower and upper bounded by linear functions defined as $\mathbf{A}^{(k),L}\mathbf{x} + \mathbf{a}^{(k),L} \leq y^{(k)} \leq \mathbf{A}^{(k),U}\mathbf{x} + \mathbf{a}^{(k),U}$ and $y^{(k),L} \leq y^{(k)} \leq y^{(k),U}$ for $x \in \mathcal{C}$.

Take the upper and lower bound functions for $\partial_{\mathbf{x}_i} u_\theta$ as $\partial_{\mathbf{x}_i} u_\theta^U$ and $\partial_{\mathbf{x}_i} u_\theta^L$, respectively, and the upper and lower bound functions for $\partial_{\mathbf{x}_i} z^{(k)}$ as $\partial_{\mathbf{x}_i} z^{(k),U}$ and $\partial_{\mathbf{x}_i} z^{(k),L}$, respectively. For the sake of simplicity of notation, we define $\mathbf{B}^{(k),+} = \mathbb{I}\left(\mathbf{B}^{(k)} \geq 0\right) \odot \mathbf{B}^{(k)}$ and $\mathbf{B}^{(k),-} = \mathbb{I}\left(\mathbf{B}^{(k)} < 0\right) \odot \mathbf{B}^{(k)}$.

Working backwards from $\partial_{\mathbf{x}_i} u_\theta$, we apply the same idea from CROWN (Zhang et al., 2018):

$$\begin{aligned}\partial_{\mathbf{x}_i} u_\theta^U &= \mathbf{W}^{(L),+} \partial_{\mathbf{x}_i} z^{(L-1),U} + \mathbf{W}^{(L),-} \partial_{\mathbf{x}_i} z^{(L-1),L} \\ \partial_{\mathbf{x}_i} u_\theta^L &= \mathbf{W}^{(L),+} \partial_{\mathbf{x}_i} z^{(L-1),L} + \mathbf{W}^{(L),-} \partial_{\mathbf{x}_i} z^{(L-1),U}\end{aligned} \tag{7}$$

We continue to apply this backwards propagation to $\partial_{\mathbf{x}_i} z^{(L-1)}$ to obtain $\partial_{\mathbf{x}_i} z^{(L-1),U}$ and $\partial_{\mathbf{x}_i} z^{(L-1),L}$. Recall that $\partial_{\mathbf{x}_i} z^{(k)} = \partial_{z^{(k-1)}} z^{(k)} \partial_{\mathbf{x}_i} z^{(k-1)}$, that is, for $j \in \{1, \ldots, d_k\}$ we have $\partial_{\mathbf{x}_i} z_j^{(k)} = \partial_{z^{(k-1)}} z_{j,:}^{(k)} \partial_{\mathbf{x}_i} z^{(k-1)} = \sum_{n=1}^{d_{k-1}} \partial_{z^{(k-1)}} z_{j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1)}$.

We resolve the bilinear dependencies of each $\partial_{\mathbf{x}_i} z_j^{(k)}$ by relaxing it using a convex combination of the upper and lower bounds obtained by the McCormick envelopes of the product. Assuming that $\partial_{z^{(k-1)}} z_{j,n}^{(k),L} \leq \partial_{z^{(k-1)}} z_{j,n}^{(k)} \leq \partial_{z^{(k-1)}} z_{j,n}^{(k),U}$ and $\partial_{\mathbf{x}_i} z_n^{(k-1),L} \leq \partial_{\mathbf{x}_i} z_n^{(k-1)} \leq \partial_{\mathbf{x}_i} z_n^{(k-1),U}$, we have that:

$$
\begin{aligned}
\partial_{\mathbf{x}_i} z_j^{(k)} \leq \partial_{\mathbf{x}_i} z_j^{(k),U} &= \sum_{n=1}^{d_{k-1}} \alpha_{0,j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1)} + \alpha_{1,j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k)} + \alpha_{2,j,n}^{(k)} \\
\partial_{\mathbf{x}_i} z_j^{(k)} \geq \partial_{\mathbf{x}_i} z_j^{(k),L} &= \sum_{n=1}^{d_{k-1}} \beta_{0,j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1)} + \beta_{1,j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k)} + \beta_{2,j,n}^{(k)},
\end{aligned} \tag{8}
$$

for:

$$
\begin{aligned}
\alpha_{0,j,n}^{(k)} &= \eta_{j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k),U} + \left(1 - \eta_{j,n}^{(k)}\right) \partial_{z^{(k-1)}} z_{j,n}^{(k),L} \\
\alpha_{1,j,n}^{(k)} &= \eta_{j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1),L} + \left(1 - \eta_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i} z_n^{(k-1),U} \\
\alpha_{2,j,n}^{(k)} &= -\eta_{j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k),U} \partial_{\mathbf{x}_i} z_n^{(k-1),L} - \left(1 - \eta_{j,n}^{(k)}\right) \partial_{z^{(k-1)}} z_{j,n}^{(k),L} \partial_{\mathbf{x}_i} z_n^{(k-1),U} \\
\beta_{0,j,n}^{(k)} &= \zeta_{j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k),L} + \left(1 - \zeta_{j,n}^{(k)}\right) \partial_{z^{(k-1)}} z_{j,n}^{(k),U} \\
\beta_{1,j,n}^{(k)} &= \zeta_{j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1),L} + \left(1 - \zeta_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i} z_n^{(k-1),U} \\
\beta_{2,j,n}^{(k)} &= -\zeta_{j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k),L} \partial_{\mathbf{x}_i} z_n^{(k-1),L} - \left(1 - \zeta_{j,n}^{(k)}\right) \partial_{z^{(k-1)}} z_{j,n}^{(k),U} \partial_{\mathbf{x}_i} z_n^{(k-1),U},
\end{aligned}
$$

where $\eta_{j,n}^{(k)}$ and $\zeta_{j,n}^{(k)}$ are convex coefficients that can be set as hyperparameters, or optimized for as in α-CROWN (Xu et al., 2020b).

To continue the backward propagation, we now need to bound the components of $\partial_{z^{(k-1)}} z^{(k)}$. Recall from Lemma 1 that $\partial_{z^{(k-1)}} z^{(k)} = \text{diag}\left[\sigma'\left(y^{(k-1)}\right)\right] \mathbf{W}^{(k)}$, and $\partial_{z^{(k-1)}} z_{j,:}^{(k)} = \sigma'\left(y_j^{(k-1)}\right) \mathbf{W}_{j,:}^{(k)}$ for $j \in \{1, \ldots, d_k\}$.

Since $y_j^{(k),L} \leq y_j^{(k)} \leq y_j^{(k),U}$, we can obtain a linear upper and lower bound relaxation for $\sigma'\left(y_j^{(k)}\right)$, such that $\gamma_j^{(k),L}\left(y_j^{(k)} + \delta_j^{(k),L}\right) \leq \sigma'\left(y_j^{(k)}\right) \leq \gamma_j^{(k),U}\left(y_j^{(k)} + \delta_j^{(k),U}\right)$. With this, we can proceed to bound $\partial_{z^{(k-1)}} z_{j,:}^{(k)}$ as:

$$
\begin{aligned}
\partial_{z^{(k-1)}} z_{j,:}^{(k)} &\leq \underbrace{\left(\gamma_j^{(k),U} \mathbf{W}_{j,:}^{(k),+} + \gamma_j^{(k),L} \mathbf{W}_{j,:}^{(k),-}\right)}_{\iota_{0,j,:}^{(k)}} y_j^{(k)} + \underbrace{\left(\gamma_j^{(k),U} \delta_j^{(k),U} \mathbf{W}_{j,:}^{(k),+} + \gamma_j^{(k),L} \delta_j^{(k),L} \mathbf{W}_{j,:}^{(k),-}\right)}_{\iota_{1,j,:}^{(k)}} \\
\partial_{z^{(k-1)}} z_{j,:}^{(k)} &\geq \underbrace{\left(\gamma_j^{(k),L} \mathbf{W}_{j,:}^{(k),+} + \gamma_j^{(k),U} \mathbf{W}_{j,:}^{(k),-}\right)}_{\lambda_{0,j,:}^{(k)}} y_j^{(k)} + \underbrace{\left(\gamma_j^{(k),L} \delta_j^{(k),L} \mathbf{W}_{j,:}^{(k),+} + \gamma_j^{(k),U} \delta_j^{(k),U} \mathbf{W}_{j,:}^{(k),-}\right)}_{\lambda_{1,j,:}^{(k)}}
\end{aligned} \tag{9}
$$

At this point, one could continue the back-substitution process using the bounds from CROWN (Zhang et al., 2018). However, for the sake of efficiency, we use instead the pre-computed inequalities from propagating bounds through u_θ: $\mathbf{A}^{(k),U}\mathbf{x} + \mathbf{a}^{(k),U} \leq y^{(k)} \leq \mathbf{A}^{(k),L}\mathbf{x} + \mathbf{a}^{(k),L}$. Substituting this in Equation 9, we obtain:

$$
\begin{aligned}
\partial_{z^{(k-1)}} z_{j,:}^{(k),U} &= \underbrace{\left(\iota_{0,j,:}^{(k),+} \mathbf{A}_{j,:}^{(k),U} + \iota_{0,j,:}^{(k),-} \mathbf{A}_{j,:}^{(k),L}\right)}_{\iota_{2,j,:}^{(k)}} \mathbf{x} + \underbrace{\iota_{0,j,:}^{(k),+} \mathbf{a}_j^{(k),U} + \iota_{0,j,:}^{(k),-} \mathbf{a}_j^{(k),L} + \iota_{1,j,:}^{(k)}}_{\iota_{3,j,:}^{(k)}} \\
\partial_{z^{(k-1)}} z_{j,:}^{(k),L} &= \underbrace{\left(\lambda_{0,j,:}^{(k),+} \mathbf{A}_{j,:}^{(k),L} + \lambda_{0,j,:}^{(k),-} \mathbf{A}_{j,:}^{(k),U}\right)}_{\lambda_{2,j,:}^{(k)}} \mathbf{x} + \underbrace{\lambda_{0,j,:}^{(k),+} \mathbf{a}_j^{(k),L} + \lambda_{0,j,:}^{(k),-} \mathbf{a}_j^{(k),U} + \lambda_{1,j,:}^{(k)}}_{\lambda_{3,j,:}^{(k)}}
\end{aligned} \tag{10}
$$

In practice, we can use Equation 10 to compute the required $\partial_{z^{(k-1)}} z_{j,n}^{(k),L}$ and $\partial_{z^{(k-1)}} z_{j,n}^{(k),U}$ for the McCormick relaxation that leads to Equation 8. By back-substituting the result of Equation 10 in

Equation 8, we obtain an expression for the upper and lower bounds on $\partial_{\mathbf{x}_i} z_j^{(k)}$ that only depends on $\partial_{\mathbf{x}_i} z^{(k-1)}$ and $\mathbf{x}$:

$$\begin{aligned}\partial_{\mathbf{x}_i} z_j^{(k),U} &= \sum_{n=1}^{d_{k-1}} \alpha_{0,j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1)} + \alpha_{3,j,n}^{(k)} \mathbf{x} + \alpha_{4,j,n}^{(k)} \\ \partial_{\mathbf{x}_i} z_j^{(k),L} &= \sum_{n=1}^{d_{k-1}} \beta_{0,j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1)} + \beta_{3,j,n}^{(k)} \mathbf{x} + \beta_{4,j,n}^{(k)},\end{aligned} \tag{11}$$

where:

$$\alpha_{3,j,n}^{(k)} = \alpha_{1,j,n}^{(k),+} \iota_{2,j,n}^{(k)} + \alpha_{1,j,n}^{(k),-} \lambda_{2,j,n}^{(k)}, \qquad \alpha_{4,j,n}^{(k)} = \alpha_{1,j,n}^{(k),+} \iota_{3,j,n}^{(k)} + \alpha_{1,j,n}^{(k),-} \lambda_{3,j,n}^{(k)} + \alpha_{2,j,n}^{(k)}$$
$$\beta_{3,j,n}^{(k)} = \beta_{1,j,n}^{(k),+} \lambda_{2,j,n}^{(k)} + \beta_{1,j,n}^{(k),-} \iota_{2,j,n}^{(k)}, \qquad \beta_{4,j,n}^{(k)} = \beta_{1,j,n}^{(k),+} \lambda_{3,j,n}^{(k)} + \beta_{1,j,n}^{(k),-} \iota_{3,j,n}^{(k)} + \alpha_{2,j,n}^{(k)}$$

