arXiv:2211.15641v1 [math.OC] 28 Nov 2022

Blow up phenomena for gradient descent optimization methods in the training of artificial neural networks

Davide Gallon[1], Arnulf Jentzen[2,3], and Felix Lindner[4]

[1]University of Padua, Italy, e-mail: davide.gallon@studenti.unipd.it

[2]School of Data Science and Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong, Shenzhen, China, e-mail: ajentzen@cuhk.edu.cn

[3]Applied Mathematics: Institute for Analysis and Numerics, University of Münster, Germany, e-mail: ajentzen@uni-muenster.de

[4]Faculty of Mathematics and Natural Sciences, University of Kassel, Germany, e-mail: lindner@mathematik.uni-kassel.de

November 29, 2022

Abstract

In this article we investigate blow up phenomena for gradient descent optimization methods in the training of artificial neural networks (ANNs). Our theoretical analysis is focused on shallow ANNs with one neuron on the input layer, one neuron on the output layer, and one hidden layer. For ANNs with ReLU activation and at least two neurons on the hidden layer we establish the existence of a target function such that there exists a lower bound for the risk values of the critical points of the associated risk function which is strictly greater than the infimum of the image of the risk function. This allows us to demonstrate that every gradient flow trajectory with an initial risk smaller than this lower bound diverges. Furthermore, we analyze and compare various popular types of activation functions with regard to the divergence of gradient flow trajectories and gradient descent trajectories in the training of ANNs and with regard to the closely related question concerning the existence of global minimum points of the risk function.

Contents

1 Introduction

While artificial neural networks (ANNs) are widely used and increasingly popular in a large variety of scientific and industrial applications, training methods for ANNs are still far from being well-understood from an analytical perspective.

In the training of ANNs one is ultimately interested in mimimizing the true risk, i.e., the expected loss of the realization function associated to the ANN. A natural direction of research aiming at a better theoretical understanding of such optimization problems concerns the analysis of the associated gradient flow (GF) differential equations. Loosely speaking, each GF trajectory represents a path of steepest descent in the risk landscape. In order to render the concept of GFs useful for the practical training of ANNs, at least two types of approximation have to be taken into account. First, the unknown true risk function has to be approximated by an empirical risk function based on the training data at hand. Second, the continuous-time GF has to be approximated by a discrete-time gradient descent (GD) optimization scheme. Discretization parameters associated to these types of approximation are the size of the training data set and the learning rate respectively. A possible further approximation regarding the gradient of the empirical risk function by means of Monte Carlo estimation leads to the class of stochastic GD optimization methods, involving the batch size as an additional approximation parameter.

In order to ensure that a GD optimization scheme produces trajectories which lie suitably close to the corresponding GF trajectories associated to the true risk, the discretization parameters, say, the learning rate and the size of the training data set, have to be chosen sufficiently small and large respectively. However, just how small or large the discretization parameters need to be chosen is generally relative to the object to be approximated, i.e., relative to the GF. In particular, if a GF trajectory is such that the norm of the ANN parameter vector specifying the realization function diverges to infinity, problems concerning an adequate choice of the discretization parameters specifying the approximation of the GF may arise. As a matter of fact, available results in the research literature pertaining to the convergence analysis of GFs and GD type optimization algorithms are typically based, either explicitly or implicitly, on suitable

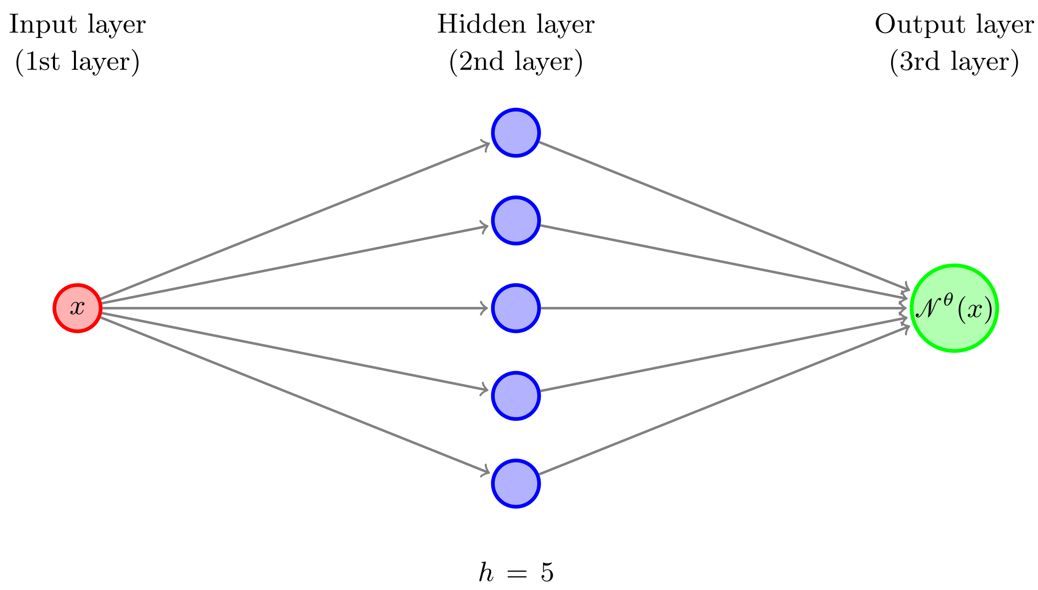

Figure 1: Graphical illustration of the shallow ANN architecture considered in Theorems 1.1–1.4 in the special case of $h = 5$ neurons on the hidden layer. In this situation every ANN parameter vector $\theta \in \mathbb{R}^{3h+1} = \mathbb{R}^{16}$ specifies a realization function $\mathscr{N}^\theta \colon \mathbb{R} \to \mathbb{R}$ which depends on the choice of the activation function $\mathbb{A} \colon \mathbb{R} \mapsto \mathbb{R}$ and maps the scalar input $x \in [\mathscr{a}, \mathscr{b}]$ to the scalar output $\mathscr{N}^\theta(x) = \theta_{3h+1} + \sum_{i=1}^{h} \theta_{2h+i} \mathbb{A}(\theta_{h+i} + \theta_i x) \in \mathbb{R}$.

boundedness assumptions concerning the GF and GD trajectories; see the literature overview in Subsection 1.4 below.

It is the contribution of this article to uncover and analyze situations where the standard boundedness assumptions on GF and GD trajectories fail to hold and blow up phenomena occur instead. In our theoretical analysis we focus on shallow ANNs with one neuron in the input layer, one neuron in the output layer, and one hidden layer made up of an arbitrary number of neurons. In this introductory section we present four selected key results of this article to elucidate the scope of the study. The first key result, see Theorem 1.1 below, concerns ANNs with ReLU activation function and establishes the existence of a target function such that every GF trajectory with an initial risk below a certain threshhold diverges, assuming that the hidden layer consists of at least two neurons. In the second key result of this article, see Theorem 1.2 below, we show for various popular types of activation functions that there exist target functions such that GF and GD trajectories diverge whenever the associated risk values satisfy certain asymptotic optimality conditions and the hidden layer consists of at least two neurons. From an analytical perspective, blow up phenomena in the training of ANNs are closely connected to the question whether there exist global minimum points of the risk function. Further key results of this article, see Theorem 1.3 and Theorem 1.4 below, therefore concern the existence and non-existence of global minima of the risk function depending on the choice of the activation function and a superior role played in this regard by piecewise affine activation functions.

In the remainder of this introductory section we provide the precise statements of the selected results mentioned above together with some additional comments in Subsections 1.1–1.3, we give a short overview of related research literature in Subsection 1.4, and we outline the overall structure of this article in Subsection 1.5.

1.1 Blow up phenomena

Before stating the key results we shortly comment on the mathematical setting and the employed notation. We consider ANNs with one neuron on the input layer, one neuron on the output layer, and one hidden layer made up of $h \in \mathbb{N}$ neurons; compare the illustration in Figure 1. The specification of such an ANN involves $2h$ real weight parameters and $h+1$ real bias parameters, so that the overall number of ANN parameters amounts to $3h+1$. The real numbers $\mathscr{a} \in \mathbb{R}$, $\mathscr{b} \in (\mathscr{a}, \infty)$ define the domain of the target function $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ and thus the interval on which the realization function associated to the ANN should approximate v.

Our first main result on blow up phenomena in the training of ANNs formulated in Theorem 1.1 below focuses on ANNs with ReLU activation function $\mathbb{R} \ni x \mapsto \max\{x, 0\} \in \mathbb{R}$ and establishes the existence of a non-decreasing target function $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ such that for every choice $h \in \mathbb{N}\backslash\{1\}$ of the number of hidden neurons there exists a positive threshold value such that every GF trajectory $\Theta\colon [0, \infty) \to \mathbb{R}^{3h+1}$ with a initial risk smaller than this threshold value diverges to infinity in the sense that $\liminf_{t\to\infty} \|\Theta_t\| = \infty$. Here we denote for every non-decreasing $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ and every $h \in \mathbb{N}$ by $\mathcal{L}^{v,h}\colon \mathbb{R}^{3h+1} \to \mathbb{R}$ the risk function associated to the ANN measuring how well the realization function approximates the target function. Theorem 1.1 is a slightly simplified version of a more general result in Theorem 2.48 in Subsection 2.6 below. Note that the lack of differentiability of the ReLU activation function at zero entails a lack of differentiability of the risk function associated to the ANN, so that we need to work with an appropriate generalized gradient of the risk function to be able to specify the GF. While the general statement in Theorem 2.48 is based on a generalized gradient defined in terms of continuously differentiable approximations of the ReLU activation function, we simplify the exposition in Theorem 1.1 below by employing the less involved concept of the left gradient of the risk function, denoted for every non-decreasing $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ and every $h \in \mathbb{N}$ by $\mathcal{G}^{v,h}\colon \mathbb{R}^{3h+1} \to \mathbb{R}^{3h+1}$. In order to ensure that both concepts of generalized gradients coincide, we additionally assume in Theorem 1.1 that $\mathscr{a} \geq 0$. This additional assumption is not needed in the general statement in Theorem 2.48 .

Theorem 1.1. *Let $\mathscr{a} \in [0, \infty)$, $\mathscr{b} \in (\mathscr{a}, \infty)$ and for every non-decreasing $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ and every $h \in \mathbb{N}$ let $\mathcal{L}^{v,h}\colon \mathbb{R}^{3h+1} \to \mathbb{R}$ satisfy for all $\theta = (\theta_1, \ldots, \theta_{3h+1}) \in \mathbb{R}^{3h+1}$ that*

$$\mathcal{L}^{v,h}(\theta) = \int_{\mathscr{a}}^{\mathscr{b}} \big(v(x) - \theta_{3h+1} - \textstyle\sum_{i=1}^{h} \theta_{2h+i}\big[\max\{\theta_{h+i} + \theta_i x, 0\}\big]\big)^2 \, \mathrm{d}x \tag{1.1}$$

and let $\mathcal{G}^{v,h}\colon \mathbb{R}^{3h+1} \to \mathbb{R}^{3h+1}$ be the left gradient[1] of $\mathcal{L}^{v,h}$. Then[2] there exists a non-decreasing $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ such that for all $h \in \mathbb{N}\backslash\{1\}$ there exists $\varepsilon \in (0, \infty)$ such that for all $\Theta \in C([0, \infty), \mathbb{R}^{3h+1})$ with $\forall\, t \in [0, \infty)\colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}^{v,h}(\Theta_s)\, \mathrm{d}s$ and $\mathcal{L}^{v,h}(\Theta_0) < \varepsilon + \inf_{\theta\in\mathbb{R}^{3h+1}} \mathcal{L}^{v,h}(\theta)$ it holds that $\liminf_{t\to\infty} \|\Theta_t\| = \infty$.

Theorem 1.1 is a direct consequence of Theorem 2.48 in Subsection 2.6 below. Theorem 2.48, in turn, follows from combining Proposition 2.45 in Subsection 2.6 below, which is a slight modification of [20, Theorem 1.3], and Theorem 2.47 in Subsection 2.6 below, which is one of the main results of this paper. Roughly speaking, Proposition 2.45 states that every GF trajectory which does not diverge to infinity converges to a critical point of the risk function and the risk values associated to the GF trajectory converge to the risk of the critical point. Theorem 2.47, rather, asserts for the case of the indicator function $\mathbb{1}_{((\mathscr{a}+\mathscr{b})/2, \mathscr{b}]}\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ being the target function that there exists a lower bound $\varepsilon \in (0, \infty)$ for the risk of critical points. In

[1] For every $h \in \mathbb{N}$ let $e_1^{(h)}, \ldots, e_{3h+1}^{(h)} \in \mathbb{R}^{3h+1}$ satisfy $e_1^{(h)} = (1, 0, \ldots, 0)$, $\ldots$, $e_{3h+1}^{(h)} = (0, \ldots, 0, 1)$. Observe that for all non-decreasing $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ and all $h \in \mathbb{N}$, $\theta \in \mathbb{R}^{3h+1}$ it holds that $\mathcal{G}^{v,h}(\theta) = (\lim_{\varepsilon\nearrow 0}(\mathcal{L}^{v,h}(\theta + \varepsilon e_1^{(h)}) - \mathcal{L}^{v,h}(\theta))\varepsilon^{-1}, \ldots, \lim_{\varepsilon\nearrow 0}(\mathcal{L}^{v,h}(\theta + \varepsilon e_{3h+1}^{(h)}) - \mathcal{L}^{v,h}(\theta))\varepsilon^{-1})$.

[2] Note that the function $\|\cdot\|\colon (\cup_{n\in\mathbb{N}}\mathbb{R}^n) \to \mathbb{R}$ satisfies for all $n \in \mathbb{N}$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ that $\|x\| = (\sum_{i=1}^{n} |x_i|^2)^{1/2}$.

this situation, Theorem 2.48 infers the divergence of every GF trajectory with an initial risk smaller than ε, which readily implies Theorem 1.1 in view of the fact that the different concepts of generalized gradients used in these results are compatible and in view of the fact that the infimum of the risk values associated to the considered target function f equals zero in the case of at least two neurons in the hidden layer; compare, e.g., Proposition 3.2.