Given Equation 11, we now have a recursive expression for each of the blocks that compose the computation of $\partial_{\mathbf{x}_i} u_\theta$, which allows us to obtain a closed form expression for $\partial_{\mathbf{x}_i} u_\theta^U$ and $\partial_{\mathbf{x}_i} u_\theta^L$ by applying recursive back-substitution starting with Equation 7. Let us begin by performing back-substitution to the result in Equation 11 for layer $L-1$:

$$\partial_{\mathbf{x}_i} z_j^{(L-1),U} = \sum_{n=1}^{d_{L-2}} \alpha_{0,j,n}^{(L-1)} \partial_{\mathbf{x}_i} z_n^{(L-2)} + \alpha_{3,j,n}^{(L-1)} \mathbf{x} + \alpha_{4,j,n}^{(L-1)} \tag{12}$$

$$= \sum_{n=1}^{d_{L-2}} \alpha_{0,j,n}^{(L-1)} \left(\sum_{r=1}^{d_{L-3}} \mu_{0,n,r}^{(L-2)} \partial_{\mathbf{x}_i} z_r^{(L-3)} + \mu_{3,n,r}^{(L-2)} \mathbf{x} + \mu_{4,n,r}^{(L-2)} \right) + \alpha_{3,j,n}^{(L-1)} \mathbf{x} + \alpha_{4,j,n}^{(L-1)} \tag{13}$$

$$= \sum_{n=1}^{d_{L-2}} \alpha_{0,j,n}^{(L-1)} \left(\sum_{r=1}^{d_{L-3}} \mu_{0,n,r}^{(L-2)} \partial_{\mathbf{x}_i} z_r^{(L-3)} \right) + \alpha_{0,j,n}^{(L-1)} \left(\sum_{r=1}^{d_{L-3}} \mu_{3,n,r}^{(L-2)} \mathbf{x} + \mu_{4,n,r}^{(L-2)} \right) + \alpha_{3,j,n}^{(L-1)} \mathbf{x} + \alpha_{4,j,n}^{(L-1)} \tag{14}$$

$$= \sum_{r=1}^{d_{L-3}} \left(\sum_{n=1}^{d_{L-2}} \alpha_{0,j,n}^{(L-1)} \mu_{0,n,r}^{(L-2)} \right) \partial_{\mathbf{x}_i} z_r^{(L-3)} + \tag{15}$$

$$+ \left(\sum_{n=1}^{d_{L-2}} \alpha_{0,j,n}^{(L-1)} \left(\mu_{3,n,r}^{(L-2)} \mathbf{x} + \mu_{4,n,r}^{(L-2)} \right) + \frac{1}{d_{L-3}} \left(\alpha_{3,j,n}^{(L-1)} \mathbf{x} + \alpha_{4,j,n}^{(L-1)} \right) \right) \tag{16}$$

$$= \sum_{r=1}^{d_{L-3}} \rho_{0,j,r}^{(L-2)} \partial_{\mathbf{x}_i} z_r^{(L-3)} + \left(\sum_{n=1}^{d_{L-2}} \alpha_{0,j,n}^{(L-1)} \mu_{3,n,r}^{(L-2)} + \frac{1}{d_{L-3}} \alpha_{3,j,n}^{(L-1)} \right) \mathbf{x} + \tag{17}$$

$$+ \left(\sum_{n=1}^{d_{L-1}} \alpha_{0,j,n}^{(L-1)} \mu_{4,n,r}^{(L-2)} + \frac{1}{d_{L-3}} \alpha_{4,j,n}^{(L-1)} \right) \tag{18}$$

$$= \sum_{r=1}^{d_{L-3}} \rho_{0,j,r}^{(L-2)} \partial_{\mathbf{x}_i} z_r^{(L-3)} + \rho_{1,j,r}^{(L-2)} \mathbf{x} + \rho_{2,j,r}^{(L-2)}, \tag{19}$$

where:

$$\begin{aligned}\rho_{0,j,r}^{(L-2)} &= \sum_{n=1}^{d_{L-2}} \alpha_{0,j,n}^{(L-1)} \mu_{0,n,r}^{(L-2)} \\ \rho_{1,j,r}^{(L-2)} &= \sum_{n=1}^{d_{L-2}} \alpha_{0,j,n}^{(L-1)} \mu_{3,n,r}^{(L-2)} + \frac{1}{d_{L-2}} \alpha_{3,j,n}^{(L-1)} \\ \rho_{2,j,r}^{(L-2)} &= \sum_{n=1}^{d_{L-2}} \alpha_{0,j,n}^{(L-1)} \mu_{4,n,r}^{(L-2)} + \frac{1}{d_{L-2}} \alpha_{4,j,n}^{(L-1)},\end{aligned}$$

and:

$$\mu_{p,n,:}^{(L-2)} = \begin{cases} \alpha_{p,n,:}^{(L-2)} & \text{if } \alpha_{0,j,n}^{(L-1)} \geq 0 \\ \beta_{p,n,:}^{(L-2)} & \text{if } \alpha_{0,j,n}^{(L-1)} < 0 \end{cases} \quad , p \in \{0, 3, 4\}$$

As in CROWN (Zhang et al., 2018), given we have put Equation 19 in the same form as Equation 12, we can now apply this argument recursively using the $\rho^{(k)}$ and $\mu^{(k)}$ coefficients to obtain:

$$\partial_{\mathbf{x}_i} z_j^{(L-1),U} = \rho_{0,j,i}^{(1)} + \sum_{r=1}^{d_0} \rho_{1,j,r}^{(1)} \mathbf{x} + \rho_{2,j,r}^{(1)},$$

where:

$$\begin{aligned} \rho_{0,j,r}^{(k-1)} &= \begin{cases} \alpha_{0,j,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \rho_{0,j,n}^{(k)} \mu_{0,n,r}^{(k-1)} & \text{if } k \in \{2, \dots, L-1\} \end{cases} \\ \rho_{1,j,r}^{(k-1)} &= \begin{cases} \alpha_{3,j,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \rho_{0,j,n}^{(k)} \mu_{3,n,r}^{(k-1)} + \frac{1}{d_{k-2}} \rho_{1,j,n}^{(k)} & \text{if } k \in \{2, \dots, L-1\} \end{cases} \\ \rho_{2,j,r}^{(k-1)} &= \begin{cases} \alpha_{4,j,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \rho_{0,j,n}^{(k)} \mu_{4,n,r}^{(k-1)} + \frac{1}{d_{k-2}} \rho_{2,j,n}^{(k)} & \text{if } k \in \{2, \dots, L-1\} \end{cases} \end{aligned} \quad ,$$

and:

$$\mu_{p,n,:}^{(k-1)} = \begin{cases} \alpha_{p,n,:}^{(k-1)} & \text{if } \rho_{0,j,n}^{(k)} \geq 0 \\ \beta_{p,n,:}^{(k-1)} & \text{if } \rho_{0,j,n}^{(k)} < 0 \end{cases} \quad , p \in \{0, 3, 4\}$$

And following the same recursive argument:

$$\partial_{\mathbf{x}_i} z_j^{(L-1),L} = \tau_{0,j,i}^{(1)} + \sum_{r=1}^{d_0} \tau_{1,j,r}^{(1)} \mathbf{x} + \tau_{2,j,r}^{(1)},$$

where:

$$\begin{aligned} \tau_{0,j,r}^{(k-1)} &= \begin{cases} \beta_{0,j,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \tau_{0,j,n}^{(k)} \omega_{0,n,r}^{(k-1)} & \text{if } k \in \{2, \dots, L-1\} \end{cases} \\ \tau_{1,j,r}^{(k-1)} &= \begin{cases} \beta_{3,j,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \tau_{0,j,n}^{(k)} \omega_{3,n,r}^{(k-1)} + \frac{1}{d_{k-2}} \tau_{1,j,n}^{(k)} & \text{if } k \in \{2, \dots, L-1\} \end{cases} \\ \tau_{2,j,r}^{(k-1)} &= \begin{cases} \beta_{4,j,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \tau_{0,j,n}^{(k)} \omega_{4,n,r}^{(k-1)} + \frac{1}{d_{k-2}} \tau_{2,j,n}^{(k)} & \text{if } k \in \{2, \dots, L-1\} \end{cases} \end{aligned} \quad ,$$

and:

$$\omega_{p,n,:}^{(k-1)} = \begin{cases} \beta_{p,n,:}^{(k-1)} & \text{if } \tau_{0,j,n}^{(k)} \geq 0 \\ \alpha_{p,n,:}^{(k-1)} & \text{if } \tau_{0,j,n}^{(k)} < 0 \end{cases} \quad , p \in \{0, 3, 4\}$$

With these expressions, we can compute the required $\partial_{\mathbf{x}_i} z_n^{(k-1),L}$ and $\partial_{\mathbf{x}_i} z_n^{(k-1),U}$ which we assumed to be known to derive Equation 8.

Finally, by back-propagating the bounds starting from Equation 7, we get:

$$\begin{aligned}
\partial_{\mathbf{x}_i} u^U_{\theta,j} &= \sum_{n=1}^{d_{L-1}} \mathbf{W}^{(L),+}_{j,n} \left(\sum_{r=1}^{d_{L-2}} \alpha^{(L-1)}_{0,n,r} \partial_{\mathbf{x}_i} z^{(L-2)}_{[r]} + \alpha^{(L-1)}_{3,n,r} \mathbf{x} + \alpha^{(L-1)}_{4,n,r} \right) + \\
&\quad + \mathbf{W}^{(L),-}_{j,n} \left(\sum_{r=1}^{d_{L-2}} \beta^{(L-1)}_{0,n,r} \partial_{\mathbf{x}_i} z^{(L-2)}_{[r]} + \beta^{(L-1)}_{3,n,r} \mathbf{x} + \beta^{(L-1)}_{4,n,r} \right) \\
&= \sum_{r=1}^{d_{L-2}} \left(\sum_{n=1}^{d_{L-1}} \mathbf{W}^{(L),+}_{j,n} \alpha^{(L-1)}_{0,n,r} + \mathbf{W}^{(L),-}_{j,n} \beta^{(L-1)}_{0,n,r} \right) \partial_{\mathbf{x}_i} z^{(L-2)}_{[r]} + \\
&\quad + \left(\sum_{n=1}^{d_{L-1}} \mathbf{W}^{(L),+}_{j,n} \alpha^{(L-1)}_{3,n,r} + \mathbf{W}^{(L),-}_{j,n} \beta^{(L-1)}_{3,n,r} \right) \mathbf{x} + \left(\sum_{n=1}^{d_{L-1}} \mathbf{W}^{(L),+}_{j,n} \alpha^{(L-1)}_{4,n,r} + \mathbf{W}^{(L),-}_{j,n} \beta^{(L-1)}_{4,n,r} \right) \\
&= \sum_{r=1}^{d_{L-2}} \phi^{(L-1),U}_{0,j,r} \partial_{\mathbf{x}_i} z^{(L-2),U}_{n} + \phi^{(L-1),U}_{1,j,r} \mathbf{x} + \phi^{(L-1),U}_{2,j,r} ,
\end{aligned}$$

where:

$$\begin{aligned}
\phi^{(L-1),U}_{0,j,r} &= \sum_{n=1}^{d_{L-1}} \mathbf{W}^{(L),+}_{j,n} \alpha^{(L-1)}_{0,n,r} + \mathbf{W}^{(L),-}_{j,n} \beta^{(L-1)}_{0,n,r} \\
\phi^{(L-1),U}_{1,j,r} &= \sum_{n=1}^{d_{L-1}} \mathbf{W}^{(L),+}_{j,n} \alpha^{(L-1)}_{3,n,r} + \mathbf{W}^{(L),-}_{j,n} \beta^{(L-1)}_{3,n,r} \\
\phi^{(L-1),U}_{2,j,r} &= \sum_{n=1}^{d_{L-1}} \mathbf{W}^{(L),+}_{j,n} \alpha^{(L-1)}_{4,n,r} + \mathbf{W}^{(L),-}_{j,n} \beta^{(L-1)}_{4,n,r} .
\end{aligned}$$

From this, using the same back-propagation logic as in the derivations of $\partial_{\mathbf{x}_i} z^{(k-1),L}_n$ and $\partial_{\mathbf{x}_i} z^{(k-1),U}_n$, we can obtain:

$$\partial_{\mathbf{x}_i} u^U_{\theta,j} = \phi^{(1),U}_{0,j,i} + \sum_{r=1}^{d_0} \phi^{(1),U}_{1,j,r} \mathbf{x} + \phi^{(1),U}_{2,j,r} , \tag{20}$$

where:

$$\begin{aligned}
\phi^{(k-1),U}_{0,j,r} &= \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}^{(k),+}_{j,n} \alpha^{(k-1)}_{0,n,r} + \mathbf{W}^{(k),-}_{j,n} \beta^{(k-1)}_{0,n,r} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \phi^{(k),U}_{0,j,n} \upsilon^{(k-1)}_{0,n,r} & \text{if } k \in \{2, \ldots, L-1\} \end{cases} \\
\phi^{(k-1),U}_{1,j,r} &= \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}^{(k),+}_{j,n} \alpha^{(k-1)}_{3,n,r} + \mathbf{W}^{(k),-}_{j,n} \beta^{(k-1)}_{3,n,r} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \phi^{(k),U}_{0,j,n} \upsilon^{(k-1)}_{3,n,r} + \frac{1}{d_{k-2}} \phi^{(k),U}_{1,j,n} & \text{if } k \in \{2, \ldots, L-1\} \end{cases} \\
\phi^{(k-1),U}_{2,j,r} &= \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}^{(k),+}_{j,n} \alpha^{(k-1)}_{4,n,r} + \mathbf{W}^{(k),-}_{j,n} \beta^{(k-1)}_{4,n,r} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \phi^{(k),U}_{0,j,n} \upsilon^{(k-1)}_{4,n,r} + \frac{1}{d_{k-2}} \phi^{(k),U}_{2,j,n} & \text{if } k \in \{2, \ldots, L-1\} \end{cases} ,
\end{aligned}$$

and:

$$\upsilon^{(k-1)}_{p,n,:} = \begin{cases} \alpha^{(k-1)}_{p,n,:} & \text{if } \phi^{(k),U}_{0,j,n} \geq 0 \\ \beta^{(k-1)}_{p,n,:} & \text{if } \phi^{(k),U}_{0,j,n} < 0 \end{cases} , p \in \{0, 3, 4\}$$

And similarly for the lower bound:

$$\partial_{\mathbf{x}_i} u^L_{\theta,j} = \phi^{(1),L}_{0,j,i} + \sum_{r=1}^{d_0} \phi^{(1),L}_{1,j,r} \mathbf{x} + \phi^{(1),L}_{2,j,r} , \tag{21}$$

where:

$$\phi_{0,j,r}^{(k-1),L} = \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \beta_{0,n,r}^{(k-1)} + \mathbf{W}_{j,n}^{(k),-} \alpha_{0,n,r}^{(k-1)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \phi_{0,j,n}^{(k),L} \chi_{0,n,r}^{(k-1)} & \text{if } k \in \{2, \dots, L-1\} \end{cases}$$

$$\phi_{1,j,r}^{(k-1),L} = \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \beta_{3,n,r}^{(k-1)} + \mathbf{W}_{j,n}^{(k),-} \alpha_{3,n,r}^{(k-1)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \phi_{0,j,n}^{(k),L} \chi_{3,n,r}^{(k-1)} + \frac{1}{d_{k-2}} \phi_{1,j,n}^{(k),L} & \text{if } k \in \{2, \dots, L-1\} \end{cases}$$

$$\phi_{2,j,r}^{(k-1),L} = \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \beta_{4,n,r}^{(k-1)} + \mathbf{W}_{j,n}^{(k),-} \alpha_{4,n,r}^{(k-1)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \phi_{0,j,n}^{(k),L} \chi_{4,n,r}^{(k-1)} + \frac{1}{d_{k-2}} \phi_{2,j,n}^{(k),L} & \text{if } k \in \{2, \dots, L-1\} \end{cases},$$

and:

$$\chi_{p,n,:}^{(k-1)} = \begin{cases} \beta_{p,n,:}^{(k-1)} & \text{if } \phi_{0,j,n}^{(k),L} \geq 0 \\ \alpha_{p,n,:}^{(k-1)} & \text{if } \phi_{0,j,n}^{(k),L} < 0 \end{cases}, p \in \{0, 3, 4\}.$$

E.4 Theorem 2 Formal Statement and Proof

Theorem 2 (∂-CROWN: linear lower and upper bounding $\partial_{\mathbf{x}_i^2} u_\theta$)**.** *Assume that through a previous computation of bounds on $\partial_{\mathbf{x}_i} u_\theta$, the components of that network required for $\partial_{\mathbf{x}_i^2} u_\theta$, i.e., $\partial_{\mathbf{x}_i} z^{(k-1)}$ and $\partial_{z^{(k-1)}} z^{(k)}$, are lower and upper bounded by linear functions. In particular, $\mathbf{C}^{(k),L}\mathbf{x} + \mathbf{c}^{(k),L} \leq \partial_{\mathbf{x}_i} z^{(k-1)} \leq \mathbf{C}^{(k),U}\mathbf{x} + \mathbf{c}^{(k),U}$ and $\mathbf{D}^{(k),L}\mathbf{x} + \mathbf{d}^{(k),L} \leq \partial_{z^{(k-1)}} z^{(k)} \leq \mathbf{D}^{(k),U}\mathbf{x} + \mathbf{d}^{(k),U}$.*

For every $j \in \{1, \dots, d_L\}$ there exist two functions $\partial_{\mathbf{x}_i^2} u_{\theta,j}^U$ and $\partial_{\mathbf{x}_i^2} u_{\theta,j}^L$ such that, $\forall \mathbf{x} \in \mathcal{C}$ it holds that $\partial_{\mathbf{x}_i^2} u_{\theta,j}^L \leq \partial_{\mathbf{x}_i^2} u_{\theta,j} \leq \partial_{\mathbf{x}_i^2} u_{\theta,j}^U$. These functions can be written as:

$$\partial_{\mathbf{x}_i^2} u_{\theta,j}^U = \psi_{0,j,i}^{(1),U} + \sum_{r=1}^{d_0} \psi_{1,j,r}^{(1),U} \mathbf{x} + \psi_{2,j,r}^{(1),U}$$

$$\partial_{\mathbf{x}_i^2} u_{\theta,j}^L = \psi_{0,j,i}^{(1),L} + \sum_{r=1}^{d_0} \psi_{1,j,r}^{(1),L} \mathbf{x} + \psi_{2,j,r}^{(1),L}$$

where for $p \in \{0, 1, 2\}$, $\psi_{p,j,r}^{(1),U}$ and $\psi_{p,j,r}^{(1),L}$ are functions of $\mathbf{W}^{(k)}$, $y^{(k),L}$, $y^{(k),U}$, $\mathbf{A}^{(k),L}$, $\mathbf{A}^{(k),U}$ $\mathbf{a}^{(k),L}$, $\mathbf{a}^{(k),U}$, $\mathbf{C}^{(k),L}$, $\mathbf{C}^{(k),U}$ $\mathbf{c}^{(k),L}$, $\mathbf{c}^{(k),U}$, $\mathbf{D}^{(k),L}$, $\mathbf{D}^{(k),U}$ $\mathbf{d}^{(k),L}$, and $\mathbf{d}^{(k),U}$, and can be computed using a recursive closed-form expression in $\mathcal{O}(L)$ time.

Proof: Assume that through the computation of the previous bounds on u_θ, the pre-activation layer outputs of u_θ, $y^{(k)}$, are lower and upper bounded by linear functions defined as $\mathbf{A}^{(k),L}\mathbf{x} + \mathbf{a}^{(k),L} \leq y^{(k)} \leq \mathbf{A}^{(k),U}\mathbf{x} + \mathbf{a}^{(k),U}$ and $y^{(k),L} \leq y^{(k)} \leq y^{(k),U}$ for $\mathbf{x} \in \mathcal{C}$. Additionally, we consider also that through a previous computation of bounds on $\partial_{\mathbf{x}_i} u_\theta$, the components of that network required for $\partial_{\mathbf{x}_i^2} u_\theta$, *i.e.*, $\partial_{\mathbf{x}_i} z^{(k-1)}$ and $\partial_{z^{(k-1)}} z^{(k)}$ are lower and upper bounded by linear functions. In particular, $\mathbf{C}^{(k),L}\mathbf{x}+\mathbf{c}^{(k),L} \leq \partial_{\mathbf{x}_i} z^{(k-1)} \leq \mathbf{C}^{(k),U}\mathbf{x}+\mathbf{c}^{(k),U}$ and $\mathbf{D}^{(k),L}\mathbf{x}+\mathbf{d}^{(k),L} \leq \partial_{z^{(k-1)}} z^{(k)} \leq \mathbf{D}^{(k),U}\mathbf{x} + \mathbf{d}^{(k),U}$.

Take the upper and lower bound functions for $\partial_{\mathbf{x}_i^2} u_\theta$ as $\partial_{\mathbf{x}_i^2} u_\theta^U$ and $\partial_{\mathbf{x}_i^2} u_\theta^L$, respectively, and the upper and lower bound functions for $\partial_{\mathbf{x}_i^2} z^{(k)}$ as $\partial_{\mathbf{x}_i^2} z^{(k),U}$ and $\partial_{\mathbf{x}_i^2} z^{(k),L}$, respectively. For the sake of simplicity of notation, we define $\mathbf{B}^{(k),+} = \mathbb{I}\left(\mathbf{B}^{(k)} \geq 0\right) \odot \mathbf{B}^{(k)}$ and $\mathbf{B}^{(k),-} = \mathbb{I}\left(\mathbf{B}^{(k)} < 0\right) \odot \mathbf{B}^{(k)}$.

Note that, unless explicitly mentioned otherwise, the non-network variables (denoted by Greek letters, as well as bold, capital and lowercase letters) used here have no relation to the ones from Appendix E.3.