Our second main result on blow up phenomena in the training of ANNs formulated in Theorem 1.2 below treats various types of activation functions and establishes for each type the existence of a target function such that all GF and GD trajectories which fulfill certain asymptotic optimality conditions w.r.t. the associated risk values diverge to infinity. Observe that the family of functions $A_{k,\gamma}\colon \mathbb{R} \to \mathbb{R}$, $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$, appearing in Theorem 1.2 is such that depending on the choice of $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ we have that $A_{k,\gamma}\colon \mathbb{R} \to \mathbb{R}$ refers to the softsign activation function in the case $k < -5$, the arctangent activation function in the case $k = -5$, the inverse square root unit activation function with parameter $\xi \in (0,3)$ in the case $k = -4$, the exponential linear unit activation function in the case $k = -3$, the hyperbolic tangent activation function in the case $k = -2$, the logistic activation function in the case $k = -1$, the softplus activation function in the case $k = 0$, the ReLU activation function in the case $k = 1$, $\gamma = 0$, the leaky ReLU activation function with parameter γ in the case $k = 1$, $\gamma \in (0,1)$, and the rectified power unit activation function with exponent k in the case $k > 1$. Here we denote for every Lebesgue square integrable target function $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ and every $h \in \mathbb{N}$, $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ the associated risk function and its left gradient by $\mathcal{L}^{v,h}_{k,\gamma}\colon \mathbb{R}^{3h+1} \to \mathbb{R}$ and $\mathcal{G}^{v,h}_{k,\gamma}\colon \mathbb{R}^{3h+1} \to \mathbb{R}^{3h+1}$ respectively. Similar to Theorem 1.1 above, the assertions in Theorem 1.2 are slightly simplified versions of more general results employing a generalized gradient of the risk function defined in terms of continuously differentiable approximations of the activation function in Section 3 below.

Theorem 1.2. *Let $\mathscr{a} \in [0,\infty)$, $\mathscr{b} \in (\mathscr{a},\infty)$, $\xi \in (0,3)$, for every $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ let $A_{k,\gamma}\colon \mathbb{R} \to \mathbb{R}$, satisfy for all $x \in \mathbb{R}$ that*

$$A_{k,\gamma}(x) = \begin{cases} x(1+|x|)^{-1}, & : k < -5 \\ \arctan(x) & : k = -5 \\ x(1+\xi x^2)^{-1/2} & : k = -4 \\ x\mathbb{1}_{(0,\infty)}(x) + (\exp(x)-1)\mathbb{1}_{(-\infty,0]}(x) & : k = -3 \\ (\exp(x)-\exp(-x))(\exp(x)+\exp(-x))^{-1} & : k = -2 \\ (1+\exp(-x))^{-1} & : k = -1 \\ \ln(1+\exp(x)) & : k = 0 \\ (\max\{x,0\})^k + \min\{\gamma x, 0\} & : k > 0, \end{cases} \tag{1.2}$$

and for every Lebesgue square integrable $v\colon [\mathscr{a},\mathscr{b}] \to \mathbb{R}$ and every $h \in \mathbb{N}$, $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ let $\mathcal{L}^{v,h}_{k,\gamma}\colon \mathbb{R}^{3h+1} \to \mathbb{R}$ satisfy for all $\theta = (\theta_1, \dots, \theta_{3h+1}) \in \mathbb{R}^{3h+1}$ that

$$\mathcal{L}^{v,h}_{k,\gamma}(\theta) = \int_{\mathscr{a}}^{\mathscr{b}} \big(v(x) - \theta_{3h+1} - \textstyle\sum_{i=1}^h \theta_{2h+i}\big[A_{k,\gamma}(\theta_{h+i} + \theta_i x)\big]\big)^2 \,\mathrm{d}x \tag{1.3}$$

and let $\mathcal{G}^{v,h}_{k,\gamma}\colon \mathbb{R}^{3h+1} \to \mathbb{R}^{3h+1}$ be the left gradient of $\mathcal{L}^{v,h}_{k,\gamma}$. Then

(i) there exists a polynomial $v\colon [\mathscr{a},\mathscr{b}] \to \mathbb{R}$ such that for all $h \in \mathbb{N}\backslash\{1\}$, $k \in \mathbb{Z}\backslash\mathbb{N}$, $\Theta \in C([0,\infty), \mathbb{R}^{3h+1})$ with $\liminf_{t\to\infty} \mathcal{L}^{v,h}_{k,0}(\Theta_t) = \inf_{\theta\in\mathbb{R}^{3h+1}} \mathcal{L}^{v,h}_{k,0}(\theta)$ and $\forall\, t \in [0,\infty)\colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}^{v,h}_{k,0}(\Theta_s)\,\mathrm{d}s$ it holds that $\liminf_{t\to\infty} \|\Theta_t\| = \infty$,

(ii) there exists a polynomial $v\colon [\mathscr{a},\mathscr{b}] \to \mathbb{R}$ such that for all $h \in \mathbb{N}\backslash\{1\}$, $k \in \mathbb{Z}\backslash\mathbb{N}$ and all $\Theta\colon \mathbb{N}_0 \to \mathbb{R}^{3h+1}$ with $\limsup_{n\to\infty} \mathcal{L}^{v,h}_{k,0}(\Theta_n) = \inf_{\theta\in\mathbb{R}^{3h+1}} \mathcal{L}^{v,h}_{k,0}(\theta)$ it holds that $\liminf_{n\to\infty} \|\Theta_n\| = \infty$,

(iii) *for all* $k \in \mathbb{Z}\backslash\{1\}$ *there exists a Lipschitz continuous* $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ *such that for all* $h \in \mathbb{N}\backslash\{1\}$, $\Theta \in C([0,\infty), \mathbb{R}^{3h+1})$ *with* $\liminf_{t\to\infty} \mathcal{L}^{v,h}_{k,0}(\Theta_t) = \inf_{\theta\in\mathbb{R}^{3h+1}} \mathcal{L}^{v,h}_{k,0}(\theta)$ *and* $\forall\, t \in [0,\infty)\colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}^{v,h}_{k,0}(\Theta_s)\,\mathrm{d}s$ *it holds that* $\liminf_{t\to\infty} \|\Theta_t\| = \infty$,

(iv) *for all* $k \in \mathbb{Z}\backslash\{1\}$ *there exists a Lipschitz continuous* $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ *such that for all* $h \in \mathbb{N}\backslash\{1\}$ *and all* $\Theta\colon \mathbb{N}_0 \to \mathbb{R}^{3h+1}$ *with* $\limsup_{n\to\infty} \mathcal{L}^{v,h}_{k,0}(\Theta_n) = \inf_{\theta\in\mathbb{R}^{3h+1}} \mathcal{L}^{v,h}_{k,0}(\theta)$ *it holds that* $\liminf_{n\to\infty} \|\Theta_n\| = \infty$,

(v) *there exists a non-decreasing* $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ *such that for all* $h \in \mathbb{N}\backslash\{1\}$, $\gamma \in \mathbb{R}\backslash\{1\}$, $\Theta \in C([0,\infty), \mathbb{R}^{3h+1})$ *with* $\liminf_{t\to\infty} \mathcal{L}^{v,h}_{1,\gamma}(\Theta_t) = \inf_{\theta\in\mathbb{R}^{3h+1}} \mathcal{L}^{v,h}_{1,\gamma}(\theta)$ *and* $\forall\, t \in [0,\infty)\colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}^{v,h}_{1,\gamma}(\Theta_s)\,\mathrm{d}s$ *it holds that* $\liminf_{t\to\infty} \|\Theta_t\| = \infty$, *and*

(vi) *there exists a non-decreasing* $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ *such that for all* $h \in \mathbb{N}\backslash\{1\}$, $\gamma \in \mathbb{R}\backslash\{1\}$ *and all* $\Theta\colon \mathbb{N}_0 \to \mathbb{R}^{3h+1}$ *with* $\limsup_{n\to\infty} \mathcal{L}^{v,h}_{1,\gamma}(\Theta_n) = \inf_{\theta\in\mathbb{R}^{3h+1}} \mathcal{L}^{v,h}_{1,\gamma}(\theta)$ *it holds that* $\liminf_{n\to\infty} \|\Theta_n\| = \infty$.

Item (i), item (iii), and item (v) in Theorem 1.2 are direct consequences of Corollary 3.29, Corollary 3.30, and Corollary 3.31 in Subsection 3.8 below. Corollary 3.29, Corollary 3.30, and Corollary 3.31, in turn, are based on non-existence results concerning global minima of the risk function, compare Theorem 1.3 below, and an abstract divergence result for GF trajectories in Lemma 3.28 in Subsection 3.8 below. Item (ii), item (iv), and item (vi) in Theorem 1.2 are direct consequences of Corollary 3.33, Corollary 3.35, and Corollary 3.37 in Subsection 3.9 below. Corollary 3.33, Corollary 3.35, and Corollary 3.37, in turn, are based on non-existence results concerning global minima of the risk function, compare Theorem 1.3 below, and an abstract divergence result for GD trajectories in Lemma 3.32 in Subsection 3.9 below. Related results can be found in [30, Proposition 3.6].

1.2 Existence and non-existence of global minima

The analysis of blow up phenomena for GFs and GD optimization methods in the training of ANNs is closely related to the question whether there exist global minimum points of the risk function associated to the ANN. In fact, the divergence results in Theorem 1.2 above heavily rely on the non-existence of global minimum points of the risk function for certain target functions. In our third key result formulated in Theorem 1.3 below we consider the activation functions introduced in Theorem 1.2 and establish the existence of several target functions $v\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ such that for every choice $h \in \mathbb{N}\backslash\{1\}$ of the number of hidden neurons and for specific choices of $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ the set of global minimum points $\mathscr{M}^{v,h}_{k,\gamma} = \{\theta \in \mathbb{R}^{3h+1}\colon \mathcal{L}^{v,h}_{k,\gamma}(\theta) = \inf_{\vartheta\in\mathbb{R}^{3h+1}} \mathcal{L}^{v,h}_{k,\gamma}(\vartheta)\}$ is empty. In particular, in the case of softsign, arctangent, inverse square root unit, exponential linear unit, hyperbolic tangent, standard logistic, and softplus activation we employ a polynomial target function, in the case of rectified power unit activation we employ a Lipschitz continuous target function, and in the case of ReLU and leaky ReLU activation we employ a non-decreasing target function.

Theorem 1.3. *Let* $\mathscr{a} \in \mathbb{R}$, $\mathscr{b} \in (\mathscr{a}, \infty)$, $\xi \in (0,3)$, *for every* $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ *let* $A_{k,\gamma}\colon \mathbb{R} \to \mathbb{R}$

satisfy for all $x \in \mathbb{R}$ *that*

$$A_{k,\gamma}(x) = \begin{cases} x(1+|x|)^{-1}, & : k < -5 \\ \arctan(x) & : k = -5 \\ x(1+\xi x^2)^{-1/2} & : k = -4 \\ x\mathbb{1}_{(0,\infty)}(x) + (\exp(x)-1)\mathbb{1}_{(-\infty,0]}(x) & : k = -3 \\ (\exp(x)-\exp(-x))(\exp(x)+\exp(-x))^{-1} & : k = -2 \\ (1+\exp(-x))^{-1} & : k = -1 \\ \ln(1+\exp(x)) & : k = 0 \\ (\max\{x,0\})^k + \min\{\gamma x, 0\} & : k > 0, \end{cases} \tag{1.4}$$

and for every measurable $v\colon [\mathscr{a},\mathscr{b}] \to \mathbb{R}$ *and every* $h \in \mathbb{N}$, $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ *let* $\mathcal{L}^{v,h}_{k,\gamma}\colon \mathbb{R}^{3h+1} \to [0,\infty]$ *satisfy for all* $\theta = (\theta_1, \ldots, \theta_{3h+1}) \in \mathbb{R}^{3h+1}$ *that*

$$\mathcal{L}^{v,h}_{k,\gamma}(\theta) = \int_{\mathscr{a}}^{\mathscr{b}} \bigl(v(x) - \theta_{3h+1} - \textstyle\sum_{i=1}^h \theta_{2h+i}\bigl[A_{k,\gamma}(\theta_{h+i} + \theta_i x)\bigr]\bigr)^2 \,\mathrm{d}x \tag{1.5}$$

and let $\mathscr{M}^{v,h}_{k,\gamma} \subseteq \mathbb{R}^{3h+1}$ *satisfy* $\mathscr{M}^{v,h}_{k,\gamma} = \{\theta \in \mathbb{R}^{3h+1}\colon \mathcal{L}^{v,h}_{k,\gamma}(\theta) = \inf_{\vartheta \in \mathbb{R}^{3h+1}} \mathcal{L}^{v,h}_{k,\gamma}(\vartheta)\}$. *Then*

(i) there exists a polynomial $v\colon [\mathscr{a},\mathscr{b}] \to \mathbb{R}$ *such that* $\cup_{h\in\mathbb{N}\backslash\{1\},\, k\in\mathbb{Z}\backslash\mathbb{N}} \mathscr{M}^{v,h}_{k,0} = \varnothing$,

(ii) it holds for all $k \in \mathbb{Z}\backslash\{1\}$ *that there exists a Lipschitz continuous* $v\colon [\mathscr{a},\mathscr{b}] \to \mathbb{R}$ *such that* $\cup_{h\in\mathbb{N}\backslash\{1\}} \mathscr{M}^{v,h}_{k,0} = \varnothing$,

(iii) there exists a non-decreasing $v\colon [\mathscr{a},\mathscr{b}] \to \mathbb{R}$ *such that* $\cup_{h\in\mathbb{N}\backslash\{1\},\, \gamma\in\mathbb{R}\backslash\{1\}} \mathscr{M}^{v,h}_{1,\gamma} = \varnothing$, *and*

(iv) it holds for all Lipschitz continuous $v\colon [\mathscr{a},\mathscr{b}] \to \mathbb{R}$ *and all* $h \in \mathbb{N}$ *that* $\mathscr{M}^{v,h}_{1,0} \neq \varnothing$.

Item (i) and item (ii) in Theorem 1.3 follow directly from combining Lemma 3.11 in Subsection 3.3, Lemma 3.21 in Subsection 3.4, Lemma 3.23 in Subsection 3.5, Lemma 3.25 in Subsection 3.6, and Lemma 3.27 in Subsection 3.7 below. Item (iii) in Theorem 1.3 is a direct consequence of Lemma 3.6 in Subsection 3.2 below. The strategy in the proofs of these results is to identify the infimum of the image of the risk function and to consequently show that the set of global minimum points is empty. Item (iv) in Theorem 1.3 has been proven in [20, Theorem 1.1]. Related results can be found in [30, Theorem 3.3].

1.3 Superiority of piecewise affine activation functions with respect to the existence of minimum points

In practical applications of ANNs the choice of the activation functions is typically guided by heuristic arguments and numerical experiments. Theorem 1.3 above suggest from an analytic perspective a superiority of piecewise affine activation functions with respect to the existence of minimum points of the risk function. Indeed, the non-existence results for global minima of the risk function in item (i) and item (ii) in Theorem 1.3 are based on polynomial and Lipschitz continuous target functions and exclusively involve continuously differentiable activation functions, whereas the existence result for global minima of the risk function in item (iv) in Theorem 1.3 holds for all all Lipschitz continuous target functions and involves the piecewise affine ReLU activation function. This aspect is further highlighted in Theorem 1.4 below.