Starting backwards from $\partial_{\mathbf{x}_i^2} z^{(k)}$, we have that:

$$\partial_{\mathbf{x}_i^2} z_j^{(k)} = \sum_{n=1}^{d_{k-1}} \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1)} + \partial_{z^{(k-1)}} z_{j,n}^{(k)} \partial_{\mathbf{x}_i^2} z_n^{(k-1)}.$$

Given the transitive property of the sum operator, we can bound $\partial_{\mathbf{x}_i^2} z_j^{(k)}$ by using a McCormick envelope around each of the multiplications. Assuming that for all $j \in \{1 \ldots, d_k\}, n \in \{1 \ldots, d_{k-1}\}$: $\partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),L} \leq \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k)} \leq \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),U}$, $\partial_{\mathbf{x}_i} z_n^{(k-1),L} \leq \partial_{\mathbf{x}_i} z_n^{(k-1)} \leq \partial_{\mathbf{x}_i} z_n^{(k-1),U}$, $\partial_{z^{(k-1)}} z_{j,n}^{(k)} \leq \partial_{z^{(k-1)}} z_{j,n}^{(k)} \leq \partial_{z^{(k-1)}} z_{j,n}^{(k)}$, and $\partial_{\mathbf{x}_i^2} z_n^{(k-1),L} \leq \partial_{\mathbf{x}_i^2} z_n^{(k-1)} \leq \partial_{\mathbf{x}_i^2} z_n^{(k-1),U}$, we obtain:

$$
\begin{aligned}
\partial_{\mathbf{x}_i^2} z_j^{(k)} \leq \partial_{\mathbf{x}_i^2} z_j^{(k),U} &= \sum_{n=1}^{d_{k-1}} \alpha_{0,j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1)} + \alpha_{1,j,n}^{(k)} \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k)} + \alpha_{2,j,n}^{(k)} \partial_{\mathbf{x}_i^2} z_n^{(k-1)} + \alpha_{3,j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k)} + \alpha_{4,j,n}^{(k)} \\
\partial_{\mathbf{x}_i^2} z_j^{(k)} \geq \partial_{\mathbf{x}_i^2} z_j^{(k),L} &= \sum_{n=1}^{d_{k-1}} \beta_{0,j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1)} + \beta_{1,j,n}^{(k)} \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k)} + \beta_{2,j,n}^{(k)} \partial_{\mathbf{x}_i^2} z_n^{(k-1)} + \beta_{3,j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k)} + \beta_{4,j,n}^{(k)}
\end{aligned}
\tag{22}
$$

for:

$$
\begin{aligned}
\alpha_{0,j,n}^{(k)} &= \eta_{j,n}^{(k)} \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),U} + \left(1 - \eta_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),L} \qquad \alpha_{1,j,n}^{(k)} = \eta_{j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1),L} + \left(1 - \eta_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i} z_n^{(k-1),U} \\
\alpha_{2,j,n}^{(k)} &= \gamma_{j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k),U} + \left(1 - \gamma_{j,n}^{(k)}\right) \partial_{z^{(k-1)}} z_{j,n}^{(k),L} \qquad \alpha_{3,j,n}^{(k)} = \gamma_{j,n}^{(k)} \partial_{\mathbf{x}_i^2} z_n^{(k-1),L} + \left(1 - \gamma_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i^2} z_n^{(k-1),U} \\
\alpha_{4,j,n}^{(k)} &= -\eta_{j,n}^{(k)} \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),U} \partial_{\mathbf{x}_i} z_n^{(k-1),L} - \left(1 - \eta_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),L} \partial_{\mathbf{x}_i} z_n^{(k-1),U} + \\
&\quad - \gamma_{j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k),U} \partial_{\mathbf{x}_i x_i} z_n^{(k-1),L} - \left(1 - \gamma_{j,n}^{(k)}\right) \partial_{z^{(k-1)}} z_{j,n}^{(k),L} \partial_{\mathbf{x}_i x_i} z_n^{(k-1),U}
\end{aligned}
$$

$$
\begin{aligned}
\beta_{0,j,n}^{(k)} &= \zeta_{j,n}^{(k)} \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),L} + \left(1 - \zeta_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),U} \qquad \beta_{1,j,n}^{(k)} = \zeta_{j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1),L} + \left(1 - \zeta_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i} z_n^{(k-1),U} \\
\beta_{2,j,n}^{(k)} &= \delta_{j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k),L} + \left(1 - \delta_{j,n}^{(k)}\right) \partial_{z^{(k-1)}} z_{j,n}^{(k),U} \qquad \beta_{3,j,n}^{(k)} = \delta_{j,n}^{(k)} \partial_{\mathbf{x}_i^2} z_n^{(k-1),L} + \left(1 - \delta_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i^2} z_n^{(k-1),U} \\
\beta_{4,j,n}^{(k)} &= -\zeta_{j,n}^{(k)} \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),L} \partial_{\mathbf{x}_i} z_n^{(k-1),L} - \left(1 - \zeta_{j,n}^{(k)}\right) \partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k),U} \partial_{\mathbf{x}_i} z_n^{(k-1),U} + \\
&\quad - \delta_{j,n}^{(k)} \partial_{z^{(k-1)}} z_{j,n}^{(k),L} \partial_{\mathbf{x}_i^2} z_n^{(k-1),L} - \left(1 - \delta_{j,n}^{(k)}\right) \partial_{z^{(k-1)}} z_{j,n}^{(k),U} \partial_{\mathbf{x}_i^2} z_n^{(k-1),U},
\end{aligned}
$$

where $\eta_{j,n}^{(k)}$, $\gamma_{j,n}^{(k)}$, $\zeta_{j,n}^{(k)}$ and $\delta_{j,n}^{(k)}$ are convex coefficients that can be set as hyperparameters, or optimized for as in α-CROWN (Xu et al., 2020b).

For the next step of the back-propagation process, we now need to bound $\partial_{\mathbf{x}_i} z_n^{(k-1)}$, $\partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k)}$, and $\partial_{z^{(k-1)}} z_{j,n}^{(k)}$, so as to eventually be able to write $\partial_{\mathbf{x}_i^2} z_j^{(k)}$ as a function of simply $\partial_{\mathbf{x}_i^2} z_n^{(k-1)}$ and $\mathbf{x}$. As per our assumptions at the beginning of this section, for the sake of computational efficiency we take $\partial_{\mathbf{x}_i} z_n^{(k-1)}$ and $\partial_{z^{(k-1)}} z_{j,n}^{(k)}$ from the computation of the bounds of $\partial_{\mathbf{x}_i} u_{\theta,j}$, and thus assume we have a linear upper and lower bound function of $\mathbf{x}$. This leaves us with $\partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k)}$ to bound as a linear function of $\mathbf{x}$.

Note that, as per Lemma 2, $\partial_{\mathbf{x}_i z^{(k-1)}} z_{j,n}^{(k)} = \sigma''\left(\mathbf{y}_j^{(k)}\right)\left(\mathbf{W}_{j,:}^{(k)} \partial_{\mathbf{x}_i} z^{(k-1)}\right) \mathbf{W}_{j,n}^{(k)}$. Since $\left(\mathbf{W}_{j,:}^{(k)} \partial_{\mathbf{x}_i} z^{(k-1)}\right) = \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k)} \partial_{\mathbf{x}_i} z_n^{(k-1)}$, and $\mathbf{C}_{n,:}^{(k),U} \mathbf{x} + \mathbf{c}_n^{(k),U} \leq \partial_{\mathbf{x}_i} z_n^{(k-1)} \leq \mathbf{C}_{n,:}^{(k),L} \mathbf{x} + \mathbf{c}_n^{(k),L}$ (from the assumptions above), we can write:

$$
\mathbf{W}_{j,:}^{(k)} \partial_{\mathbf{x}_i} z^{(k-1)} \leq \underbrace{\left(\sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \mathbf{C}_{n,:}^{(k),U} + \mathbf{W}_{j,n}^{(k),-} \mathbf{C}_{n,:}^{(k),L}\right)}_{\mathbf{E}_j^{(k),U}} \mathbf{x} + \underbrace{\left(\sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \mathbf{c}^{(k),U} + \mathbf{W}_{j,n}^{(k),-} \mathbf{c}^{(k),L}\right)}_{\mathbf{e}_j^{(k),U}}
$$

$$
\mathbf{W}_{j,n}^{(k)} \partial_{\mathbf{x}_i} z^{(k-1)} \geq \underbrace{\left(\sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \mathbf{C}_{n,:}^{(k),L} + \mathbf{W}_{j,n}^{(k),-} \mathbf{C}_{n,:}^{(k),U}\right)}_{\mathbf{E}_j^{(k),L}} \mathbf{x} + \underbrace{\left(\sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \mathbf{c}_n^{(k),L} + \mathbf{W}_{j,n}^{(k),-} \mathbf{c}_n^{(k),U}\right)}_{\mathbf{e}_j^{(k),L}}.
$$

We define $\theta_j^{(k),U} = \max_{\mathbf{x} \in \mathcal{C}} \mathbf{E}_j^{(k),U} \mathbf{x} + \mathbf{e}_j^{(k),U}$ and $\theta_j^{(k),L} = \min_{\mathbf{x} \in \mathcal{C}} \mathbf{E}_j^{(k),L} \mathbf{x} + \mathbf{e}_j^{(k),L}$. As with the first derivative case, since $y_j^{(k),L} \leq y_j^{(k)} \leq y_j^{(k),U}$, we can obtain a linear upper and lower bound relaxation for $\sigma''\left(y_j^{(k)}\right)$, such that $\lambda_j^{(k),L}\left(y_j^{(k)}+\mu_j^{(k),L}\right)\leq\sigma''\left(y_j^{(k)}\right)\leq\lambda_j^{(k),U}\left(y_j^{(k)}+\mu_j^{(k),U}\right)$, as well as the values $\iota_j^{(k),L}\leq\sigma''\left(y_j^{(k)}\right)\leq\iota_j^{(k),U}$. By considering the assumption that $\mathbf{A}_{j,:}^{(k),U}\mathbf{x}+\mathbf{a}_j^{(k),U}\leq y^{(k)}\leq\mathbf{A}_{j,:}^{(k),L}\mathbf{x}+\mathbf{a}_j^{(k),L}$, we can obtain:

$$\sigma''\left(y_j^{(k)}\right)\leq\underbrace{\left(\lambda_j^{(k),U,+}\mathbf{A}_{j,:}^{(k),U}+\lambda_j^{(k),U,-}\mathbf{A}_{j,:}^{(k),L}\right)}_{\mathbf{H}_j^{(k),U}}\mathbf{x}+\underbrace{\left(\lambda_j^{(k),U,+}\mathbf{a}_j^{(k),U}+\lambda_j^{(k),U,-}\mathbf{a}_j^{(k),L}+\lambda_j^{(k),U}\mu_j^{(k),U}\right)}_{\mathbf{h}_j^{(k),U}}$$

$$\sigma''\left(y_j^{(k)}\right)\geq\underbrace{\left(\lambda_j^{(k),L,+}\mathbf{A}_{j,:}^{(k),L}+\lambda_j^{(k),L,-}\mathbf{A}_{j,:}^{(k),U}\right)}_{\mathbf{H}_j^{(k),L}}\mathbf{x}+\underbrace{\left(\lambda_j^{(k),L,+}\mathbf{a}_j^{(k),L}+\lambda_j^{(k),L,-}\mathbf{a}_j^{(k),U}+\lambda_j^{(k),L}\mu_j^{(k),L}\right)}_{\mathbf{h}_j^{(k),L}}.$$

This allows us to relax $\sigma''\left(y_j^{(k)}\right)\left(\mathbf{W}_{j,:}^{(k)}\partial_{\mathbf{x}_i}z^{(k-1)}\right)$ using McCormick envelopes:

$$\begin{aligned}
\sigma''\left(y_j^{(k)}\right)\left(\mathbf{W}_{j,:}^{(k)}\partial_{\mathbf{x}_i}z^{(k-1)}\right)&\leq\nu_{0,j}^{(k),U}\left(\mathbf{W}_{j,:}^{(k)}\partial_{\mathbf{x}_i}z^{(k-1)}\right)+\nu_{1,j}^{(k),U}\sigma''\left(y_j^{(k)}\right)+\nu_{2,j}^{(k),U}\\
\sigma''\left(y_j^{(k)}\right)\left(\mathbf{W}_{j,:}^{(k)}\partial_{\mathbf{x}_i}z^{(k-1)}\right)&\geq\nu_{0,j}^{(k),L}\left(\mathbf{W}_{j,:}^{(k)}\partial_{\mathbf{x}_i}z^{(k-1)}\right)+\nu_{1,j}^{(k),L}\sigma''\left(y_j^{(k)}\right)+\nu_{2,j}^{(k),L},
\end{aligned}$$

for:

$$\begin{aligned}
\nu_{0,j}^{(k),U}&=\rho_j^{(k)}\iota_j^{(k),U}+\left(1-\rho_j^{(k)}\right)\iota_j^{(k),L}\qquad\nu_{1,j,n}^{(k),U}=\rho_j^{(k)}\theta_j^{(k),L}+\left(1-\rho_j^{(k)}\right)\iota_j^{(k),U}\\
\nu_{2,j}^{(k),U}&=-\rho_j^{(k)}\iota_j^{(k),U}\theta_j^{(k),L}-\left(1-\rho_j^{(k)}\right)\iota_j^{(k),L}\theta_j^{(k),U}
\end{aligned}$$

$$\begin{aligned}
\nu_{0,j}^{(k),L}&=\tau_j^{(k)}\iota_j^{(k),L}+\left(1-\tau_j^{(k)}\right)\iota_j^{(k),U}\qquad\nu_{1,j}^{(k),L}=\tau_j^{(k)}\theta_j^{(k),L}+\left(1-\tau_j^{(k)}\right)\theta_j^{(k),U}\\
\nu_{2,j}^{(k),L}&=-\tau_j^{(k)}\iota_j^{(k),L}\theta_j^{(k),L}-\left(1-\tau_j^{(k)}\right)\iota_j^{(k),U}\theta_j^{(k),U},
\end{aligned}$$

where $\rho_j^{(k)}$ and $\tau_j^{(k)}$ are convex coefficients that can be set as hyperparameters, or optimized for as in α-CROWN (Xu et al., 2020b). By replacing this multiplication in the expression from Lemma 2, we bound $\partial_{\mathbf{x}_i z^{(k-1)}}z_{j,n}^{(k)}$ as:

$$\begin{aligned}
\partial_{\mathbf{x}_i z^{(k-1)}}z_{j,n}^{(k)}&\leq\upsilon_{0,j,n}^{(k),U}\left(\mathbf{W}_{j,:}^{(k)}\partial_{\mathbf{x}_i}z^{(k-1)}\right)+\upsilon_{1,j,n}^{(k),U}\sigma''\left(y_j^{(k)}\right)+\upsilon_{2,j}^{(k),U}\\
\partial_{\mathbf{x}_i z^{(k-1)}}z_{j,n}^{(k)}&\geq\upsilon_{0,j,n}^{(k),L}\left(\mathbf{W}_{j,:}^{(k)}\partial_{\mathbf{x}_i}z^{(k-1)}\right)+\upsilon_{1,j,n}^{(k),L}\sigma''\left(y_j^{(k)}\right)+\upsilon_{2,j}^{(k),L},
\end{aligned}$$

for:

$$\upsilon_{i,j,n}^{(k),U}=\nu_{i,j}^{(k),U}\mathbf{W}_{j,n}^{(k),+}+\nu_{i,j}^{(k),L}\mathbf{W}_{j,n}^{(k),-},\quad\upsilon_{i,j,n}^{(k),L}=\nu_{i,j}^{(k),L}\mathbf{W}_{j,n}^{(k),+}+\nu_{i,j}^{(k),U}\mathbf{W}_{j,n}^{(k),-}\qquad i\in\{0,1,2\}.$$

By replacing the lower and upper bounds for $\sigma''(y_j^{(k)})$ and $\left(\mathbf{W}_{j,:}^{(k)}\partial_{\mathbf{x}_i}z^{(k-1)}\right)$ in the previous inequality, we obtain the expression:

$$\begin{aligned}
\partial_{\mathbf{x}_i z^{(k-1)}}z_{j,n}^{(k)}&\leq\mathbf{M}_{j,n}^{(k),U}\mathbf{x}+\mathbf{m}_{j,n}^{(k),U}\\
\partial_{\mathbf{x}_i z^{(k-1)}}z_{j,n}^{(k)}&\geq\mathbf{M}_{j,n}^{(k),L}\mathbf{x}+\mathbf{m}_{j,n}^{(k),L},
\end{aligned}$$

for:

$$\begin{aligned}
\mathbf{M}_{j,n}^{(k),U}&=\upsilon_{0,j,n}^{(k),U,+}\mathbf{E}_j^{(k),U}+\upsilon_{0,j,n}^{(k),U,-}\mathbf{E}_j^{(k),L}+\upsilon_{1,j,n}^{(k),U,+}\mathbf{H}_j^{(k),U}+\upsilon_{1,j,n}^{(k),U,-}\mathbf{H}_j^{(k),L}\\
\mathbf{m}_{j,n}^{(k),U}&=\upsilon_{0,j,n}^{(k),U,+}\mathbf{e}_j^{(k),U}+\upsilon_{0,j,n}^{(k),U,-}\mathbf{e}_j^{(k),L}+\upsilon_{1,j,n}^{(k),U,+}\mathbf{h}_j^{(k),U}+\upsilon_{1,j,n}^{(k),U,-}\mathbf{h}_j^{(k),L}+\upsilon_{2,j,n}^{(k),U}\\
\mathbf{M}_{j,n}^{(k),L}&=\upsilon_{0,j,n}^{(k),L,+}\mathbf{E}_j^{(k),L}+\upsilon_{0,j,n}^{(k),L,-}\mathbf{E}_j^{(k),U}+\upsilon_{1,j,n}^{(k),L,+}\mathbf{H}_j^{(k),L}+\upsilon_{1,j,n}^{(k),L,-}\mathbf{H}_j^{(k),U}\\
\mathbf{m}_{j,n}^{(k),L}&=\upsilon_{0,j,n}^{(k),L,+}\mathbf{e}_j^{(k),L}+\upsilon_{0,j,n}^{(k),L,-}\mathbf{e}_j^{(k),U}+\upsilon_{1,j,n}^{(k),L,+}\mathbf{h}_j^{(k),L}+\upsilon_{1,j,n}^{(k),L,-}\mathbf{h}_j^{(k),U}+\upsilon_{2,j,n}^{(k),L}.
\end{aligned}$$

Finally in the derivation of $\partial_{\mathbf{x}_i^2} z_j^{(k)}$ as a function of $\mathbf{x}$ and $\partial_{\mathbf{x}_i^2} z^{(k-1)}$, we just have to replace all the quantities in Equation 22 (recalling from the assumptions that $\mathbf{C}^{(k),U}\mathbf{x} + \mathbf{c}^{(k),U} \leq \partial_{\mathbf{x}_i} z^{(k-1)} \leq \mathbf{C}^{(k),L}\mathbf{x} + \mathbf{c}^{(k),L}$ and $\mathbf{D}^{(k),U}\mathbf{x} + \mathbf{d}^{(k),U} \leq \partial_{z^{(k-1)}} z^{(k)} \leq \mathbf{D}^{(k),L}\mathbf{x} + \mathbf{d}^{(k),L}$) to obtain:

$$\begin{aligned}
\partial_{\mathbf{x}_i^2} z_j^{(k)} \leq \partial_{\mathbf{x}_i^2} z_j^{(k),U} &= \sum_{n=1}^{d_{k-1}} \alpha_{2,j,n}^{(k)} \partial_{\mathbf{x}_i^2} z_n^{(k-1)} + \alpha_{5,j,n}^{(k)} \mathbf{x} + \alpha_{6,j,n}^{(k)} \\
\partial_{\mathbf{x}_i^2} z_j^{(k)} \geq \partial_{\mathbf{x}_i^2} z_j^{(k),L} &= \sum_{n=1}^{d_{k-1}} \beta_{2,j,n}^{(k)} \partial_{\mathbf{x}_i^2} z_n^{(k-1)} + \beta_{5,j,n}^{(k)} \mathbf{x} + \beta_{6,j,n}^{(k)},
\end{aligned} \tag{23}$$

where:

$$\begin{aligned}
\alpha_{5,j,n}^{(k)} &= \alpha_{0,j,n}^{(k),+}\mathbf{C}_n^{(k),U} + \alpha_{0,j,n}^{(k),-}\mathbf{C}_n^{(k),L} + \alpha_{1,j,n}^{(k),+}\mathbf{M}_{j,n}^{(k),U} + \alpha_{1,j,n}^{(k),-}\mathbf{M}_{j,n}^{(k),L} + \alpha_{3,j,n}^{(k),+}\mathbf{D}_{j,n}^{(k),U} + \alpha_{3,j,n}^{(k),-}\mathbf{D}_{j,n}^{(k),L} \\
\alpha_{6,j,n}^{(k)} &= \alpha_{0,j,n}^{(k),+}\mathbf{c}_n^{(k),U} + \alpha_{0,j,n}^{(k),-}\mathbf{c}_n^{(k),L} + \alpha_{1,j,n}^{(k),+}\mathbf{m}_{j,n}^{(k),U} + \alpha_{1,j,n}^{(k),-}\mathbf{m}_{j,n}^{(k),L} + \alpha_{3,j,n}^{(k),+}\mathbf{d}_{j,n}^{(k),U} + \alpha_{3,j,n}^{(k),-}\mathbf{d}_{j,n}^{(k),L} + \alpha_{4,j,n}^{(k)} \\
\beta_{5,j,n}^{(k)} &= \beta_{0,j,n}^{(k),+}\mathbf{C}_n^{(k),L} + \beta_{0,j,n}^{(k),-}\mathbf{C}_n^{(k),U} + \beta_{1,j,n}^{(k),+}\mathbf{M}_{j,n}^{(k),L} + \beta_{1,j,n}^{(k),-}\mathbf{M}_{j,n}^{(k),U} + \beta_{3,j,n}^{(k),+}\mathbf{D}_{j,n}^{(k),L} + \beta_{3,j,n}^{(k),-}\mathbf{D}_{j,n}^{(k),U} \\
\beta_{6,j,n}^{(k)} &= \beta_{0,j,n}^{(k),+}\mathbf{c}_n^{(k),L} + \beta_{0,j,n}^{(k),-}\mathbf{c}_n^{(k),U} + \beta_{1,j,n}^{(k),+}\mathbf{m}_{j,n}^{(k),L} + \beta_{1,j,n}^{(k),-}\mathbf{m}_{j,n}^{(k),U} + \beta_{3,j,n}^{(k),+}\mathbf{d}_{j,n}^{(k),L} + \beta_{3,j,n}^{(k),-}\mathbf{d}_{j,n}^{(k),U} + \beta_{4,j,n}^{(k)}
\end{aligned}$$

This forms a recursion of exactly the same form as Equation 11 from Appendix E.3, where only the coefficients of $\partial_{\mathbf{x}_i^2} z_n^{(k-1)}$ and $\mathbf{x}$ are different ($\alpha_{0,j,n}^{(k)}$ in this case is referred by $\alpha_{2,j,n}^{(k)}$, $\alpha_{3,j,n}^{(k)}$ by $\alpha_{5,j,n}^{(k)}$, and $\alpha_{4,j,n}^{(k)}$ by $\alpha_{6,j,n}^{(k)}$, and similarly for the β values). This yields:

$$\partial_{\mathbf{x}_i x_i} z_j^{(L-1),U} = \rho_{0,j,i}^{(1),U} + \sum_{r=1}^{d_0} \rho_{1,j,r}^{(1),U}\mathbf{x} + \rho_{2,j,r}^{(1),U},$$

where:

$$\begin{aligned}
\rho_{0,j,r}^{(k-1),U} &= \begin{cases} \alpha_{2,n,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \rho_{0,j,n}^{(k),U} \mu_{2,n,r}^{(k-1),U} & \text{if } k \in \{2,\dots,L-1\} \end{cases} \\
\rho_{1,j,r}^{(k-1),U} &= \begin{cases} \alpha_{5,n,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \rho_{0,j,n}^{(k),U} \mu_{5,n,r}^{(k-1),U} + \frac{1}{d_{k-2}}\rho_{1,j,n}^{(k),U} & \text{if } k \in \{2,\dots,L-1\} \end{cases} \\
\rho_{2,j,r}^{(k-1),U} &= \begin{cases} \alpha_{6,n,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \rho_{0,j,n}^{(k),U} \mu_{6,n,r}^{(k-1),U} + \frac{1}{d_{k-2}}\rho_{2,j,n}^{(k),U} & \text{if } k \in \{2,\dots,L-1\} \end{cases},
\end{aligned}$$

and:

$$\mu_{p,n,:}^{(k-1),U} = \begin{cases} \alpha_{p,n,:}^{(k-1)} & \text{if } \rho_{0,j,n}^{(k),U} \geq 0 \\ \beta_{p,n,:}^{(k-1)} & \text{if } \rho_{0,j,n}^{(k),U} < 0 \end{cases}, p \in \{2,5,6\}.$$