Theorem 1.4. *Let $h \in \mathbb{N}\backslash\{1\}$, $\mathscr{a} \in \mathbb{R}$, $\mathscr{b} \in (\mathscr{a}, \infty)$, $\xi \in (0,3)$, for every $k \in \mathbb{Z}$ let $A_k \in C(\mathbb{R}, \mathbb{R})$ satisfy for all $x \in \mathbb{R}$ that*

$$A_k(x) = \begin{cases} x(1+|x|)^{-1}, & : k < -5 \\ \arctan(x) & : k = -5 \\ x(1+\xi x^2)^{-1/2} & : k = -4 \\ x\mathbb{1}_{(0,\infty)}(x) + (\exp(x)-1)\mathbb{1}_{(-\infty,0]}(x) & : k = -3 \\ (\exp(x)-\exp(-x))(\exp(x)+\exp(-x))^{-1} & : k = -2 \\ (1+\exp(-x))^{-1} & : k = -1 \\ \ln(1+\exp(x)) & : k = 0 \\ (\max\{x,0\})^k & : k > 0, \end{cases} \tag{1.6}$$

let $k \in \mathbb{Z}$, and for every measurable $v \colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ let $\mathscr{L}^v \colon \mathbb{R}^{3h+1} \to [0, \infty]$ satisfy for all $\theta = (\theta_1, \ldots, \theta_{3h+1}) \in \mathbb{R}^{3h+1}$ that

$$\mathscr{L}^v(\theta) = \int_{\mathscr{a}}^{\mathscr{b}} \big(v(x) - \theta_{3h+1} - \textstyle\sum_{i=1}^h \theta_{2h+i}\big[A_k(\theta_{h+i} + \theta_i x)\big]\big)^2 \, \mathrm{d}x. \tag{1.7}$$

Then the following three statements are equivalent:

(i) It holds for every Lipschitz continuous $v \colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ that there exists $\theta \in \mathbb{R}^{3h+1}$ such that $\mathscr{L}^v(\theta) = \inf_{\vartheta \in \mathbb{R}^{3h+1}} \mathscr{L}^v(\vartheta)$.

(ii) It holds that $A_k \notin C^1(\mathbb{R}, \mathbb{R})$.

(iii) It holds that $k = 1$.

Theorem 1.4 is a direct consequence of Theorem 1.3 and the elementary observation that the ReLU activation function is the only activation function appearing in Theorem 1.4 which is not continuously differentiable.

1.4 Literature overview

Let us complement the presentation of the findings of our work by a short review of related research literature. Despite the lack of a full-fledged convergence analysis for GFs and GD optimization schemes in the training of ANNs in literature, there are several promising mathematical approaches. For the convergence of GFs and GD type methods in the case of convex target functions we refer, e.g, to [18, 6, 29] and the references mentioned therein. More complications are encountered in the case of non-convex problems: in principle there could be many local minima. For more details on abstract convergence results for GD and GF optimization methods we refer, e.g., to [2, 16, 25, 27, 7, 11].

Another promising direction of research considers the overparametrized regime, where the number of parameters of the model exceeds the number of training points; see, e.g., [19, 3, 4, 13, 14, 31, 26]. Under Lojasiewicz type assumptions convergence results for GD and GF type optimization schemes can be found, e.g, in [15, 25, 12, 1, 5]. A further interesting method is to consider only special target functions; see, e.g.,[22, 9] for a convergence analysis for GF and GD processes in the case of constant target functions and [21] for a convergence analysis for GF and GD processes in the training of ANNs with piecewise linear target functions.

For lower bounds and divergence results for GD and GF optimization methods we refer, e.g., to [17, 10, 28]. Results related to the findings of the present article can be found in [30, Section 3].

1.5 Structure of the article

The remainder of this article is structured as follows. In Section 2 we present the mathematical framework used to prove Theorem 1.1, we establish several properties for critical points of the risk function in order to find a strictly positive lower bound for the risk values of critical points, and we demonstrate that the considered GF trajectories blow up. In Section 3 we introduce the mathematical framework needed for the proof of Theorem 1.2 and Theorem 1.3, we establish the non-existence of global minima employing various target functions, and we finally verify the assiciated blow up phenomena, thus proving Theorem 1.3 and Theorem 1.2. In Section 4 we complement our findings be investigating the non-existence of global minima in the case in which the risk is defined using a discrete measure, the activation function is the standard logistic function, and there is one neuron in the hidden layer.

2 Blow up phenomena for gradient flows (GFs) in the training of artificial neural networks (ANNs) with ReLU activation

In this section we investigate blow up phenomena for GFs in the training of shallow ANNs with ReLU activation function in the case where the target function is given by the indicator function $\mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\mathscr{b}]}\colon [\mathscr{a},\mathscr{b}] \to \mathbb{R}$. In particular, in Theorem 2.48 in Subsection 2.6 below we demonstrate that every GF trajectory $\Theta\colon [0,\infty) \to \mathbb{R}^{3h+1}$ with an initial risk smaller than a certain threshold diverges to infinity in the sense that $\liminf_{t\to\infty} \|\Theta_t\| = \infty$. Theorem 1.1 in the introduction is a direct consequence of Theorem 2.48.

The two main ingredients in our proof of Theorem 2.48 are Proposition 2.45 and Theorem 2.47 in Subsection 2.6 below. Proposition 2.45 is a slight modification of [20, Theorem 1.3] and states that for every GF trajectory $\Theta\colon [0,\infty) \to \mathbb{R}^{3h+1}$ with $\liminf_{t\to\infty} \|\Theta_t\| < \infty$ there exists $\beta \in (0,\infty)$ such that the GF trajectory converges with order β to a critical point of the risk function and the risk values associated to the GF trajectory converge with order 1 to the risk of the critical point. Theorem 2.47 is one of the main results of this article and establishes a positive lower bound $\varepsilon \in (0,\infty)$ for the risk of critical points. In the proof of Theorem 2.48 we combine Proposition 2.45, Theorem 2.47, and the well-known fact that the risk values associated to a GF trajectory are non-increasing (see, e.g., [23, Lemma 3.1]) to conclude that every GF trajectory with an initial risk smaller than ε diverges to infinity. The positive lower bound for the risk of critical points in Theorem 2.47 is based on an analogous result for the specific case $[\mathscr{a},\mathscr{b}] = [0,1]$ in Lemma 2.46 in Subsection 2.6 in combination with an affine coordinate transformation. Lemma 2.46, in turn, is proved by induction w.r.t. the number of hidden neurons $h \in \mathbb{N}$ and relies on a series of auxiliary results in Subsections 2.2–2.5 below.

In Subsection 2.5 we provide in Proposition 2.42 and Lemma 2.44 the base step and the induction step for our proof of the positive lower bound for the risk of critical points in Lemma 2.46. More precisely, in Lemma 2.44 we establish a positive lower bound for the risk of critical points in the case in which all hidden neurons are active and no combination of parameters allows for a representation of the same realization function using less neurons. Lemma 2.44 is built on a detailed analysis of all possible parameter constellations of such critical points in Corollary 2.34, Corollary 2.35, Proposition 2.36, Lemma 2.37, Lemma 2.38, Corollary 2.39, Lemma 2.40, and Lemma 2.41 in Subsection 2.5. These results, in turn, employ properties of critical points of the risk function derived in Subsection 2.4 below and several elementary and well-known estimates and conclusions regarding specific integrals associated to the risk function provided in Subsection 2.2 and Subsection 2.3 below.

In Subsection 2.1 below we specify in Setting 2.1 the mathematical framework regarding the training of shallow ANNs with ReLU activation used repeatedly throughout this section. Note that here for every ANN parameter vector $\theta \in \mathbb{R}^{3h+1}$ and every index $j \in \{1,2,\dots,h\}$ associated to a hidden neuron with weight parameter $\mathfrak{w}^{\theta}_j \neq 0$ we have that the real number

$\mathfrak{q}_j^\theta \in \mathbb{R}$ represents a possible breakpoint of the piecewise affine realization function $\mathscr{N}_\infty^{h,\theta} \colon \mathbb{R} \to \mathbb{R}$ associated to the ANN.

Finally, in Subsection 2.7 and Subsection 2.8 below we complement our findings in this section by providing an explicit lower bound for the risk of critical points in the specific case of $h = 2$ hidden neurons in Lemma 2.52 in Subsection 2.7 and by providing a general upper bound for the norms of GF trajectories in Proposition 2.54 in Subsection 2.8.

2.1 Mathematical description of ANNs

Setting 2.1. *For every $h \in \mathbb{N}$ let $\mathfrak{d}_h \in \mathbb{N}$ satisfy $\mathfrak{d}_h = 3h+1$, for every $h \in \mathbb{N}$, $\theta = (\theta_1, \dots, \theta_{\mathfrak{d}_h}) \in \mathbb{R}^{\mathfrak{d}_h}$ let $\mathfrak{w}_0^\theta, \mathfrak{w}_1^\theta, \dots, \mathfrak{w}_{h+1}^\theta, \mathfrak{b}_1^\theta, \mathfrak{b}_2^\theta \dots, \mathfrak{b}_h^\theta, \mathfrak{v}_1^\theta, \mathfrak{v}_2^\theta, \dots, \mathfrak{v}_h^\theta, \mathfrak{c}^\theta, \mathfrak{q}_0^\theta, \mathfrak{q}_1^\theta, \dots, \mathfrak{q}_{h+1}^\theta \in (-\infty, \infty]$ satisfy for all $j \in \{1, 2, \dots, h\}$ that $\mathfrak{w}_0^\theta = -\mathfrak{w}_{h+1}^\theta = -1$, $\mathfrak{q}_0^\theta = 1 - \mathfrak{q}_{h+1}^\theta = 0$, $\mathfrak{w}_j^\theta = \theta_j$, $\mathfrak{b}_j^\theta = \theta_{h+j}$, $\mathfrak{v}_j^\theta = \theta_{2h+j}$, $\mathfrak{c}^\theta = \theta_{\mathfrak{d}_h}$, and*

$$\mathfrak{q}_j^\theta = \begin{cases} -\mathfrak{b}_j^\theta/\mathfrak{w}_j^\theta & : \mathfrak{w}_j^\theta \neq 0 \\ \infty & : \mathfrak{w}_j^\theta = 0, \end{cases} \tag{2.1}$$

for every $h \in \mathbb{N}$, $\theta \in \mathbb{R}^{\mathfrak{d}_h}$ let $M_v^\theta \subseteq \mathbb{N}$, $v \in \{0, 1\}$, satisfy $M_0^\theta = \{k \in \{1, 2, \dots, h\} \colon 0 \le \mathfrak{q}_k^\theta \le 1/2\}$ and $M_1^\theta = \{k \in \{1, 2, \dots, h\} \colon 1/2 \le \mathfrak{q}_k^\theta \le 1\}$, for every $h \in \mathbb{N}$, $\theta \in \mathbb{R}^{\mathfrak{d}_h}$, $v \in \{0, 1\}$ let $m_{v,1}^\theta, m_{v,2}^\theta \in \mathbb{R}$ satisfy

$$(m_{v,1}^\theta, m_{v,2}^\theta) = \begin{cases} \big(\min(M_v^\theta), \max(M_v^\theta)\big) & : M_v^\theta \neq \varnothing \\ (0, h+1) & : M_v^\theta = \varnothing, \end{cases} \tag{2.2}$$

let $\mathbb{A}_r \in C(\mathbb{R}, \mathbb{R})$, $r \in \mathbb{N} \cup \{\infty\}$, satisfy for all $x \in \mathbb{R}$ that $(\bigcup_{r \in \mathbb{N}} \{\mathbb{A}_r\}) \subseteq C^1(\mathbb{R}, \mathbb{R})$, $\mathbb{A}_\infty(x) = \max\{x, 0\}$, $\sup_{r \in \mathbb{N}} \sup_{y \in [-|x|, |x|]} |(\mathbb{A}_r)'(y)| < \infty$, and

$$\limsup_{r \to \infty} \big(|\mathbb{A}_r(x) - \mathbb{A}_\infty(x)| + |(\mathbb{A}_r)'(x) - \mathbb{1}_{(0,\infty)}(x)|\big) = 0, \tag{2.3}$$

for every $h \in \mathbb{N}$, $r \in \mathbb{N} \cup \{\infty\}$, $\theta \in \mathbb{R}^{\mathfrak{d}_h}$ let $\mathscr{N}_r^{h,\theta} \colon \mathbb{R} \to \mathbb{R}$ satisfy for all $x \in \mathbb{R}$ that

$$\mathscr{N}_r^{h,\theta}(x) = \mathfrak{c}^\theta + \textstyle\sum_{i=1}^h \mathfrak{v}_i^\theta \big[\mathbb{A}_r(\mathfrak{w}_i^\theta x + \mathfrak{b}_i^\theta)\big], \tag{2.4}$$

for every $h \in \mathbb{N}$, $r \in \mathbb{N} \cup \{\infty\}$ let $\mathcal{L}_r^h \colon \mathbb{R}^{\mathfrak{d}_h} \to \mathbb{R}$ satisfy for all $\theta \in \mathbb{R}^{\mathfrak{d}_h}$ that

$$\mathcal{L}_r^h(\theta) = \int_0^1 \big(\mathscr{N}_r^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x)\big)^2 \, \mathrm{d}x, \tag{2.5}$$

for every $h \in \mathbb{N}$, $\theta \in \mathbb{R}^{\mathfrak{d}_h}$, $i \in \{1, 2, \dots, h\}$ let $I_i^\theta \subseteq \mathbb{R}$ satisfy $I_i^\theta = \{x \in [0,1] \colon \mathfrak{w}_i^\theta x + \mathfrak{b}_i^\theta > 0\}$, for every $h \in \mathbb{N}$, $\theta \in \mathbb{R}^{\mathfrak{d}_h}$ with $m_{0,1}^\theta \neq 0$ let $\alpha^\theta \in \mathbb{R}$ satisfy for all $x \in \big[0, \mathfrak{q}_{m_{0,1}^\theta}^\theta\big]$ that $\mathscr{N}_\infty^{h,\theta}(x) = \alpha^\theta x + \mathscr{N}_\infty^{h,\theta}(0)$, for every $h \in \mathbb{N}$, $\theta \in \mathbb{R}^{\mathfrak{d}_h}$ with $m_{1,2}^\theta \neq 0$ let $\beta^\theta \in \mathbb{R}$ satisfy for all $x \in \big[\mathfrak{q}_{m_{1,2}^\theta}^\theta, 1\big]$ that $\mathscr{N}_\infty^{h,\theta}(x) = \beta^\theta (x - 1) + \mathscr{N}_\infty^{h,\theta}(1)$, and for every $h \in \mathbb{N}$ let $\mathcal{G}^h = (\mathcal{G}_1^h, \dots, \mathcal{G}_{\mathfrak{d}_h}^h) \colon \mathbb{R}^{\mathfrak{d}_h} \to \mathbb{R}^{\mathfrak{d}_h}$ satisfy for all $\theta \in \{\vartheta \in \mathbb{R}^{\mathfrak{d}_h} \colon ((\nabla \mathcal{L}_r^h)(\vartheta))_{r \in \mathbb{N}}$ is convergent$\}$ that $\mathcal{G}^h(\theta) = \lim_{r \to \infty} (\nabla \mathcal{L}_r^h)(\theta)$.

2.2 Estimates of integrals

Proposition 2.2. *Let $\alpha, \beta \in \mathbb{R}$. Then*

$$\int_0^1 (\alpha x + \beta - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \ge \int_0^1 \left(\frac{3x}{2} - \frac{1}{4} - \mathbb{1}_{(1/2,\infty)}(x) \right)^2 \mathrm{d}x = \frac{1}{16}. \tag{2.6}$$

Proof of Proposition 2.2. Throughout this proof let $\Phi \colon \mathbb{R}^2 \to \mathbb{R}$ satisfy for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ that

$$\Phi(\mathbf{a}, \mathbf{b}) = \int_0^1 (\mathbf{a}x + \mathbf{b} - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x. \tag{2.7}$$

Observe that (2.7) ensures that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ it holds that

$$\begin{aligned}\Phi(\mathbf{a}, \mathbf{b}) &= \int_0^1 (\mathbf{a}x + \mathbf{b} - \mathbb{1}_{(1/2,\infty)}(x))^2 \,\mathrm{d}x \\ &= \int_0^1 [(\mathbf{a}x + \mathbf{b})^2 - 2(\mathbf{a}x + \mathbf{b})\mathbb{1}_{(1/2,\infty)}(x) + \mathbb{1}_{(1/2,\infty)}(x)] \,\mathrm{d}x \\ &= \int_0^1 (\mathbf{a}^2x^2 + 2\mathbf{a}\mathbf{b}x + \mathbf{b}^2) \,\mathrm{d}x - 2\int_{\frac{1}{2}}^1 (\mathbf{a}x + \mathbf{b}) \,\mathrm{d}x + \frac{1}{2} \\ &= \frac{\mathbf{a}^2}{3} + \mathbf{a}\mathbf{b} + \mathbf{b}^2 - \frac{3\mathbf{a}}{4} - \mathbf{b} + \frac{1}{2}.\end{aligned} \tag{2.8}$$

Therefore, we obtain for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ that

$$(\nabla\Phi)(\mathbf{a}, \mathbf{b}) = \left(\frac{2\mathbf{a}}{3} + \mathbf{b} - \frac{3}{4}, \mathbf{a} + 2\mathbf{b} - 1\right). \tag{2.9}$$

This implies that

$$\begin{aligned}&\{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^2 \colon (\nabla\Phi)(\mathbf{a}, \mathbf{b}) = 0\} \\ &= \{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^2 \colon [(\mathbf{a} = 1 - 2\mathbf{b}) \wedge (8\mathbf{a} + 12\mathbf{b} - 9 = 0)]\} \\ &= \{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^2 \colon [(\mathbf{a} = 1 - 2\mathbf{b}) \wedge (8(1 - 2\mathbf{b}) + 12\mathbf{b} - 9 = 0)]\} \\ &= \{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^2 \colon [(\mathbf{a} = 1 - 2\mathbf{b}) \wedge (\mathbf{b} = -1/4)]\} = \{(3/2, -1/4)\}.\end{aligned} \tag{2.10}$$

Combining this with the fact that

$$\Phi(3/2, -1/4) = \frac{3}{4} - \frac{3}{8} + \frac{1}{16} - \frac{9}{8} + \frac{1}{4} + \frac{1}{2} = \frac{1}{16} \tag{2.11}$$

and the fact that

$$\lim\nolimits_{\|(\mathbf{a},\mathbf{b})\|\to\infty} \Phi(\mathbf{a}, \mathbf{b}) = \infty \tag{2.12}$$

establishes that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ it holds that

$$\Phi(\mathbf{a}, \mathbf{b}) \geq \Phi(3/2, -1/4) = \int_0^1 \left(\frac{3x}{2} - \frac{1}{4} - \mathbb{1}_{(1/2,\infty)}(x)\right)^2 \,\mathrm{d}x = \frac{1}{16}. \tag{2.13}$$

The proof of Proposition 2.2 is thus complete. ☐

Proposition 2.3. *Let $\alpha, \beta \in \mathbb{R}$, $\mathscr{a} \in [0, 1/2]$ satisfy $\alpha\mathscr{a} + \beta = 0$. Then*

$$\int_{\mathscr{a}}^{\frac{3}{4}} (\alpha x + \beta - \mathbb{1}_{(1/2,\infty)}(x))^2 \,\mathrm{d}x \geq \int_{\frac{3}{8}}^{\frac{3}{4}} \left(\frac{32x}{9} - \frac{4}{3} - \mathbb{1}_{(1/2,\infty)}(x)\right)^2 \,\mathrm{d}x = \frac{1}{36}. \tag{2.14}$$

Proof of Proposition 2.3. Throughout this proof let $\Phi \colon \mathbb{R}^2 \times [0, 1/2] \to \mathbb{R}$ satisfy for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$, $\mathfrak{a} \in [0, 1/2]$ that

$$\Phi(\mathbf{a}, \mathbf{b}, \mathfrak{a}) = \int_{\mathfrak{a}}^{\frac{3}{4}} (\mathbf{a}x + \mathbf{b} - \mathbb{1}_{(1/2,\infty)}(x))^2 \tag{2.15}$$

and let $L \colon \mathbb{R}^3 \times [0, 1/2]$ satisfy for all $\mathbf{a}, \mathbf{b}, \lambda \in \mathbb{R}$, $\mathfrak{a} \in [0, 1/2]$ that

$$L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{a}) = \Phi(\mathbf{a}, \mathbf{b}, \mathfrak{a}) - \lambda(\mathbf{a}\mathfrak{a} + \mathbf{b}). \tag{2.16}$$

Note that (2.15) ensures that

$$\begin{aligned}\Phi(\mathbf{a},\mathbf{b},\mathfrak{a}) &= \int_{\mathfrak{a}}^{\frac{3}{4}} (\mathbf{a}x+\mathbf{b}-\mathbb{1}_{(1/2,\infty)}(x))^2 \,\mathrm{d}x \\ &= \int_{\mathfrak{a}}^{\frac{3}{4}} [(\mathbf{a}x+\mathbf{b})^2 - 2(\mathbf{a}x+\mathbf{b})\mathbb{1}_{(1/2,\infty)}(x) + \mathbb{1}_{(1/2,\infty)}(x)] \,\mathrm{d}x \\ &= \int_{\mathfrak{a}}^{\frac{3}{4}} (\mathbf{a}^2x^2+2\mathbf{a}\mathbf{b}x+\mathbf{b}^2)\,\mathrm{d}x - 2\int_{\frac{1}{2}}^{\frac{3}{4}}(\mathbf{a}x+\mathbf{b})\,\mathrm{d}x + \frac{1}{4} \\ &= \frac{\mathbf{a}^2}{3}\Big(\frac{27}{64}-\mathfrak{a}^3\Big) + \mathbf{a}\mathbf{b}\Big(\frac{9}{16}-\mathfrak{a}^2\Big) + \mathbf{b}^2\Big(\frac{3}{4}-\mathfrak{a}\Big) - \frac{5\mathbf{a}}{16} - \frac{\mathbf{b}}{2} + \frac{1}{4}.\end{aligned} \tag{2.17}$$

This implies for all $\mathbf{a},\mathbf{b},\lambda\in\mathbb{R}$, $\mathfrak{a}\in[0,1/2]$ that

$$\begin{aligned}&\frac{\partial}{\partial\mathbf{a}}L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{a}) = \frac{\partial}{\partial\mathbf{a}}\Phi(\mathbf{a},\mathbf{b},\mathfrak{a}) - \lambda\mathfrak{a} = \frac{2\mathbf{a}}{3}\Big(\frac{27}{64}-\mathfrak{a}^3\Big) + \mathbf{b}\Big(\frac{9}{16}-\mathfrak{a}^2\Big) - \frac{5}{16} - \lambda\mathfrak{a}, \\ &\frac{\partial}{\partial\mathbf{b}}L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{a}) = \frac{\partial}{\partial\mathbf{b}}\Phi(\mathbf{a},\mathbf{b},\mathfrak{a}) - \lambda = \mathbf{a}\Big(\frac{9}{16}-\mathfrak{a}^2\Big) + 2\mathbf{b}\Big(\frac{3}{4}-\mathfrak{a}\Big) - \frac{1}{2} - \lambda, \\ &\frac{\partial}{\partial\lambda}L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{a}) = -\mathbf{a}\mathfrak{a} - \mathbf{b}, \qquad \text{and} \\ &\frac{\partial}{\partial\mathfrak{a}}L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{a}) = \frac{\partial}{\partial\mathfrak{a}}\Phi(\mathbf{a},\mathbf{b},\mathfrak{a}) - \lambda\mathbf{a} = -\mathbf{a}^2\mathfrak{a}^2 - 2\mathbf{a}\mathbf{b}\mathfrak{a} - \mathbf{b}^2 - \lambda\mathbf{a}.\end{aligned} \tag{2.18}$$

Hence, we obtain that

$$\begin{aligned}&\begin{cases}\frac{\partial}{\partial\mathbf{a}}L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{a})=0\\ \frac{\partial}{\partial\mathbf{b}}L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{a})=0\\ \frac{\partial}{\partial\lambda}L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{a})=0\\ \frac{\partial}{\partial\mathfrak{a}}L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{a})=0\end{cases} = \begin{cases}\frac{2\mathbf{a}}{3}\left(\frac{27}{64}-\mathfrak{a}^3\right)+\mathbf{b}\left(\frac{9}{16}-\mathfrak{a}^2\right)-\frac{5}{16}-\lambda\mathfrak{a}=0\\ \mathbf{a}\left(\frac{9}{16}-\mathfrak{a}^2\right)+2\mathbf{b}\left(\frac{3}{4}-\mathfrak{a}\right)-\frac{1}{2}-\lambda=0\\ -\mathbf{a}\mathfrak{a}-\mathbf{b}=0\\ -\mathbf{a}^2\mathfrak{a}^2-2\mathbf{a}\mathbf{b}\mathfrak{a}-\mathbf{b}^2-\lambda\mathbf{a}=0\end{cases}\\ &= \begin{cases}\frac{2\mathbf{a}}{3}\left(\frac{27}{64}-\mathfrak{a}^3\right)-\mathbf{a}\mathfrak{a}\left(\frac{9}{16}-\mathfrak{a}^2\right)-\frac{5}{16}-\lambda\mathfrak{a}=0\\ \mathbf{a}\left(\frac{9}{16}-\mathfrak{a}^2\right)-2\mathbf{a}\mathfrak{a}\left(\frac{3}{4}-\mathfrak{a}\right)-\frac{1}{2}-\lambda=0\\ \mathbf{b}=-\mathbf{a}\mathfrak{a}\\ \lambda\mathbf{a}=0\end{cases}\\ &= \begin{cases}\frac{9\mathbf{a}}{32}+\frac{\mathbf{a}\mathfrak{a}^3}{3}-\frac{9\mathbf{a}\mathfrak{a}}{16}-\frac{5}{16}-\lambda\mathfrak{a}=0\\ \frac{9\mathbf{a}}{16}+\mathbf{a}\mathfrak{a}^2-\frac{3\mathbf{a}\mathfrak{a}}{2}-\frac{1}{2}-\lambda=0\\ \mathbf{b}=-\mathbf{a}\mathfrak{a}\\ \lambda\mathbf{a}=0\end{cases}\\ &= \begin{cases}\mathbf{a}=\left(\frac{5}{16}+\lambda\mathfrak{a}\right)\left(\frac{9}{32}-\frac{9\mathfrak{a}}{16}+\frac{\mathfrak{a}^3}{3}\right)^{-1}\\ \left(\frac{5}{16}+\lambda\mathfrak{a}\right)\left(\frac{9}{32}-\frac{9\mathfrak{a}}{16}+\frac{\mathfrak{a}^3}{3}\right)^{-1}\left(\frac{9}{16}+\mathfrak{a}^2-\frac{3\mathfrak{a}}{2}\right)-\frac{1}{2}=\lambda\\ \mathbf{b}=-\mathbf{a}\mathfrak{a}\\ \lambda\mathbf{a}=0.\end{cases}\end{aligned} \tag{2.19}$$

This implies in the case $\nabla = L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{a}) = \mathbf{a} = 0$ that it holds that $\mathbf{a}=\mathbf{b}=0$, $\lambda=-1/2$, and $\mathfrak{a}=5/8$ which is not in the domain of L. In the following we distinguish between the case $\lambda=0$, the case $\mathfrak{a}=0$, and the case $\mathfrak{a}=1/2$. We first show (2.14) in the case

$$\lambda = 0. \tag{2.20}$$

Observe that (2.19) and (2.20) ensure that

$$\begin{aligned}0 &= \left(\frac{5}{16}\right)\left(\frac{9}{32}-\frac{9\mathfrak{a}}{16}+\frac{\mathfrak{a}^3}{3}\right)^{-1}\left(\frac{9}{16}+\mathfrak{a}^2-\frac{3\mathfrak{a}}{2}\right)-\frac{1}{2}=5\left(\frac{9}{16}+\mathfrak{a}^2-\frac{3\mathfrak{a}}{2}\right)\\ &\quad-8\left(\frac{9}{32}-\frac{9\mathfrak{a}}{16}+\frac{\mathfrak{a}^3}{3}\right)=\frac{9}{16}+5\mathfrak{a}^2-3\mathfrak{a}-\frac{8\mathfrak{a}^3}{3}=-\frac{1}{48}(3-4\mathfrak{a})^2(8\mathfrak{a}-3).\end{aligned} \tag{2.21}$$

Therefore, we obtain that $\mathfrak{a}={}^3/_8$. Combining this with (2.19) shows that

$$\nabla L({}^{32}/_9,-{}^4/_3,0,{}^3/_8)=0. \tag{2.22}$$

This, Lagrange multiplier theorem, and the fact that for all $\mathfrak{a}\in(0,{}^1/_2)$ it holds that

$$\lim_{\|(\mathbf{a},\mathbf{b})\|\to\infty}\Phi(\mathbf{a},\mathbf{b},\mathfrak{a})=\infty \tag{2.23}$$

imply for all $\mathbf{a},\mathbf{b}\in\mathbb{R}$, $\mathfrak{a}\in(0,{}^1/_2)$ with $\mathbf{a}\mathfrak{a}+\mathbf{b}=0$ that

$$\Phi(\mathbf{a},\mathbf{b},\mathfrak{a})\geq\Phi({}^{32}/_9,-{}^4/_3,{}^3/_8)=\frac{1}{36}. \tag{2.24}$$

This establishes (2.14) in the case $\lambda=0$. In the next step we prove (2.14) in the case

$$\mathfrak{a}=0. \tag{2.25}$$

Note that (2.17) and (2.25) assure for all $\mathbf{a},\mathbf{b}\in\mathbb{R}$ with $\mathbf{a}\mathfrak{a}+\mathbf{b}=0$ that

$$\Phi(\mathbf{a},\mathbf{b},0)=\Phi(\mathbf{a},0,0)=\frac{9\mathbf{a}^2}{64}-\frac{5\mathbf{a}}{16}+\frac{1}{4}\geq\frac{9}{64}\left(\frac{10}{9}\right)^2-\frac{5}{16}\left(\frac{10}{9}\right)+\frac{1}{4}=\frac{11}{144}. \tag{2.26}$$