And following the same argument:

$$\partial_{\mathbf{x}_i} z_j^{(L-1),L} = \rho_{0,j,i}^{(1),L} + \sum_{r=1}^{d_0} \rho_{1,j,r}^{(1),L}\mathbf{x} + \rho_{2,j,r}^{(1),L},$$

where:

$$\begin{aligned}
\rho_{0,j,r}^{(k-1),L} &= \begin{cases} \beta_{2,n,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \rho_{0,j,n}^{(k),L} \mu_{2,n,r}^{(k-1),L} & \text{if } k \in \{2,\dots,L-1\} \end{cases} \\
\rho_{1,j,r}^{(k-1),L} &= \begin{cases} \beta_{5,n,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \rho_{0,j,n}^{(k),L} \mu_{5,n,r}^{(k-1),L} + \frac{1}{d_{k-2}}\rho_{1,j,n}^{(k),L} & \text{if } k \in \{2,\dots,L-1\} \end{cases} \\
\rho_{2,j,r}^{(k-1),L} &= \begin{cases} \beta_{6,n,r}^{(k)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \rho_{0,j,n}^{(k),L} \mu_{6,n,r}^{(k-1),L} + \frac{1}{d_{k-2}}\rho_{2,j,n}^{(k),L} & \text{if } k \in \{2,\dots,L-1\} \end{cases},
\end{aligned}$$

and:

$$\mu_{p,n,:}^{(k-1),L} = \begin{cases} \beta_{p,n,:}^{(k-1)} & \text{if } \rho_{0,j,n}^{(k),L} \geq 0 \\ \alpha_{p,n,:}^{(k-1)} & \text{if } \rho_{0,j,n}^{(k),L} < 0 \end{cases} \quad , p \in \{2,5,6\}$$

With these expressions, we can compute the required $\partial_{\mathbf{x}_i^2} z_n^{(k-1),L}$ and $\partial_{\mathbf{x}_i^2} z_n^{(k-1),U}$ which we assumed to be known to derive Equation 22.

Finally, with the exact same argument as in Appendix E.3, we obtain:

$$\partial_{\mathbf{x}_i} u_{\theta,j}^U = \psi_{0,j,i}^{(1),U} + \sum_{r=1}^{d_0} \psi_{1,j,r}^{(1),U} \mathbf{x} + \psi_{2,j,r}^{(1),U},$$

where:

$$\begin{aligned}
\psi_{0,j,r}^{(k-1),U} &= \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \alpha_{2,n,r}^{(k-1)} + \mathbf{W}_{j,n}^{(k),-} \beta_{2,n,r}^{(k-1)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \psi_{0,j,n}^{(k),U} \psi_{2,n,r}^{(k-1),U} & \text{if } k \in \{2,\ldots,L-1\} \end{cases} \\
\psi_{1,j,r}^{(k-1),U} &= \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \alpha_{5,n,r}^{(k-1)} + \mathbf{W}_{j,n}^{(k),-} \beta_{5,n,r}^{(k-1)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \psi_{0,j,n}^{(k),U} \psi_{5,n,r}^{(k-1),U} + \frac{1}{d_{k-2}} \psi_{1,j,n}^{(k),U} & \text{if } k \in \{2,\ldots,L-1\} \end{cases} \\
\psi_{2,j,r}^{(k-1),U} &= \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \alpha_{6,n,r}^{(k-1)} + \mathbf{W}_{j,n}^{(k),-} \beta_{6,n,r}^{(k-1)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \psi_{0,j,n}^{(k),U} \psi_{6,n,r}^{(k-1),U} + \frac{1}{d_{k-2}} \psi_{2,j,n}^{(k),U} & \text{if } k \in \{2,\ldots,L-1\} \end{cases}
\end{aligned} \quad ,$$

and:

$$\psi_{p,n,:}^{(k-1)} = \begin{cases} \alpha_{p,n,:}^{(k-1)} & \text{if } \psi_{0,j,n}^{(k),U} \geq 0 \\ \beta_{p,n,:}^{(k-1)} & \text{if } \psi_{0,j,n}^{(k),U} < 0 \end{cases} \quad , p \in \{2,5,6\}.$$

And similarly for the lower bound:

$$\partial_{\mathbf{x}_i} u_{\theta,j}^L = \psi_{0,j,i}^{(1),L} + \sum_{r=1}^{d_0} \psi_{1,j,r}^{(1),L} \mathbf{x} + \psi_{2,j,r}^{(1),L},$$

where:

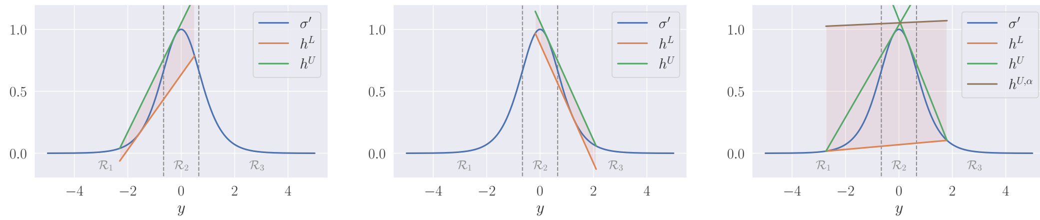
$$\begin{aligned}
\psi_{0,j,r}^{(k-1),L} &= \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \beta_{2,n,r}^{(k-1)} + \mathbf{W}_{j,n}^{(k),-} \alpha_{2,n,r}^{(k-1)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \psi_{0,j,n}^{(k),L} \psi_{2,n,r}^{(k-1),L} & \text{if } k \in \{2,\ldots,L-1\} \end{cases} \\
\psi_{1,j,r}^{(k-1),L} &= \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \beta_{5,n,r}^{(k-1)} + \mathbf{W}_{j,n}^{(k),-} \alpha_{5,n,r}^{(k-1)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \psi_{0,j,n}^{(k),L} \psi_{5,n,r}^{(k-1),L} + \frac{1}{d_{k-2}} \psi_{1,j,n}^{(k),L} & \text{if } k \in \{2,\ldots,L-1\} \end{cases} \\
\psi_{2,j,r}^{(k-1),L} &= \begin{cases} \sum_{n=1}^{d_{k-1}} \mathbf{W}_{j,n}^{(k),+} \beta_{6,n,r}^{(k-1)} + \mathbf{W}_{j,n}^{(k),-} \alpha_{6,n,r}^{(k-1)} & \text{if } k = L \\ \sum_{n=1}^{d_{k-1}} \psi_{0,j,n}^{(k),L} \psi_{6,n,r}^{(k-1),L} + \frac{1}{d_{k-2}} \psi_{2,j,n}^{(k),L} & \text{if } k \in \{2,\ldots,L-1\} \end{cases}
\end{aligned} \quad ,$$

and:

$$\psi_{p,n,:}^{(k-1),L} = \begin{cases} \beta_{p,n,:}^{(k-1)} & \text{if } \psi_{0,j,n}^{(k),L} \geq 0 \\ \alpha_{p,n,:}^{(k-1)} & \text{if } \psi_{0,j,n}^{(k),L} < 0 \end{cases} \quad , p \in \{2,5,6\}.$$

E.5 Formulation and proof of closed-form global bounds on $\partial_{\mathbf{x}_i} u_\theta$

Lemma 3 (Closed-form global bounds on $\partial_{\mathbf{x}_i} u_\theta$)**.** *For every $j \in \{1,\ldots,d_L\}$ there exist two values $\kappa_j^U \in \mathbb{R}$ and $\kappa_j^L \in \mathbb{R}$, such that $\forall \mathbf{x} \in \mathcal{C} = \{\mathbf{x} \in \mathbb{R}^{d_0} : \mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U\}$ it holds that*



(a) $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_2$ (b) $l_b \in \mathcal{R}_2$ and $u_b \in \mathcal{R}_3$ (c) $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_3$

Figure 7: *Relaxing* $\sigma'(y) = 1 - \tanh^2(y)$: examples of the linear relaxations of σ' for different sets of l_b and u_b.

$\kappa_j^L \leq \partial_{\mathbf{x}_i} u_{\theta,j} \leq \kappa_j^U$, *with:*

$$\kappa_j^U = \mathbf{B}^{U,+}\mathbf{x}^U + \mathbf{B}^{U,-}\mathbf{x}^L + \phi_{0,j,i}^{(1)} + \sum_{r=1}^{d_0} \phi_{2,j,r}^{(1)}$$

$$\kappa_j^L = \mathbf{B}^{L,+}\mathbf{x}^L + \mathbf{B}^{L,-}\mathbf{x}^U + \psi_{0,j,i}^{(1)} + \sum_{r=1}^{d_0} \psi_{2,j,r}^{(1)},$$

where $\mathbf{B}^U = \sum_{r=1}^{d_0} \phi_{1,j,r}^{(1)}$, $\mathbf{B}^L = \sum_{r=1}^{d_0} \psi_{1,j,r}^{(1)}$, *and* $\mathbf{B}^{\cdot,+} = \mathbb{I}(\mathbf{B}^{\cdot} \geq 0) \odot \mathbf{B}^{\cdot}$ *and* $\mathbf{B}^{\cdot,-} = \mathbb{I}(\mathbf{B}^{\cdot} < 0) \odot \mathbf{B}^{\cdot}$.

Proof. Take a function $f : \mathbb{R}^{d_0} \to \mathbb{R}$ defined as $f(\mathbf{x}) = \mathbf{v}^\top \mathbf{x} + c$ for $\mathbf{v} \in \mathbb{R}^{d_0}$ and $c \in \mathbb{R}$, as well as a domain $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^{d_0} : \mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U\}$. Given the perpendicularity of the constraints in $\mathcal{C}$, by separating each component of f we obtain:

$$\max_{\mathbf{x} \in \mathcal{C}} f(\mathbf{x}) = (\mathbf{v}^+)^\top \mathbf{x}^U + (\mathbf{v}^-)^\top \mathbf{x}^L + c, \qquad \min_{\mathbf{x} \in \mathcal{C}} f(\mathbf{x}) = (\mathbf{v}^+)^\top \mathbf{x}^L + (\mathbf{v}^-)^\top \mathbf{x}^U + c,$$

where $\mathbf{v}^+ = \mathbb{I}(\mathbf{v} \geq 0) \odot \mathbf{v}$ and $\mathbf{v}^- = \mathbb{I}(\mathbf{v} < 0) \odot \mathbf{v}$. □

F CORRECTNESS CERTIFICATION FOR PINNS WITH $\tanh$ ACTIVATIONS

∂-CROWN allows one to compute lower and upper bounds on the outputs of $\partial_{\mathbf{x}_i} u_\theta$, $\partial_{\mathbf{x}_i^2} u_\theta$ and f_θ as long as we can obtain linear bounds for u_θ's activations, σ, $\partial_{\mathbf{x}_i} u_\theta$'s activations, σ', and $\partial_{\mathbf{x}_i^2} u_\theta$'s activations, σ'', assuming previously computed bounds on the input of those activations. In this section we explore how to compute those bounds when u_θ has $\tanh$ activations.