This establishes (2.14) in the case $\mathfrak{a}=0$. Finally we demonstrate (2.14) in the case

$$\mathfrak{a}=\frac{1}{2}. \tag{2.27}$$

Observe that (2.17) and (2.27) assure for all $\mathbf{a},\mathbf{b}\in\mathbb{R}$ with $\mathbf{a}\mathfrak{a}+\mathbf{b}=0$ that

$$\begin{aligned}\Phi(\mathbf{a},\mathbf{b},{}^1/_2)=\Phi(\mathbf{a},-{}^{\mathbf{a}}/_2,{}^1/_2)&=\frac{19\mathbf{a}^2}{192}-\frac{5\mathbf{a}^2}{32}+\frac{\mathbf{a}^2}{16}-\frac{5\mathbf{a}}{16}+\frac{\mathbf{a}}{4}+\frac{1}{4}\\ &=\frac{\mathbf{a}^2}{192}-\frac{\mathbf{a}}{16}+\frac{1}{4}\geq\frac{6^2}{192}-\frac{6}{16}+\frac{1}{4}=\frac{1}{16}.\end{aligned} \tag{2.28}$$

This establishes (2.14) in the case $\mathfrak{a}={}^1/_2$. The proof of Proposition 2.3 is thus complete. ☐

Proposition 2.4. *Let $\alpha,\beta\in\mathbb{R}$, $\mathscr{b}\in[{}^1/_2,1]$ satisfy $\alpha\mathscr{b}+\beta=1$. Then*

$$\int_{\frac{1}{4}}^{\mathscr{b}}(\alpha x+\beta-\mathbb{1}_{({}^1/_2,\infty)}(x))^2\,\mathrm{d}x\geq\int_{\frac{1}{4}}^{\frac{5}{8}}\left(\frac{32x}{9}-\frac{11}{9}-\mathbb{1}_{({}^1/_2,\infty)}(x)\right)^2\mathrm{d}x=\frac{1}{36}. \tag{2.29}$$

Proof of Proposition 2.4. Throughout this proof let $\Phi\colon\mathbb{R}^2\times[{}^1/_2,1]\to\mathbb{R}$ satisfy for all $\mathbf{a},\mathbf{b}\in\mathbb{R}$, $\mathfrak{b}\in[{}^1/_2,1]$ that

$$\Phi(\mathbf{a},\mathbf{b},\mathfrak{b})=\int_{\frac{1}{4}}^{\mathfrak{b}}(\mathbf{a}x+\mathbf{b}-\mathbb{1}_{({}^1/_2,\infty)}(x))^2 \tag{2.30}$$

and let $L\colon\mathbb{R}^3\times[{}^1/_2,1]$ satisfy for all $\mathbf{a},\mathbf{b},\lambda\in\mathbb{R}$, $\mathfrak{b}\in[{}^1/_2,1]$ that

$$L(\mathbf{a},\mathbf{b},\lambda,\mathfrak{b})=\Phi(\mathbf{a},\mathbf{b},\mathfrak{b})-\lambda(\mathbf{a}\mathfrak{b}+\mathbf{b}-1). \tag{2.31}$$

Note that (2.30) ensures for all $\mathbf{a}, \mathbf{b}, \lambda \in \mathbb{R}$, $\mathfrak{b} \in [1/2, 1]$ that

$$\begin{aligned} \Phi(\mathbf{a}, \mathbf{b}, \mathfrak{b}) &= \int_{\frac{1}{4}}^{\mathfrak{b}} (\mathbf{a}x + \mathbf{b} - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \\ &= \int_{\frac{1}{4}}^{\mathfrak{b}} [(\mathbf{a}x + \mathbf{b})^2 - 2(\mathbf{a}x + \mathbf{b})\mathbb{1}_{(1/2,\infty)}(x) + \mathbb{1}_{(1/2,\infty)}(x)] \, \mathrm{d}x \\ &= \int_{\frac{1}{4}}^{\mathfrak{b}} (\mathbf{a}^2 x^2 + 2\mathbf{a}\mathbf{b}x + \mathbf{b}^2) \, \mathrm{d}x - 2\int_{\frac{1}{2}}^{\mathfrak{b}} (\mathbf{a}x + \mathbf{b}) \, \mathrm{d}x + \mathfrak{b} - \frac{1}{2} \\ &= \frac{\mathbf{a}^2}{3}\Big(\mathfrak{b}^3 - \frac{1}{64}\Big) + \mathbf{a}\mathbf{b}\Big(\mathfrak{b}^2 - \frac{1}{16}\Big) + \mathbf{b}^2\Big(\mathfrak{b} - \frac{1}{4}\Big) - \mathbf{a}\left(\mathfrak{b}^2 - \frac{1}{4}\right) \\ &\quad - 2\mathbf{b}\left(\mathfrak{b} - \frac{1}{2}\right) + \mathfrak{b} - \frac{1}{2}. \end{aligned} \tag{2.32}$$

This implies for all $\mathbf{a}, \mathbf{b}, \lambda \in \mathbb{R}$, $\mathfrak{b} \in [1/2, 1]$ that

$$\begin{aligned} \frac{\partial}{\partial \mathbf{a}} L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) &= \frac{\partial}{\partial \mathbf{a}} \Phi(\mathbf{a}, \mathbf{b}, \mathfrak{b}) - \lambda \mathfrak{b} = \frac{2\mathbf{a}}{3}\Big(\mathfrak{b}^3 - \frac{1}{64}\Big) + \mathbf{b}\Big(\mathfrak{b}^2 - \frac{1}{16}\Big) - \mathfrak{b}^2 + \frac{1}{4} - \lambda \mathfrak{b}, \\ \frac{\partial}{\partial \mathbf{b}} L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) &= \frac{\partial}{\partial \mathbf{b}} \Phi(\mathbf{a}, \mathbf{b}, \mathfrak{b}) - \lambda = \mathbf{a}\Big(\mathfrak{b}^2 - \frac{1}{16}\Big) + 2\mathbf{b}\Big(\mathfrak{b} - \frac{1}{4}\Big) - 2\mathfrak{b} + 1 - \lambda, \\ \frac{\partial}{\partial \lambda} L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) &= -\mathbf{a}\mathfrak{b} - \mathbf{b} + 1, \qquad \text{and} \\ \frac{\partial}{\partial \mathfrak{b}} L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) &= \frac{\partial}{\partial \mathfrak{b}} \Phi(\mathbf{a}, \mathbf{b}, \mathfrak{b}) - \lambda \mathbf{a} = \mathbf{a}^2\mathfrak{b}^2 + 2\mathbf{a}\mathbf{b}\mathfrak{b} + \mathbf{b}^2 - 2\mathbf{a}\mathfrak{b} - 2\mathbf{b} + 1 - \lambda \mathbf{a}. \end{aligned} \tag{2.33}$$

Hence, we obtain that

$$\begin{aligned} \begin{cases} \frac{\partial}{\partial \mathbf{a}} L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) = 0 \\ \frac{\partial}{\partial \mathbf{b}} L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) = 0 \\ \frac{\partial}{\partial \lambda} L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) = 0 \\ \frac{\partial}{\partial \mathfrak{b}} L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) = 0 \end{cases} &= \begin{cases} \frac{2\mathbf{a}}{3}\left(\mathfrak{b}^3 - \frac{1}{64}\right) + \mathbf{b}\left(\mathfrak{b}^2 - \frac{1}{16}\right) - \mathfrak{b}^2 + \frac{1}{4} - \lambda\mathfrak{b} = 0 \\ \mathbf{a}\left(\mathfrak{b}^2 - \frac{1}{16}\right) + 2\mathbf{b}\left(\mathfrak{b} - \frac{1}{4}\right) - 2\mathfrak{b} + 1 - \lambda = 0 \\ -\mathbf{a}\mathfrak{b} - \mathbf{b} + 1 = 0 \\ \mathbf{a}^2\mathfrak{b}^2 + 2\mathbf{a}\mathbf{b}\mathfrak{b} + \mathbf{b}^2 - 2\mathbf{a}\mathfrak{b} - 2\mathbf{b} + 1 - \lambda\mathbf{a} = 0 \end{cases} \\ &= \begin{cases} \frac{2\mathbf{a}}{3}\left(\mathfrak{b}^3 - \frac{1}{64}\right) + \mathbf{b}\left(\mathfrak{b}^2 - \frac{1}{16}\right) - \mathfrak{b}^2 + \frac{1}{4} - \lambda\mathfrak{b} = 0 \\ \mathbf{a}\left(\mathfrak{b}^2 - \frac{1}{16}\right) + 2\mathbf{b}\left(\mathfrak{b} - \frac{1}{4}\right) - 2\mathfrak{b} + 1 - \lambda = 0 \\ \mathbf{b} = -\mathbf{a}\mathfrak{b} + 1 \\ \lambda\mathbf{a} = 0 \end{cases} \\ &= \begin{cases} \frac{2\mathbf{a}\mathfrak{b}^3}{3} - \frac{\mathbf{a}}{96} - \mathbf{a}\mathfrak{b}^3 + \mathfrak{b}^2 + \frac{\mathbf{a}\mathfrak{b}}{16} - \frac{1}{16} - \mathfrak{b}^2 + \frac{1}{4} - \lambda\mathfrak{b} = 0 \\ \mathbf{a}\mathfrak{b}^2 - \frac{\mathbf{a}}{16} - 2\mathbf{a}\mathfrak{b}^2 + 2\mathfrak{b} + \frac{\mathbf{a}\mathfrak{b}}{2} - \frac{1}{2} - 2\mathfrak{b} + 1 - \lambda = 0 \\ \mathbf{b} = -\mathbf{a}\mathfrak{b} + 1 \\ \lambda\mathbf{a} = 0 \end{cases} \\ &= \begin{cases} -\frac{\mathbf{a}\mathfrak{b}^3}{3} - \frac{\mathbf{a}}{96} + \frac{\mathbf{a}\mathfrak{b}}{16} + \frac{3}{16} - \lambda\mathfrak{b} = 0 \\ -\mathbf{a}\mathfrak{b}^2 - \frac{\mathbf{a}}{16} + \frac{\mathbf{a}\mathfrak{b}}{2} + \frac{1}{2} - \lambda = 0 \\ \mathbf{b} = -\mathbf{a}\mathfrak{b} + 1 \\ \lambda\mathbf{a} = 0 \end{cases} \\ &= \begin{cases} \mathbf{a} = \left(-\frac{3}{16} + \lambda\mathfrak{b}\right)\left(-\frac{1}{96} + \frac{\mathfrak{b}}{16} - \frac{\mathfrak{b}^3}{3}\right)^{-1} \\ \left(\frac{\mathfrak{b}}{2} - \mathfrak{b}^2 - \frac{1}{16}\right)\left(-\frac{3}{16} + \lambda\mathfrak{b}\right)\left(-\frac{1}{96} + \frac{\mathfrak{b}}{16} - \frac{\mathfrak{b}^3}{3}\right)^{-1} + \frac{1}{2} = \lambda \\ \mathbf{b} = -\mathbf{a}\mathfrak{b} + 1 \\ \lambda\mathbf{a} = 0. \end{cases} \end{aligned} \tag{2.34}$$

This implies that in the case $\nabla L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) = \lambda - {}^1\!/_2 = 0$ it holds that $\mathbf{a} = 0$, $\mathbf{b} = 1$, and $\mathfrak{b} = {}^3\!/_8$ which is not in the domain of L. In the following we distinguish between the case $\lambda = 0$, the case $\mathfrak{b} = {}^1\!/_2$, and the case $\mathfrak{b} = 1$. We first demonstrate (2.29) in the case

$$\lambda = 0. \tag{2.35}$$

Observe that (2.34) and (2.35) ensure that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$, $\mathfrak{b} \in [{}^1\!/_2, 1]$ such that $\nabla L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) = 0$ it holds that

$$0 = \left(\frac{\mathfrak{b}}{2} - \mathfrak{b}^2 - \frac{1}{16}\right)\left(-\frac{3}{16}\right)\left(-\frac{1}{96} + \frac{\mathfrak{b}}{16} - \frac{\mathfrak{b}^3}{3}\right)^{-1} + \frac{1}{2}. \tag{2.36}$$

Therefore, we obtain that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$, $\mathfrak{b} \in [{}^1\!/_2, 1]$ such that $\nabla L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) = 0$ it holds that

$$0 = \left(48\mathfrak{b} - 96\mathfrak{b}^2 - 6\right) - \frac{8}{3}\left(-1 + 6\mathfrak{b} - 32\mathfrak{b}^3\right) = \frac{2}{3}(4\mathfrak{b} - 1)^2(8\mathfrak{b} - 5). \tag{2.37}$$

Combining this with (2.34) shows that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$, $\mathfrak{b} \in [{}^1\!/_2, 1]$ such that $\nabla L(\mathbf{a}, \mathbf{b}, \lambda, \mathfrak{b}) = 0$ it holds that $\mathfrak{b} = {}^5\!/_8$, $\mathbf{a} = {}^{32}\!/_9$, and $\mathbf{b} = -{}^{11}\!/_9$. This, Lagrange multiplier theorem, and the fact that for all $\mathfrak{b} \in ({}^1\!/_2, 1)$ it holds that

$$\lim\nolimits_{\|(\mathbf{a},\mathbf{b})\| \to \infty} \Phi(\mathbf{a}, \mathbf{b}, \mathfrak{b}) = \infty \tag{2.38}$$

imply for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$, $\mathfrak{b} \in ({}^1\!/_2, 1)$ with $\mathbf{a}\mathfrak{b} + \mathbf{b} = 1$ that

$$\Phi(\mathbf{a}, \mathbf{b}, \mathfrak{b}) \geq \Phi({}^{32}\!/_9, -{}^{11}\!/_9, {}^5\!/_8) = \frac{1}{36}. \tag{2.39}$$

This establishes (2.29) in the case $\lambda = 0$. In the next step we prove (2.29) in the case

$$\mathfrak{b} = 1. \tag{2.40}$$

Note that (2.32) and (2.40) assure for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ with $\mathbf{a}\mathfrak{b} + \mathbf{b} = 1$ that

$$\begin{aligned}\Phi(\mathbf{a}, \mathbf{b}, 1) = \Phi(\mathbf{a}, 1 - \mathbf{a}, 1) &= \frac{21\mathbf{a}^2}{64} + \frac{15\mathbf{a}(1 - \mathbf{a})}{16} + \frac{3(1 - \mathbf{a})^2}{4} - \frac{3\mathbf{a}}{4} - (1 - \mathbf{a}) + 1 - \frac{1}{2} \\ &= \frac{1}{64}(9\mathbf{a}^2 - 20\mathbf{a} + 16) \geq \frac{1}{64}\left(9\left(\frac{10}{9}\right)^2 - 20\left(\frac{10}{9}\right) + 16\right) = \frac{11}{144}.\end{aligned} \tag{2.41}$$

This establishes (2.29) in the case $\mathfrak{b} = 1$. Finally we show (2.29) in the case

$$\mathfrak{b} = \frac{1}{2}. \tag{2.42}$$

Observe that (2.32) and (2.42) assure for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ with $\mathbf{a}\mathfrak{b} + \mathbf{b} = 1$ that

$$\begin{aligned}\Phi(\mathbf{a}, \mathbf{b}, {}^1\!/_2) = \Phi(\mathbf{a}, 1 - {}^{\mathbf{a}}\!/_2, {}^1\!/_2) &= \frac{\mathbf{a}^2}{64} + \frac{3\mathbf{a}(1 - \frac{\mathbf{a}}{2})}{16} + \frac{(1 - \frac{\mathbf{a}}{2})^2}{4} \\ &= \frac{\mathbf{a}^2}{192} - \frac{\mathbf{a}}{16} + \frac{1}{4} \geq \frac{6^2}{192} - \frac{6}{16} + \frac{1}{4} = \frac{1}{16}.\end{aligned} \tag{2.43}$$

This establishes (2.29) in the case $\mathfrak{b} = {}^1\!/_2$. The proof of Proposition 2.4 is thus complete. ◻

Proposition 2.5. *Let $\alpha, \beta \in \mathbb{R}$. Then*

$$\int_{\frac{1}{4}}^{\frac{3}{4}} (\alpha x + \beta - \mathbb{1}_{({}^1\!/_2, \infty)}(x))^2 \, \mathrm{d}x \geq \int_{\frac{1}{4}}^{\frac{3}{4}} \left(3x - 1 - \mathbb{1}_{({}^1\!/_2, \infty)}(x)\right)^2 \, \mathrm{d}x = \frac{1}{32}. \tag{2.44}$$

Proof of Proposition 2.5. Throughout this proof let $\Phi\colon \mathbb{R}^2 \to \mathbb{R}$ satisfy for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ that

$$\Phi(\mathbf{a}, \mathbf{b}) = \int_{\frac{1}{4}}^{\frac{3}{4}} (\mathbf{a}x + \mathbf{b} - \mathbb{1}_{(1/2,\infty)}(x))^2 \, dx. \tag{2.45}$$

Note that (2.45) ensures that

$$\begin{aligned}
\Phi(\mathbf{a}, \mathbf{b}) &= \int_{\frac{1}{4}}^{\frac{3}{4}} (\mathbf{a}x + \mathbf{b} - \mathbb{1}_{(1/2,\infty)}(x))^2 \, dx \\
&= \int_{\frac{1}{4}}^{\frac{3}{4}} [(\mathbf{a}x + \mathbf{b})^2 - 2(\mathbf{a}x + \mathbf{b})\mathbb{1}_{(1/2,\infty)}(x) + \mathbb{1}_{(1/2,\infty)}(x)] \, dx \\
&= \int_{\frac{1}{4}}^{\frac{3}{4}} (\mathbf{a}^2 x^2 + 2\mathbf{a}\mathbf{b}x + \mathbf{b}^2) \, dx - 2 \int_{\frac{1}{2}}^{\frac{3}{4}} (\mathbf{a}x + \mathbf{b}) \, dx + \frac{1}{4} \\
&= \frac{26\mathbf{a}^2}{192} + \frac{\mathbf{a}\mathbf{b}}{2} + \frac{\mathbf{b}^2}{2} - \frac{5\mathbf{a}}{16} - \frac{\mathbf{b}}{2} + \frac{1}{4}.
\end{aligned} \tag{2.46}$$

Hence, we obtain for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ that

$$(\nabla\Phi)(\mathbf{a}, \mathbf{b}) = \left(\frac{13\mathbf{a}}{48} + \frac{\mathbf{b}}{2} - \frac{5}{16}, \frac{\mathbf{a}}{2} + \mathbf{b} - \frac{1}{2} \right). \tag{2.47}$$

This implies that

$$\begin{aligned}
&\{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^2 \colon (\nabla\Phi)(\mathbf{a}, \mathbf{b}) = 0\} \\
&= \{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^2 \colon [(\mathbf{a} = 1 - 2\mathbf{b}) \wedge (13\mathbf{a} + 24\mathbf{b} - 15 = 0)]\} \\
&= \{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^2 \colon [(\mathbf{a} = 1 - 2\mathbf{b}) \wedge (13(1 - 2\mathbf{b}) + 24\mathbf{b} - 15 = 0)]\} \\
&= \{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^2 \colon [(\mathbf{a} = 1 - 2\mathbf{b}) \wedge (\mathbf{b} = -1)]\} = \{(3, -1)\}.
\end{aligned} \tag{2.48}$$

Combining this with the fact that

$$\Phi(3, -1) = \frac{39}{32} - \frac{3}{2} + \frac{1}{2} - \frac{15}{16} + \frac{1}{2} + \frac{1}{4} = \frac{39}{32} - \frac{30}{32} - \frac{1}{4} = \frac{1}{32} \tag{2.49}$$

and the fact that

$$\lim\nolimits_{\|(\mathbf{a},\mathbf{b})\| \to \infty} \Phi(\mathbf{a}, \mathbf{b}) = \infty \tag{2.50}$$

establishes that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ it holds that

$$\Phi(\mathbf{a}, \mathbf{b}) \geq \Phi(3, -1) = \int_{\frac{1}{4}}^{\frac{3}{4}} \left(3x - 1 - \mathbb{1}_{(1/2,\infty)}(x)\right)^2 \, dx = \frac{1}{32}. \tag{2.51}$$

The proof of Proposition 2.5 is thus complete. □

Lemma 2.6. *Let $\alpha, \beta, \mathscr{a} \in \mathbb{R}$, $\mathscr{b} \in (\mathscr{a}, \infty)$. Then*

$$\int_{\mathscr{a}}^{\mathscr{b}} (\alpha x + \beta)^2 \, dx \geq \int_{\mathscr{a}}^{\mathscr{b}} \left[\alpha x - \frac{\alpha(\mathscr{b} + \mathscr{a})}{2} \right]^2 dx = \frac{\alpha^2 (\mathscr{b} - \mathscr{a})^3}{12}. \tag{2.52}$$

Proof of Lemma 2.6. Observe that, e.g., [23, Lemma 5.1] establishes (2.52). The proof of Lemma 2.6 is thus complete. □

2.3 Properties of integrands

Lemma 2.7. *Let* $\mathscr{a} \in \mathbb{R}$*,* $\mathscr{b} \in (\mathscr{a}, \infty)$*,* $\alpha, \beta \in \mathbb{R}$ *satisfy* $1/2 \notin (\mathscr{a}, \mathscr{b})$ *and*

$$\int_{\mathscr{a}}^{\mathscr{b}} x(\alpha x + \beta - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathscr{a}}^{\mathscr{b}} (\alpha x + \beta - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.53}$$

Then $\alpha = 0$ *and* $\beta = \mathbb{1}_{(-\infty,\mathscr{a}]}(1/2)$*.*

Proof of Lemma 2.7. Note that the assumption that $1/2 \notin (\mathscr{a}, \mathscr{b})$ and, e.g., [23, Lemma 6.6] demonstrate that $\alpha = 0$ and $\beta = \mathbb{1}_{(-\infty,\mathscr{a}]}(1/2)$. The proof of Lemma 2.7 is thus complete. ☐

Lemma 2.8. *Let* $\mathscr{a} \in [0, 1/2)$*,* $\mathscr{b} \in (1/2, 1]$*,* $\alpha, \beta \in \mathbb{R}$ *satisfy*

$$\int_{\mathscr{a}}^{\mathscr{b}} x(\alpha x + \beta - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathscr{a}}^{\mathscr{b}} (\alpha x + \beta - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.54}$$

Then

$$\alpha = \frac{3(2\mathscr{a} - 1)(2\mathscr{b} - 1)}{2(\mathscr{a} - \mathscr{b})^3} \qquad \text{and} \qquad \beta = -\frac{(2\mathscr{b} - 1)(8\mathscr{a}^2 + \mathscr{a}(2\mathscr{b} - 3) + \mathscr{b}(2\mathscr{b} - 3))}{4(\mathscr{a} - \mathscr{b})^3}. \tag{2.55}$$

Proof of Lemma 2.8. Observe that (2.54) implies that

$$0 = \int_{\mathscr{a}}^{\mathscr{b}} (\alpha x + \beta - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \frac{\alpha}{2}(\mathscr{b}^2 - \mathscr{a}^2) + \beta(\mathscr{b} - \mathscr{a}) - (\mathscr{b} - 1/2). \tag{2.56}$$

Therefore, we obtain that

$$\beta = (\mathscr{b} - 1/2)(\mathscr{b} - \mathscr{a})^{-1} - \frac{\alpha}{2}(\mathscr{b} + \mathscr{a}). \tag{2.57}$$

Furthermore, note that (2.54) assures that

$$0 = \int_{\mathscr{a}}^{\mathscr{b}} [\alpha x^2 + \beta x - x\mathbb{1}_{(1/2,\infty)}(x)]\,\mathrm{d}x = \frac{\alpha}{3}(\mathscr{b}^3 - \mathscr{a}^3) + \frac{\beta}{2}(\mathscr{b}^2 - \mathscr{a}^2) - \frac{1}{2}(\mathscr{b}^2 - 1/4). \tag{2.58}$$

This and (2.57) demonstrate that

$$\begin{aligned}
0 &= \frac{\alpha}{3}(\mathscr{b}^3 - \mathscr{a}^3) + \frac{\beta}{2}(\mathscr{b}^2 - \mathscr{a}^2) - \frac{1}{2}(\mathscr{b}^2 - 1/4) \\
&= \frac{\alpha}{3}(\mathscr{b}^3 - \mathscr{a}^3) + \frac{1}{2}(\mathscr{b} - 1/2)(\mathscr{b} + \mathscr{a}) - \frac{\alpha}{4}(\mathscr{b}^2 - \mathscr{a}^2)(\mathscr{b} + \mathscr{a}) - \frac{1}{2}(\mathscr{b}^2 - 1/4) \\
&= \frac{\alpha}{12}(\mathscr{b} - \mathscr{a})[4(\mathscr{b}^2 + \mathscr{a}\mathscr{b} + \mathscr{a}^2) - 3(\mathscr{b} + \mathscr{a})^2] + \frac{1}{2}(\mathscr{b} - 1/2)[(\mathscr{b} + \mathscr{a}) - (\mathscr{b} + 1/2)] \\
&= \frac{\alpha}{12}(\mathscr{b} - \mathscr{a})^3 + \frac{1}{2}(\mathscr{b} - 1/2)(\mathscr{a} - 1/2).
\end{aligned} \tag{2.59}$$

Hence, we obtain that

$$\alpha = \frac{3(2\mathscr{a} - 1)(2\mathscr{b} - 1)}{2(\mathscr{a} - \mathscr{b})^3}. \tag{2.60}$$

This and (2.57) establish that

$$\begin{aligned}
\beta &= (\mathscr{b} - \frac{1}{2})(\mathscr{b} - \mathscr{a})^{-1} - \frac{\alpha}{2}(\mathscr{b} + \mathscr{a}) = (\mathscr{b} - \frac{1}{2})(\mathscr{b} - \mathscr{a})^{-1} - \frac{3(2\mathscr{b} - 1)(2\mathscr{a} - 1)(\mathscr{b} + \mathscr{a})}{4(\mathscr{a} - \mathscr{b})^3} \\
&= -\frac{(2\mathscr{b} - 1)[2(\mathscr{a} - \mathscr{b})^2 + 3(2\mathscr{a} - 1)(\mathscr{b} + \mathscr{a})]}{4(\mathscr{a} - \mathscr{b})^3} \\
&= -\frac{(2\mathscr{b} - 1)(8\mathscr{a}^2 + \mathscr{a}(2\mathscr{b} - 3) + \mathscr{b}(2\mathscr{b} - 3))}{4(\mathscr{a} - \mathscr{b})^3}.
\end{aligned} \tag{2.61}$$

The proof of Lemma 2.8 is thus complete. ☐

Corollary 2.9. *Let $a \in [0, 1/2)$, $b \in (1/2, 1]$, $\alpha, \beta \in \mathbb{R}$ satisfy $\alpha a + \beta = 0$ and*

$$\int_a^b x((\alpha x + \beta) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_a^b ((\alpha x + \beta) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.62}$$

Then

$$\alpha = \frac{16}{9(2b-1)} \qquad \textit{and} \qquad \beta = \frac{4(2b-3)}{9(2b-1)}. \tag{2.63}$$

Proof of Corollary 2.9. Observe that Lemma 2.8 and the assumption that $\alpha a + \beta = 0$ prove that

$$0 = \frac{3a(2a-1)(2b-1)}{2(a-b)^3} - \frac{(2b-1)(8a^2 + a(2b-3) + b(2b-3))}{4(a-b)^3}. \tag{2.64}$$

Therefore, we obtain that

$$\begin{aligned} 0 &= 6a(2a-1) - (8a^2 + a(2b-3) + b(2b-3)) \\ &= 4a^2 - a(2b+3) - b(2b-3) = 4a^2 - 2ab - 3a - 2b^2 + 3b \\ &= 4a(a-b) + 2ab - 2b^2 - 3(a-b) = (a-b)(4a + 2b - 3). \end{aligned} \tag{2.65}$$

This, the assumption that $a \in [0, 1/2)$, and the assumption that $b \in (1/2, 1]$ imply that $a = -b/2 + 3/4$. Combining this with Lemma 2.8 demonstrates that

$$\begin{aligned} \alpha &= \frac{3(-b+\frac{1}{2})(2b-1)}{2(-\frac{3b}{2}+\frac{3}{4})^3} = \frac{-\frac{3}{2}(2b-1)^2}{\frac{27}{32}(-2b+1)^3} = \frac{-\frac{3}{2}}{-\frac{27}{32}(2b-1)} = \frac{16}{9(2b-1)} \qquad \text{and} \\ \beta &= -\frac{(2b-1)[8(-\frac{b}{2}+\frac{3}{4})^2 + (-\frac{b}{2}+\frac{3}{4})(2b-3) + b(2b-3)]}{4(-\frac{3b}{2}+\frac{3}{4})^3} \\ &= \frac{\frac{1}{2}(-2b+3)^2 + \frac{1}{4}(2b+3)(2b-3)}{\frac{27}{16}(-2b+1)^2} = \frac{4(2b-3)[2(2b-3)+(2b+3)]}{27(-2b+1)^2} \\ &= \frac{4(2b-3)[4b-6+2b+3]}{27(2b-1)^2} = \frac{4(2b-3)[6b-3]}{27(2b-1)^2} = \frac{4(2b-3)}{9(2b-1)}. \end{aligned} \tag{2.66}$$

The proof of Corollary 2.9 is thus complete. □

Corollary 2.10. *Let $a \in [0, 1/2)$, $b \in (1/2, 1]$, $\alpha, \beta \in \mathbb{R}$ satisfy $\alpha b + \beta = 1$ and*

$$\int_a^b x((\alpha x + \beta) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_a^b ((\alpha x + \beta) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.67}$$