Throughout, we assume the activation's input (y) is lower bounded by l_b and upper bounded by u_b (*i.e.*, $l_b \leq y \leq u_b$), and define the upper bound line as $h^U(y) = \alpha^U(y + \beta^U)$, and the lower bound line as $h^L(y) = \alpha^L(y + \beta^L)$. For the sake of brevity, we define for a function $h : \mathbb{R} \to \mathbb{R}$, and points $p, d \in \mathbb{R}$ the function $\tau(h, p, d) = {}^{(h(p) - h(d))}/_{(p-d)} - h'(d)$. This is useful as for a given h and p, if there exists a $d \in [d_l, d_u]$, such that $\tau_{d_l,d_u}(h, p, d) = 0$, then $h'(d)$ is the slope of a tangent line to h that passes through p and d.

Bounding $\sigma(y) = \tanh(y)$ We follow the bounds provided in CROWN (Zhang et al., 2018), by observing that $\tanh$ is a convex function for $y < 0$ and concave for $y > 0$. For $l_b \leq u_b \leq 0$ we let h^U be the line that connects l_b and u_b, and for an arbitrary $d \in [l_b, u_b]$ we let h^L be the tangent line at that point. Similarly, for $0 \leq l_b \leq u_b$ we let h^L be the line that connects l_b and u_b, and for an arbitrary $d \in [l_b, u_b]$ we let h^U be the tangent line at that point. For the last case where $l_b \leq 0 \leq u_b$, we let h^U be the tangent line at $d_1 \geq 0$ that passes through $(l_b, \sigma(l_b))$, and h^L be the tangent line at $d_2 \leq 0$ that passes through $(u_b, \sigma(u_b))$. Given these bounds were given in Zhang et al. (2018), we omit visual representations of them.

Table 5: *Relaxing* $\sigma'(y) = 1 - \tanh^2(y)$: linear upper and lower bounds for a given l_b and u_b.

l_b	u_b	α^U	β^U	α^L	β^L
$\mathcal{R}_1$ $\mathcal{R}_3$	$\mathcal{R}_1$ $\mathcal{R}_3$	$(\sigma(u_b)-\sigma(l_b))/(u_b-l_b)$	$\sigma(l_b)/\alpha^U - l_b$	$\sigma'(d), d \in [l_b, u_b]$	$\sigma(d)/\alpha^L - d$
$\mathcal{R}_2$	$\mathcal{R}_2$	$\sigma'(d), d \in [l_b, u_b]$	$\sigma(d)/\alpha^U - d$	$(\sigma(u_b)-\sigma(l_b))/(u_b-l_b)$	$\sigma(l_b)/\alpha^L - l_b$
$\mathcal{R}_1$	$\mathcal{R}_2$	$\sigma'(d_1)$, $\tau_{y_1,u_b}(\sigma', l_b, d_1) = 0$	$\sigma(l_b)/\alpha^U - l_b$	$\sigma'(d_2)$, $\tau_{l_b,y_1}(\sigma', u_b, d_2) = 0$	$\sigma(u_b)/\alpha^L - u_b$
$\mathcal{R}_2$	$\mathcal{R}_3$	$\sigma'(d_1)$, $\tau_{l_b,y_2}(\sigma', u_b, d_1) = 0$	$\sigma(u_b)/\alpha^U - u_b$	$\sigma'(d_2)$, $\tau_{y_2,u_b}(\sigma', l_b, d_2) = 0$	$\sigma(l_b)/\alpha^L - l_b$
$\mathcal{R}_1$	$\mathcal{R}_3$	$\alpha\sigma'(d_1) + (1-\alpha)\sigma'(d_2)$, $\tau_{l_b,0}(\sigma', l_b, d_1) = 0$, $\tau_{0,u_b}(\sigma', u_b, d_2) = 0$	$\alpha\beta_1^U + (1-\alpha)\beta_2^U$, $\beta_1^U = \sigma(l_b)/\sigma'(d_1) - l_b$, $\beta_2^U = \sigma(u_b)/\sigma'(d_2) - u_b$	$\begin{cases}\sigma'(d_3), & -l_b \geq u_b \\ \sigma'(d_4), & -l_b < u_b\end{cases}$, $\tau_{l_b,y_1}(\sigma', u_b, d_3) = 0$, $\tau_{y_2,u_b}(\sigma', l_b, d_4) = 0$	$\begin{cases}\frac{\sigma'(u_b)}{\sigma'(d_3)} - u_b, & -l_b \geq u_b \\ \frac{\sigma'(l_b)}{\sigma'(d_4)} - l_b, & -l_b < u_b\end{cases}$

Bounding $\sigma'(y) = 1 - \tanh^2(y)$ The derivative of $\tanh(y)$, $1 - \tanh^2(y)$, is a more complicated function. By inspecting it's derivative, $\sigma''(y) = -2\tanh(y)(1 - \tanh^2(y))$, we conclude that there are two inflection points at $y_1 = \max \sigma''(y)$ and $y_2 = \min \sigma''(y)$, leading to three different regions: $y \in]-\infty, y_1]$ ($\mathcal{R}_1$, the first convex region), $y \in]y_1, y_2]$ ($\mathcal{R}_2$, the concave region), and $y \in]y_2, +\infty[$ ($\mathcal{R}_3$, the second convex region). As a result, there are 6 combinations for the location of l_b and u_b which must be resolved.

The first two cases are the straightforward: if $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_1$ or $l_b \in \mathcal{R}_3$ and $u_b \in \mathcal{R}_3$, *i.e.*, if both ends are in the same convex region, then we use the same relaxation as in the bounding of $\tanh$ in the convex region - h^U is the line that connects l_b and u_b, while h^L is a tangent line at a point $d \in [l_b, u_b]$. Similarly for the case where $l_b \in \mathcal{R}_2$ and $u_b \in \mathcal{R}_2$, we take the solution from the $\tanh$ concave side and use h^L to be the line that connects l_b and u_b, and h^U to be the tangent line at a point $d \in [l_b, u_b]$. The next case is $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_2$, *i.e.*, l_b in the first convex region and u_b in the concave one. In this case we use the same bounding as in the $\tanh$ case when $l_b \leq 0 \leq u_b$: h^U is the tangent line at $d_1 \geq y_1$ that passes through $(l_b, \sigma'(l_b))$, and h^L is the tangent line at $d_2 \leq y_1$ that passes through $(u_b, \sigma'(u_b))$. In a similar fashion, for the case in which $l_b \in \mathcal{R}_2$ and $u_b \in \mathcal{R}_3$, *i.e.*, l_b in the concave region and u_b in the second convex region, we take the opposite approach: h^U is the tangent line at $d_1 \leq y_2$ that passes through $(u_b, \sigma'(u_b))$, and h^L is the tangent line at $d_2 \geq y_2$ that passes through $(l_b, \sigma'(l_b))$. These two cases are plotted in Figures 7a and 7b.

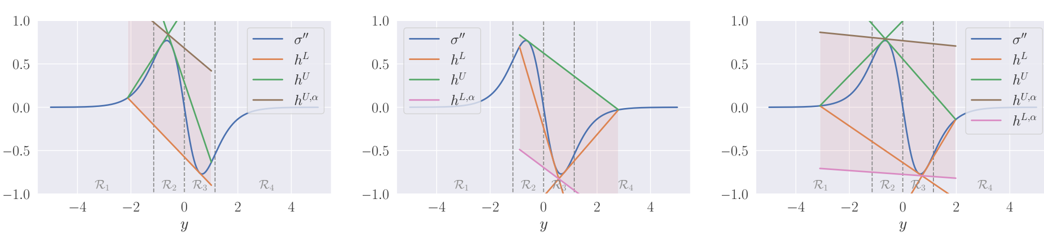
Finally, we tackle the case where $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_3$, *i.e.*, where l_b is in the first convex region and u_b is in the second convex region. Given there is a concave region in between them, two valid upper bounds would be the ones considered previously for $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_2$, and $l_b \in \mathcal{R}_2$ and $u_b \in \mathcal{R}_3$. To obtain these bounds, we shift the upper bound in the first case to 0, and the lower bound in the second case to 0 (see h^U in Figure 7c). As our bounding requires a single h^U, we take a convex combination of the two bounds obtained, $h^{U,\alpha}$. For the lower bound, we use a line that passes by either $(u_b, \sigma'(u_b))$, if $-l_b \geq u_b$, or by $(l_b, \sigma'(l_b))$, otherwise, as well as by a tangent point $d_3 \in \mathcal{R}_1$, if $-l_b \geq u_b$, or by $d_4 \in \mathcal{R}_3$, otherwise. See the line $h^{U,\alpha}$ in Figure 7c for a visual representation.

Bounding $\sigma''(y) = -2\tanh(y)\left(1 - \tanh(y)^2\right)$ By inspecting the derivative of σ'', $\sigma''(y) = -2 + 8\tanh^2(y) - 6\tanh^4(y)$, we conclude there are three inflection points for this function, one at $y_1 = \arg\max_{y \leq 0} \sigma'''(y)$, another at $y_2 = 0$, and finally at $y_3 = -y_1$. Take also, for the sake of bounding, $y_{\max} = \arg\max_{y \leq 0} \sigma''(y)$ and $y_{\min} = \arg\min_{y \leq 0} \sigma''(y)$. This leads to four different regions of σ'': $y \in]-\infty, y_1]$ ($\mathcal{R}_1$, the first convex region), $y \in]y_1, y_2]$ ($\mathcal{R}_2$, the first concave region), $y \in]y_2, y_3]$ ($\mathcal{R}_3$, the second convex region), and $y \in]y_3, +\infty[$ ($\mathcal{R}_4$, the second concave region). This leads to 10 combinations for the location of l_b and u_b.

The first four are straightforward: if $l_b \in \mathcal{R}_i$ and $u_b \in \mathcal{R}_i$ for $i \in \{1, \ldots, 4\}$, then we use exactly the same approximations as for σ and σ'', varying only based on the convexity of $\mathcal{R}_i$. Similarly,

Table 6: *Relaxing* $\sigma''(y) = -2\tanh(y)\left(1 - \tanh^2(y)\right)$: linear upper and lower bounds for a given l_b and u_b.