Then

$$\alpha = -\frac{16}{9(2a-1)} \qquad \textit{and} \qquad \beta = \frac{10a+3}{9(2a-1)}. \tag{2.68}$$

Proof of Corollary 2.10. Note that Lemma 2.8 and the assumption that $\alpha b + \beta = 1$ prove that

$$0 = -1 + \frac{3b(2b-1)(2a-1)}{2(a-b)^3} - \frac{(2b-1)(8a^2 + a(2b-3) + b(2b-3))}{4(a-b)^3}. \tag{2.69}$$

Hence, we obtain that

$$\begin{aligned} 0 &= -4(a-b)^3 + 6b(2b-1)(2a-1) - (2b-1)(8a^2 + a(2b-3) + b(2b-3)) \\ &= -4(a-b)^3 + (2b-1)(12ab - 6b - 8a^2 - 2ab + 3a - 2b^2 + 3b) \\ &= -4a^3 + 4b^3 + 12a^2b - 12ab^2 + (2b-1)(10ab - 3b - 8a^2 + 3a - 2b^2) \\ &= -4a^3 - 4a^2b + 8ab^2 - 6b^2 + 6ab - 10ab + 3b + 8a^2 - 3a + 2b^2 \\ &= -4a^3 - 4a^2b + 8ab^2 - 4ab + 8a^2 - 4b^2 - 3a + 3b \\ &= a(-4a^2 - 8ab + 8a + 4b - 3) + b(4a^2 + 8ab - 8a - 4b + 3) \\ &= (a-b)(4b - 8ab + 8a - 4a^2 - 3) \\ &= (a-b)(4b(1-2a) + 2a(1-2a) + 6a - 3) = (a-b)(1-2a)(2a+4b-3). \end{aligned} \tag{2.70}$$

This, the assumption that $a \in [0, 1/2)$, and the assumption that $b \in (1/2, 1]$ imply that $b = -a/2 + 3/4$. Combining this with Lemma 2.8 demonstrates that

$$\begin{aligned} \alpha &= \frac{3(2a-1)(-a+\frac{1}{2})}{2(\frac{3a}{2}-\frac{3}{4})^3} = -\frac{\frac{3}{2}(2a-1)^2}{\frac{27}{32}(2a-1)^3} = -\frac{16}{9(2a-1)} \qquad \text{and} \\ \beta &= -\frac{(-a+\frac{1}{2})[8a^2 + a(-a-\frac{3}{2}) + (-\frac{a}{2}+\frac{3}{4})(-a-\frac{3}{2})]}{4(\frac{3a}{2}-\frac{3}{4})^3} \\ &= -\frac{\frac{1}{2}(-2a+1)(\frac{15a^2}{2}-\frac{3a}{2}-\frac{9}{8})}{\frac{27}{16}(2a-1)^3} = \frac{8(2a-1)(\frac{15a^2}{2}-\frac{3a}{2}-\frac{9}{8})}{27(2a-1)^3} = \frac{8(\frac{5a^2}{2}-\frac{a}{2}-\frac{3}{8})}{9(2a-1)^3} \\ &= \frac{20a^2-4a-3}{9(2a-1)^2} = \frac{(2a-1)(10a+3)}{9(2a-1)^2} = \frac{10a+3}{9(2a-1)}. \end{aligned} \tag{2.71}$$

The proof of Corollary 2.10 is thus complete. □

Lemma 2.11. *Let $N \in \mathbb{N}$, $x_0, x_1, \dots, x_N, \alpha_1, \alpha_2, \dots, \alpha_N, \beta_1, \beta_2, \dots, \beta_N, a, b \in \mathbb{R}$, $f \in C([a, b], \mathbb{R})$ satisfy for all $i \in \{1, 2, \dots, N\}$, $x \in [x_{i-1}, x_i]$ that $a = x_0 < x_1 < \dots < x_N = b$, $f(x) = \alpha_i x + \beta_i$, and*

$$\int_{x_{i-1}}^{x_i} f(x)\, dx = 0. \tag{2.72}$$

Then it holds for all $i \in \{1, 2, \dots, N\}$ that $f(x_0) = (-1)^i f(x_i)$.

Proof of Lemma 2.11. Observe that (2.72) implies for all $i \in \{1, 2, \dots, N\}$ that

$$\beta_i = -\frac{\alpha_i}{2}(x_i + x_{i-1}). \tag{2.73}$$

This proves for all $i \in \{1, 2, \dots, N\}$, $x \in [x_{i-1}, x_i]$ that $f(x) = \alpha_i x - \alpha_i/2(x_i + x_{i-1})$. Therefore, we obtain for all $i \in \{1, 2, \dots, N\}$ that $f(x_{i-1}) = -f(x_i)$. This establishes for all $i \in \{1, 2, \dots, N\}$ that $f(x_0) = (-1)^i f(x_i)$. The proof of Lemma 2.11 is thus complete. □

Lemma 2.12. *Let $a \in \mathbb{R}$, $b \in (a, \infty)$, $c \in (b, \infty)$, $\alpha_1, \alpha_2 \in \mathbb{R}\backslash\{0\}$, $\beta_1, \beta_2 \in \mathbb{R}$ satisfy*

$$\int_a^b (\alpha_1 x + \beta_1)\, \mathrm{d}x = \int_b^c (\alpha_2 x + \beta_2)\, \mathrm{d}x = \int_a^b x(\alpha_1 x + \beta_1)\, \mathrm{d}x + \int_b^c x(\alpha_2 x + \beta_2)\, \mathrm{d}x = 0 \tag{2.74}$$

and $\alpha_1 b + \beta_1 = \alpha_2 b + \beta_2$. Then $b - a = c - b$, $\beta_1 = -\alpha_1(a+b)/2$, and $\beta_2 = -\alpha_2(b+c)/2$.

Proof of Lemma 2.12. Note that the assumption that

$$\int_a^b (\alpha_1 x + \beta_1)\, \mathrm{d}x = \int_b^c (\alpha_2 x + \beta_2)\, \mathrm{d}x = 0 \tag{2.75}$$

implies that

$$\beta_1 = -\frac{\alpha_1(a+b)}{2} \qquad \text{and} \qquad \beta_2 = -\frac{\alpha_2(b+c)}{2}. \tag{2.76}$$

Combining this with the assumption that $\alpha_1 b + \beta_1 = \alpha_2 b + \beta_2$ demonstrates that

$$\frac{\alpha_1(b-a)}{2} = \alpha_1 b - \frac{\alpha_1(a+b)}{2} = \alpha_2 b - \frac{\alpha_2(b+c)}{2} = \frac{\alpha_2(b-c)}{2}. \tag{2.77}$$

Hence, we obtain that

$$\alpha_2 = \frac{\alpha_1(b-a)}{b-c}. \tag{2.78}$$

This, (2.76), and the assumption that

$$\int_{\mathscr{a}}^{\mathscr{b}} x(\alpha_1 x + \beta_1)\,\mathrm{d}x + \int_{\mathscr{b}}^{\mathscr{c}} x(\alpha_2 x + \beta_2)\,\mathrm{d}x = 0 \tag{2.79}$$

ensure that

$$\begin{aligned}
0 &= \int_{\mathscr{a}}^{\mathscr{b}} x\left(\alpha_1 x - \frac{\alpha_1(\mathscr{a}+\mathscr{b})}{2}\right)\mathrm{d}x + \int_{\mathscr{b}}^{\mathscr{c}} x\left(\alpha_2 x - \frac{\alpha_2(\mathscr{b}+\mathscr{c})}{2}\right)\mathrm{d}x \\
&= \int_{\mathscr{a}}^{\mathscr{b}} \alpha_1 x\left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right)\mathrm{d}x + \int_{\mathscr{b}}^{\mathscr{c}} \alpha_2 x\left(x - \frac{\mathscr{b}+\mathscr{c}}{2}\right)\mathrm{d}x \\
&= \int_{\mathscr{a}}^{\mathscr{b}} \alpha_1 x\left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right)\mathrm{d}x + \frac{\mathscr{b}-\mathscr{a}}{\mathscr{b}-\mathscr{c}}\int_{\mathscr{b}}^{\mathscr{c}} \alpha_1 x\left(x - \frac{\mathscr{b}+\mathscr{c}}{2}\right)\mathrm{d}x.
\end{aligned} \tag{2.80}$$

Therefore, we obtain that

$$\begin{aligned}
0 &= \int_{\mathscr{a}}^{\mathscr{b}} x\left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right)\mathrm{d}x - \frac{\mathscr{b}-\mathscr{a}}{\mathscr{c}-\mathscr{b}}\int_{\mathscr{b}}^{\mathscr{c}} x\left(x - \frac{\mathscr{b}+\mathscr{c}}{2}\right)\mathrm{d}x \\
&= \int_{\mathscr{a}}^{\mathscr{b}} (x-\mathscr{a})\left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right)\mathrm{d}x - \frac{\mathscr{b}-\mathscr{a}}{\mathscr{c}-\mathscr{b}}\int_{\mathscr{b}}^{\mathscr{c}} (x-\mathscr{b})\left(x - \frac{\mathscr{b}+\mathscr{c}}{2}\right)\mathrm{d}x \\
&= \int_0^{\mathscr{b}-\mathscr{a}} x\left(x - \frac{\mathscr{b}-\mathscr{a}}{2}\right)\mathrm{d}x - \frac{\mathscr{b}-\mathscr{a}}{\mathscr{c}-\mathscr{b}}\int_0^{\mathscr{c}-\mathscr{b}} x\left(x - \frac{\mathscr{c}-\mathscr{b}}{2}\right)\mathrm{d}x \\
&= (\mathscr{b}-\mathscr{a})^3\int_0^1 x\left(x-\frac{1}{2}\right)\mathrm{d}x - (\mathscr{b}-\mathscr{a})(\mathscr{c}-\mathscr{b})^2\int_0^1 x\left(x-\frac{1}{2}\right)\mathrm{d}x.
\end{aligned} \tag{2.81}$$

Hence, we obtain that

$$0 = [(\mathscr{b}-\mathscr{a})^2 - (\mathscr{c}-\mathscr{b})^2]\int_0^1 x\left(x-\frac{1}{2}\right)\mathrm{d}x = \frac{1}{12}[(\mathscr{b}-\mathscr{a})^2 - (\mathscr{c}-\mathscr{b})^2]. \tag{2.82}$$

Therefore, we obtain that $\mathscr{b}-\mathscr{a} = \mathscr{c}-\mathscr{b}$. Combining this with (2.76) establishes that $\mathscr{b}-\mathscr{a} = \mathscr{c}-\mathscr{b}$, $\beta_1 = -\alpha_1(\mathscr{a}+\mathscr{b})/2$, and $\beta_2 = -\alpha_2(\mathscr{b}+\mathscr{c})/2$. The proof of Lemma 2.12 is thus complete. ☐

Lemma 2.13. *Let* $f\colon \mathbb{R}^3 \to \mathbb{R}$ *satisfy for all* $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [0, 1/2]$ *that satisfy*

$$f(\mathbf{a},\mathbf{n},\mathbf{q}) = \frac{\mathbf{q}\mathbf{n}^2}{3} + \frac{\mathbf{a}^2(1-\mathbf{q}^3)}{3} + \frac{1}{2} - \frac{3\mathbf{a}}{4} + (-\mathbf{a}\mathbf{q}+\mathbf{n})^2(1-\mathbf{q}) + (\mathbf{a}\mathbf{q}-\mathbf{n}) + \mathbf{a}(-\mathbf{a}\mathbf{q}+\mathbf{n})(1-\mathbf{q}^2). \tag{2.83}$$

Then

$$f(\mathbf{a},\mathbf{n},\mathbf{q}) \geq f(16/9, 0, 1/4) = 1/18. \tag{2.84}$$

Proof of Lemma 2.13. Observe that for all $\mathbf{q} \in [0, 1/2]$ it holds that

$$\liminf_{\|(\mathbf{a},\mathbf{n})\|\to\infty} f(\mathbf{a},\mathbf{n},\mathbf{q}) = \infty. \tag{2.85}$$

This implies that the minimum of f occurs in the case $\nabla f(\mathbf{a},\mathbf{n},\mathbf{q}) = 0$, in the case $\mathbf{q} = 0$, or in the case $\mathbf{q} = 1/2$. Furthermore, note that (2.83) assures for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [0, 1/2]$ that

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{a}} f(\mathbf{a},\mathbf{n},\mathbf{q}) &= \frac{2\mathbf{a}(1-\mathbf{q}^3)}{3} - \frac{3}{4} - 2\mathbf{q}(-\mathbf{a}\mathbf{q}+\mathbf{n})(1-\mathbf{q}) + \mathbf{q} + (-\mathbf{a}\mathbf{q}+\mathbf{n})(1-\mathbf{q}^2) \\
&\quad + \mathbf{a}(-\mathbf{q})(1-\mathbf{q}^2) = \frac{2\mathbf{a}}{3} - \frac{2\mathbf{a}\mathbf{q}^3}{3} - \frac{3}{4} + 2\mathbf{a}\mathbf{q}^2 - 2\mathbf{n}\mathbf{q} - 2\mathbf{a}\mathbf{q}^3 + 2\mathbf{n}\mathbf{q}^2 \\
&\quad + \mathbf{q} - \mathbf{a}\mathbf{q} + \mathbf{n} + \mathbf{a}\mathbf{q}^3 - \mathbf{n}\mathbf{q}^2 - \mathbf{a}\mathbf{q} + \mathbf{a}\mathbf{q}^3 \\
&= \frac{2\mathbf{a}}{3} - \frac{2\mathbf{a}\mathbf{q}^3}{3} - \frac{3}{4} + 2\mathbf{a}\mathbf{q}^2 - 2\mathbf{n}\mathbf{q} + \mathbf{n}\mathbf{q}^2 + \mathbf{q} - 2\mathbf{a}\mathbf{q} + \mathbf{n}, \\
\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a},\mathbf{n},\mathbf{q}) &= \frac{2\mathbf{n}\mathbf{q}}{3} + 2(-\mathbf{a}\mathbf{q}+\mathbf{n})(1-\mathbf{q}) - 1 + \mathbf{a}(1-\mathbf{q}^2) \\
&= \frac{2\mathbf{n}\mathbf{q}}{3} - 2\mathbf{a}\mathbf{q} + 2\mathbf{n} + 2\mathbf{a}\mathbf{q}^2 - 2\mathbf{n}\mathbf{q} - 1 + \mathbf{a} - \mathbf{a}\mathbf{q}^2 \\
&= -\frac{4\mathbf{n}\mathbf{q}}{3} - 2\mathbf{a}\mathbf{q} + 2\mathbf{n} + \mathbf{a}\mathbf{q}^2 - 1 + \mathbf{a}, \qquad \text{and} \\
\frac{\partial}{\partial \mathbf{q}} f(\mathbf{a},\mathbf{n},\mathbf{q}) &= \frac{\mathbf{n}^2}{3} - \mathbf{a}^2\mathbf{q}^2 - 2\mathbf{a}(-\mathbf{a}\mathbf{q}+\mathbf{n})(1-\mathbf{q}) - (-\mathbf{a}\mathbf{q}+\mathbf{n})^2 + \mathbf{a} - \mathbf{a}^2(1-\mathbf{q}^2) \\
&\quad - 2\mathbf{a}(-\mathbf{a}\mathbf{q}+\mathbf{n})\mathbf{q} = \frac{\mathbf{n}^2}{3} - \mathbf{a}^2\mathbf{q}^2 + 2\mathbf{a}^2\mathbf{q} - 2\mathbf{a}\mathbf{n} - 2\mathbf{a}^2\mathbf{q}^2 + 2\mathbf{a}\mathbf{n}\mathbf{q} - \mathbf{a}^2\mathbf{q}^2 \\
&\quad - \mathbf{n}^2 + 2\mathbf{a}\mathbf{n}\mathbf{q} + \mathbf{a} - \mathbf{a}^2 + \mathbf{a}^2\mathbf{q}^2 + 2\mathbf{a}^2\mathbf{q}^2 - 2\mathbf{a}\mathbf{n}\mathbf{q} \\
&= -\frac{2\mathbf{n}^2}{3} - \mathbf{a}^2\mathbf{q}^2 + 2\mathbf{a}^2\mathbf{q} - 2\mathbf{a}\mathbf{n} + 2\mathbf{a}\mathbf{n}\mathbf{q} + \mathbf{a} - \mathbf{a}^2.
\end{aligned} \tag{2.86}$$