l_b	u_b	α^U	β^U	α^L	β^L
$\mathcal{R}_1$ $\mathcal{R}_3$	$\mathcal{R}_1$ $\mathcal{R}_3$	$(\sigma''(u_b)-\sigma''(l_b))/(u_b-l_b)$	$\sigma''(l_b)/\alpha^U - l_b$	$\sigma'''(d)$, $d \in [l_b, u_b]$	$\sigma''(d)/\alpha^L - d$
$\mathcal{R}_2$ $\mathcal{R}_4$	$\mathcal{R}_2$ $\mathcal{R}_4$	$\sigma'''(d)$, $d \in [l_b, u_b]$	$\sigma''(d)/\alpha^U - d$	$(\sigma''(u_b)-\sigma''(l_b))/(u_b-l_b)$	$\sigma''(l_b)/\alpha^L - l_b$
$\mathcal{R}_1$	$\mathcal{R}_2$	$\sigma'''(d_1)$, $\tau_{y_1,u_b}(\sigma'', l_b, d_1) = 0$	$\sigma(l_b)/\alpha^U - l_b$	$\sigma'''(d_2)$, $\tau_{l_b,y_1}(\sigma'', u_b, d_2) = 0$	$\sigma''(u_b)/\alpha^L - u_b$
$\mathcal{R}_3$	$\mathcal{R}_4$	$\sigma'''(d_1)$, $\tau_{y_3,u_b}(\sigma'', l_b, d_1) = 0$	$\sigma''(l_b)/\alpha^U - l_b$	$\sigma'''(d_2)$, $\tau_{l_b,y_3}(\sigma'', u_b, d_2) = 0$	$\sigma''(u_b)/\alpha^L - u_b$
$\mathcal{R}_2$	$\mathcal{R}_3$	$\sigma'''(d_1)$, $\tau_{l_b,y_2}(\sigma'', u_b, d_1) = 0$	$\sigma''(u_b)/\alpha^U - u_b$	$\sigma'''(d_2)$, $\tau_{y_2,u_b}(\sigma'', l_b, d_2) = 0$	$\sigma''(l_b)/\alpha^L - l_b$
$\mathcal{R}_1$	$\mathcal{R}_3$	$\alpha\sigma'''(d_1) + (1-\alpha)\sigma'''(d_2)$, $\tau_{l_b,y_{\max}}(\sigma'', l_b, d_1) = 0$, $\tau_{y_{\max},u_b}(\sigma'', u_b, d_2) = 0$	$\alpha\beta_1^U + (1-\alpha)\beta_2^U$, $\beta_1^U = \sigma''(l_b)/\sigma'''(d_1) - l_b$, $\beta_2^U = \sigma''(u_b)/\sigma'''(d_2) - u_b$	$\sigma'''(d_3)$, $\tau_{y_1,u_b}(\sigma', l_b, d_3) = 0$	$\sigma''(l_b)/\alpha^L - l_b$
$\mathcal{R}_2$	$\mathcal{R}_4$	$\sigma'''(d_1)$, $\tau_{l_b,y_2}(\sigma', u_b, d_1) = 0$	$\sigma''(u_b)/\alpha^U - u_b$	$\alpha\sigma'''(d_2) + (1-\alpha)\sigma'''(d_3)$, $\tau_{l_b,y_{\min}}(\sigma'', l_b, d_2) = 0$, $\tau_{y_{\min},u_b}(\sigma'', u_b, d_3) = 0$	$\alpha\beta_1^L + (1-\alpha)\beta_2^L$, $\beta_1^L = \sigma''(l_b)/\sigma'''(d_2) - l_b$, $\beta_2^L = \sigma''(u_b)/\sigma'''(d_3) - u_b$
$\mathcal{R}_1$	$\mathcal{R}_4$	$\alpha\sigma'''(d_1) + (1-\alpha)\sigma'''(d_2)$, $\tau_{l_b,y_{\max}}(\sigma'', l_b, d_1) = 0$, $\tau_{y_{\max},u_b}(\sigma'', u_b, d_2) = 0$	$\alpha\beta_1^U + (1-\alpha)\beta_2^U$, $\beta_1^U = \sigma''(l_b)/\sigma'''(d_1) - l_b$, $\beta_2^U = \sigma''(u_b)/\sigma'''(d_2) - u_b$	$\alpha\sigma'''(d_3) + (1-\alpha)\sigma'''(d_4)$, $\tau_{l_b,y_{\min}}(\sigma'', l_b, d_3) = 0$, $\tau_{y_{\min},u_b}(\sigma'', u_b, d_4) = 0$	$\alpha\beta_1^L + (1-\alpha)\beta_2^L$, $\beta_1^L = \sigma''(l_b)/\sigma'''(d_3) - l_b$, $\beta_2^L = \sigma''(u_b)/\sigma'''(d_4) - u_b$

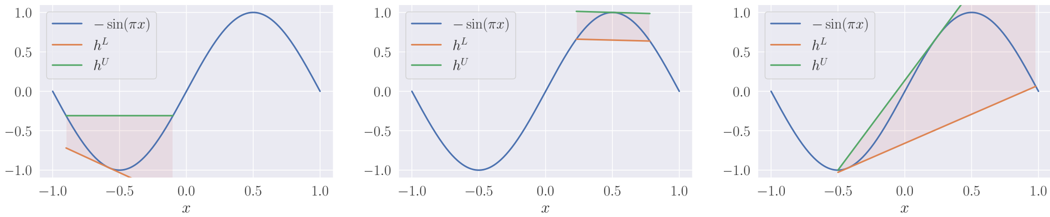


(a) $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_3$ (b) $l_b \in \mathcal{R}_2$ and $u_b \in \mathcal{R}_4$ (c) $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_4$

Figure 8: *Relaxing* $\sigma''(y) = -2\tanh(y)\left(1 - \tanh^2(y)\right)$: examples of the linear relaxations of σ'' for different sets of l_b and u_b.

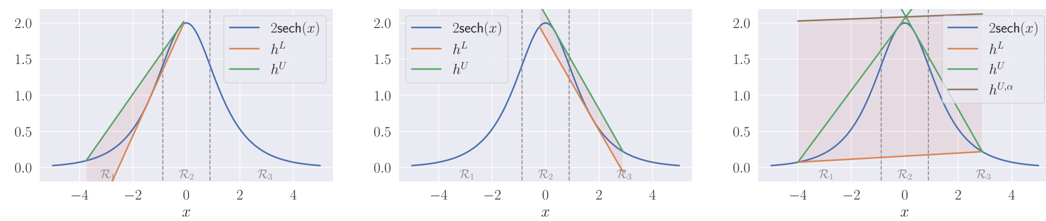
if $l_b \in \mathcal{R}_i$ and $u_b \in \mathcal{R}_{i+1}$ for $i \in \{1, 2, 3\}$, then we are also in the same situation as the adjacent regions of different convexity from σ', so we use exactly the same relaxation.

We are left with three cases where l_b and u_b are in non-adjacent regions. For $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_3$, we are in the same scenario as in the bounding of σ', since $\mathcal{R}_1$ and $\mathcal{R}_3$ are convex regions separated by a concave one. In that case we follow the bounding procedure outlined before for σ' - see Figure 8a for an example of it applied in this setting. For the case where $l_b \in \mathcal{R}_2$ and $u_b \in \mathcal{R}_4$, we are in an analogous case where $\mathcal{R}_2$ and $\mathcal{R}_4$ are concave regions separated by a convex one. As such, we



(a) $l_b \leq 0$ and $u_b \leq 0$ (b) $l_b \geq 0$ and $u_b \geq 0$ (c) $l_b \leq 0$ and $u_b \geq 0$

Figure 9: *Relaxing* $-\sin(\pi x)$: examples of the linear relaxations for different sets of l_b and u_b.



(a) $l_b \leq 0$ and $u_b \leq 0$ (b) $l_b \geq 0$ and $u_b \geq 0$ (c) $l_b \leq 0$ and $u_b \geq 0$

Figure 10: *Relaxing* $2sech(x)$: examples of the linear relaxations for different sets of l_b and u_b.

consider the two valid lower bounds computed previously for $l_b \in \mathcal{R}_2$ and $u_b \in \mathcal{R}_3$, and $l_b \in \mathcal{R}_3$ and $u_b \in \mathcal{R}_4$. To obtain these bounds, we shift the upper bound in the first case to $\arg\min \sigma''(y)$, and the lower bound in the same case to the same value (see h^L in Figure 8b). As our bounding requires a single h^L, we take a convex combination of the two bounds, $h^{L,\alpha}$. For the upper bound, we simply assume l_b is in a concave region while u_b is in a convex region, and take the tangent at d for $\arg\max \sigma''(y) \geq d \leq 0$ (see h^U in Figure 8b). Finally, we are left with the case where $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_4$. In that case, we take the upper bound lines from the case where $l_b \in \mathcal{R}_1$ and $u_b \in \mathcal{R}_3$, and the lower bound ones from where $l_b \in \mathcal{R}_2$ and $u_b \in \mathcal{R}_4$. As before, given the requirement of one lower and upper bound functions, we take a convex combination of both in $h^{L,\alpha}$ and $h^{U,\alpha}$, respectively. See Figure 8c for a visual representation.

G Linear lower and upper bounding nonlinear functions

Throughout, we assume the function's input (x) is lower bounded by l_b and upper bounded by u_b (*i.e.*, $l_b \leq x \leq u_b$), and define the upper bound line as $h^U(x) = \alpha^U(x + \beta^U)$, and the lower bound line as $h^L(x) = \alpha^L(x + \beta^L)$. For the sake of brevity, we define for a function $h : \mathbb{R} \to \mathbb{R}$, and points $p, d \in \mathbb{R}$ the function $\tau(h, p, d) = {}^{(h(p)-h(d))}/_{(p-d)} - h'(d)$. This is useful as for a given h and p, if there exists a $d \in [d_l, d_u]$, such that $\tau_{d_l,d_u}(h, p, d) = 0$, then $h'(d)$ is the slope of a tangent line to h that passes through p and d.

G.1 Case study: $-\sin(\pi x)$ for $x \in [-1, 1]$

As in Appendix F, we observe the convexity of the function $-\sin(\pi x)$ for $x \in [-1, 1]$, noticing that the function is convex for $x \leq 0$ and concave for $x \geq 0$. For $l_b \leq u_b \leq 0$ we let h^U be the line that connects l_b and u_b, and for an arbitrary $d \in [l_b, u_b]$ we let h^L be the tangent line at that point. Similarly, for $0 \leq l_b \leq u_b$ we let h^L be the line that connects l_b and u_b, and for an arbitrary $d \in [l_b, u_b]$ we let h^U be the tangent line at that point. For the last case where $l_b \leq 0 \leq u_b$, we let h^U be the tangent line at $d_1 \geq 0$ that passes through $(l_b, \sigma(l_b))$, and h^L be the tangent line at $d_2 \leq 0$ that passes through $(u_b, \sigma(u_b))$. Given the similarity of to the $\tanh$ bounds from Zhang et al. (2018), we omit a summary table, but present 3 examples of the possible cases in Figure 9.

G.2 CASE STUDY: 2SECH(x) FOR $x \in [-5, 5]$

We start by observing that the function 2sech(x) is similar to the derivative of $\tanh$, whose relaxation we presented in Appendix F. By inspecting it's derivative, $f'(x) = 2\text{sech}(x)\tanh(x)$, we conclude that there are two inflection points at $x_1 = \max f'(x)$ and $x_2 = \min f'(x)$, leading to three different regions: $x \in]-\infty, x_1]$ ($\mathcal{R}_1$, the first convex region), $x \in]x_1, x_2]$ ($\mathcal{R}_2$, the concave region), and $x \in]x_2, +\infty[$ ($\mathcal{R}_3$, the second convex region). As a result, there are 6 combinations for the location of l_b and u_b which must be resolved. This is exactly the same case as the first derivative of $\tanh$, simply with x_1 and x_2 instead of y_1 and y_2. Due to the similarities, we can use exactly the same relaxations as presented in Table 5. We present visual examples of 3 cases of this relaxation in Figure 10.

H FURTHER DETAILS ON GREEDY INPUT BRANCHING

In Section 5.3 we motivated and described at a high-level greedy input branching. In the following we provide a step-by-step analysis of Algorithm 1.

We start by initializing a lower and upper bound list of pairs $\mathcal{B}$ (line 3) as well as a list for storing the maximum error between the empirical and certified bounds $\mathcal{B}_\Delta$ (line 4). To initialize them (line 7 and 8), we first compute the empirical lower and upper bounds across the domain by sampling N_s points within the full domain $\mathcal{C}$ using $\text{SAMPLE}(\mathcal{C}, N_s)$ and evaluating the function h there (line 5) yielding $\hat{h}_{lb}$ and $\hat{h}_{ub}$, as well as the first version of the certified lower and upper bounds using ∂-CROWN on h (line 6) yielding h_{lb}, h_{ub}. Next, we pop from $\mathcal{B}$ and $\mathcal{B}_\Delta$ as C_i the interval which has the maximum error between the empirical and certified bounds (line 10), which we then proceed to split into N_d parts following a policy defined by DOMAINSPLIT (line 11). Importantly, DOMAINSPLIT must be complete, i.e., it must be that $C_i = \cup C'$. For each of those split subdomains $\mathcal{C}'$ we compute new bounds using ∂-CROWN (line 12) and add this subdomain along with its bounds and error to the empirical estimates to $\mathcal{B}$ and $\mathcal{B}_\Delta$, respectively (line 13 and 14). This process is repeated using the updated lists until the branching budget is spent, at which point the global lower bound is the minimum of all of lower bounds in $\mathcal{B}$ (defined as the list $\mathcal{B}_0$), and the global upper bound is the maximum of all upper bounds in $\mathcal{B}$ (defined as the list $\mathcal{B}_1$). These are computed in line 17. This algorithm is greedy as increasing the branching budget is expected to improve the bounds, since ∂-CROWN's bounds are guaranteed to monotonically decrease with smaller input domains.