This implies that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [0, 1/2]$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = 0$ it holds that

$$0 = -\frac{4\mathbf{q}\mathbf{n}}{3} - 2\mathbf{a}\mathbf{q} + 2\mathbf{n} + \mathbf{a}\mathbf{q}^2 - 1 + \mathbf{a} = 2\mathbf{n}\left(-\frac{2\mathbf{q}}{3} + 1\right) - 2\mathbf{a}\mathbf{q} + \mathbf{a}\mathbf{q}^2 - 1 + \mathbf{a}. \tag{2.87}$$

Hence, we obtain that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [0, 1/2]$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = 0$ it holds that

$$\mathbf{n} = \frac{1}{2}(2\mathbf{a}\mathbf{q} - \mathbf{a}\mathbf{q}^2 + 1 - \mathbf{a})\left(-\frac{2\mathbf{q}}{3} + 1\right)^{-1}. \tag{2.88}$$

This and (2.86) assure that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [0, 1/2]$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = \frac{\partial}{\partial \mathbf{q}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = 0$ it holds that

$$\begin{aligned}
0 &= -\frac{1}{6}(2\mathbf{a}\mathbf{q} - \mathbf{a}\mathbf{q}^2 + 1 - \mathbf{a})^2\left(-\frac{2\mathbf{q}}{3} + 1\right)^{-2} - \mathbf{a}^2\mathbf{q}^2 + 2\mathbf{a}^2\mathbf{q} + \mathbf{a} - \mathbf{a}^2 \\
&\quad - \mathbf{a}(2\mathbf{a}\mathbf{q} - \mathbf{a}\mathbf{q}^2 + 1 - \mathbf{a})\left(-\frac{2\mathbf{q}}{3} + 1\right)^{-1} + \mathbf{a}\mathbf{q}(2\mathbf{a}\mathbf{q} - \mathbf{a}\mathbf{q}^2 + 1 - \mathbf{a})\left(-\frac{2\mathbf{q}}{3} + 1\right)^{-1} \\
&= -\frac{1}{6}(2\mathbf{a}\mathbf{q} - \mathbf{a}\mathbf{q}^2 + 1 - \mathbf{a})^2\left(-\frac{2\mathbf{q}}{3} + 1\right)^{-2} + (-\mathbf{a}^2\mathbf{q}^2 + 2\mathbf{a}^2\mathbf{q} + \mathbf{a} - \mathbf{a}^2) \\
&\quad + (2\mathbf{a}\mathbf{q} - \mathbf{a}\mathbf{q}^2 + 1 - \mathbf{a})\left(-\frac{2\mathbf{q}}{3} + 1\right)^{-1}(-\mathbf{a} + \mathbf{a}\mathbf{q}).
\end{aligned} \tag{2.89}$$

Therefore, we obtain that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [0, 1/2]$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = \frac{\partial}{\partial \mathbf{q}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = 0$

it holds that

$$\begin{aligned} 0 &= (2\mathbf{aq} - \mathbf{aq}^2 + 1 - \mathbf{a})\bigg(- 2\mathbf{aq} + \mathbf{aq}^2 - 1 + \mathbf{a} + 6\mathbf{a}\left(-\frac{2\mathbf{q}}{3} + 1\right)^2 \\ &\quad + 6(-\mathbf{a} + \mathbf{aq})\left(-\frac{2\mathbf{q}}{3} + 1\right)\bigg) = (2\mathbf{aq} - \mathbf{aq}^2 + 1 - \mathbf{a})\bigg(- 2\mathbf{aq} + \mathbf{aq}^2 - 1 + \mathbf{a} \\ &\quad + \frac{8\mathbf{aq}^2}{3} + 6\mathbf{a} - 8\mathbf{aq} + 4\mathbf{aq} - 6\mathbf{a} - 4\mathbf{aq}^2 + 6\mathbf{aq}\bigg) \\ &= (2\mathbf{aq} - \mathbf{aq}^2 + 1 - \mathbf{a})\bigg(- \frac{\mathbf{aq}^2}{3} - 1 + \mathbf{a}\bigg) \\ &= \left(\mathbf{q} - 1 - \frac{\sqrt{\mathbf{a}}}{\mathbf{a}}\right)\left(\mathbf{q} - 1 + \frac{\sqrt{\mathbf{a}}}{\mathbf{a}}\right)\left(\mathbf{q} - \sqrt{3 - \frac{3}{\mathbf{a}}}\right)\left(\mathbf{q} + \sqrt{3 - \frac{3}{\mathbf{a}}}\right). \end{aligned} \tag{2.90}$$

Observe that (2.88) assures that in the case $\mathbf{q} = 1 + \sqrt{\mathbf{a}}/\mathbf{a}$ for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, \mathbf{q}) = 0$ it holds that $\mathbf{n} = 0$ and

$$\begin{aligned} \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, \mathbf{q}) &= \frac{2\mathbf{a}}{3} - \frac{2\mathbf{aq}^3}{3} - \frac{3}{4} + 2\mathbf{aq}^2 - 2\mathbf{nq} + \mathbf{nq}^2 + \mathbf{q} - 2\mathbf{aq} + \mathbf{n} \\ &= \frac{2\mathbf{a}}{3} - \frac{2\mathbf{a}(1 + \sqrt{\mathbf{a}}/\mathbf{a})^3}{3} - \frac{3}{4} + 2\mathbf{a}(1 + \sqrt{\mathbf{a}}/\mathbf{a})^2 + 1 + \sqrt{\mathbf{a}}/\mathbf{a} - 2\mathbf{a}(1 + \sqrt{\mathbf{a}}/\mathbf{a}) \\ &= \frac{3\sqrt{\mathbf{a}} + 4}{12\sqrt{\mathbf{a}}} \geq \frac{3}{12}. \end{aligned} \tag{2.91}$$

In the following we distinguish between the case $\mathbf{q} = 1 - \sqrt{\mathbf{a}}/\mathbf{a}$, the case $\mathbf{q} = (3 - 3/\mathbf{a})^{1/2}$, the case $\mathbf{q} = 0$, and the case $\mathbf{q} = 1/2$. We first prove (2.84) in the case

$$\mathbf{q} = 1 - \frac{\sqrt{\mathbf{a}}}{\mathbf{a}}. \tag{2.92}$$

Note that (2.86), (2.88), and (2.92) imply that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\nabla f(\mathbf{a}, \mathbf{n}, \mathbf{q}) = 0$ it holds that $\mathbf{n} = 0$ and

$$\begin{aligned} 0 = \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, \mathbf{q}) &= \frac{2\mathbf{a}}{3} - \frac{2\mathbf{aq}^3}{3} - \frac{3}{4} + 2\mathbf{aq}^2 - 2\mathbf{nq} + \mathbf{nq}^2 + \mathbf{q} - 2\mathbf{aq} + \mathbf{n} \\ &= \frac{2\mathbf{a}}{3} - \frac{2\mathbf{a}(1 - \sqrt{\mathbf{a}}/\mathbf{a})^3}{3} - \frac{3}{4} + 2\mathbf{a}(1 - \sqrt{\mathbf{a}}/\mathbf{a})^2 + 1 - \sqrt{\mathbf{a}}/\mathbf{a} - 2\mathbf{a}(1 - \sqrt{\mathbf{a}}/\mathbf{a}) \\ &= \frac{3\sqrt{\mathbf{a}} - 4}{12\sqrt{\mathbf{a}}}. \end{aligned} \tag{2.93}$$

Hence, we obtain that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\nabla f(\mathbf{a}, \mathbf{n}, \mathbf{q}) = 0$ it holds that $\mathbf{a} = 16/9$. This and the fact that $\mathbf{n} = 0$ ensure that

$$f\left(16/9, 0, 1 - \frac{4/3}{16/9}\right) = f(16/9, 0, 1/4) = \frac{1}{18}. \tag{2.94}$$

This implies (2.84) in the case $\mathbf{q} = 1 - \frac{\sqrt{\mathbf{a}}}{\mathbf{a}}$. We show (2.84) in the case

$$\mathbf{q} = (3 - 3/\mathbf{a})^{1/2}. \tag{2.95}$$

Observe that (2.88) and (2.95) show that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, \mathbf{q}) = 0$ it holds that

$$\mathbf{n} = \frac{3\sqrt{\mathbf{a}}(-2\mathbf{a} + \sqrt{3\mathbf{a}^2 - 3\mathbf{a}} + 2)}{2\sqrt{3\mathbf{a} - 3} - 3\sqrt{\mathbf{a}}}. \tag{2.96}$$

This and (2.86) ensure that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\nabla f(\mathbf{a}, \mathbf{n}, \mathbf{q}) = 0$ it holds that

$$\mathbf{a} = \frac{9}{8}, \qquad \mathbf{n} = \frac{3\sqrt{3}}{8}, \qquad \text{and} \qquad \mathbf{q} = \frac{1}{\sqrt{3}}. \tag{2.97}$$

Therefore, we obtain that

$$f(9/8, 3\sqrt{3}/8, 1/\sqrt{3}) = \frac{5}{64}. \tag{2.98}$$

This demonstrates (2.84) in the case $\mathbf{q} = (3 - 3/\mathbf{a})^{1/2}$. We now prove (2.84) in the case

$$\mathbf{q} = 0. \tag{2.99}$$

Note that (2.88) and (2.99) ensure that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, 0) = 0$ it holds that

$$\mathbf{n} = \frac{1}{2}(1 - \mathbf{a}). \tag{2.100}$$

Combining this and (2.86) shows that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, 0) = \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, 0) = 0$ it holds that

$$0 = \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, 0) = \frac{2\mathbf{a}}{3} - \frac{3}{4} + \mathbf{n} = \frac{2\mathbf{a}}{3} - \frac{3}{4} + \frac{1}{2}(1 - \mathbf{a}) = \frac{\mathbf{a}}{6} - \frac{1}{4}. \tag{2.101}$$

This and (2.85) imply that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ it holds that

$$f(\mathbf{a}, \mathbf{n}, 0) \geq f(3/2, -1/4, 0) = \frac{1}{16}. \tag{2.102}$$

This assures (2.84) in the case $\mathbf{q} = 0$. We demonstrate (2.84) in the case

$$\mathbf{q} = \frac{1}{2}. \tag{2.103}$$

Observe that (2.88) and (2.103) ensure that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, 1/2) = 0$ it holds that

$$\mathbf{n} = \frac{1}{2}\left(\mathbf{a} - \frac{\mathbf{a}}{4} + 1 - \mathbf{a}\right)\left(-\frac{2}{6} + 1\right)^{-1} = -\frac{3\mathbf{a}}{16} + \frac{3}{4}. \tag{2.104}$$

Combining this and (2.86) shows that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, 1/2) = \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, 1/2) = 0$ it holds that

$$\begin{aligned} 0 = \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, 1/2) &= \frac{2\mathbf{a}}{3} - \frac{\mathbf{a}}{12} - \frac{3}{4} + \frac{\mathbf{a}}{2} - \mathbf{n} + \frac{\mathbf{n}}{4} + \frac{1}{2} - \mathbf{a} + \mathbf{n} \\ &= \frac{\mathbf{a}}{12} + \frac{\mathbf{n}}{4} - \frac{1}{4} = \frac{\mathbf{a}}{12} - \frac{3\mathbf{a}}{64} + \frac{3}{16} - \frac{1}{4} = \frac{7\mathbf{a}}{192} - \frac{1}{16}. \end{aligned} \tag{2.105}$$

This and (2.85) imply that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ it holds that

$$f(\mathbf{a}, \mathbf{n}, 1/2) \geq f(12/7, 3/7, 1/2) = \frac{1}{14}. \tag{2.106}$$

This ensures (2.84) in the case $\mathbf{q} = 0$. The proof of Lemma 2.13 is thus complete. □

Lemma 2.14. *Let $f \colon \mathbb{R}^3 \to \mathbb{R}$ satisfy for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [1/2, 1]$ that satisfy*

$$\begin{aligned} f(\mathbf{a}, \mathbf{n}, \mathbf{q}) = \frac{\mathbf{a}^2\mathbf{q}^3}{3} &+ (-\mathbf{a}\mathbf{q} + \mathbf{n})^2\mathbf{q} - \mathbf{a}\left(\mathbf{q}^2 - \frac{1}{4}\right) + \mathbf{a}(-\mathbf{a}\mathbf{q} + \mathbf{n})\mathbf{q}^2 + \frac{1}{3}(1 - \mathbf{q})(\mathbf{n} - 1)^2 \\ &+ (1 + 2\mathbf{a}\mathbf{q} - 2\mathbf{n})\left(\mathbf{q} - \frac{1}{2}\right). \end{aligned} \tag{2.107}$$

Then

$$f(\mathbf{a}, \mathbf{n}, \mathbf{q}) \geq f(16/9, 1, 3/4) = 1/18. \tag{2.108}$$

Proof of Lemma 2.14. Note that for all $\mathbf{q} \in [1/2, 1]$ it holds that

$$\liminf_{\|(\mathbf{a},\mathbf{n})\|\to\infty} f(\mathbf{a},\mathbf{n},\mathbf{q}) = \infty. \tag{2.109}$$

This implies that the minimum of f occurs in the case $\nabla f(\mathbf{a},\mathbf{n},\mathbf{q}) = 0$, in the case $\mathbf{q} = 1/2$, or in the case $\mathbf{q} = 1$. Furthermore, observe that (2.107) assures for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [0, 1/2]$ that

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{a}} f(\mathbf{a},\mathbf{n},\mathbf{q}) &= \frac{2\mathbf{a}\mathbf{q}^3}{3} - 2(-\mathbf{a}\mathbf{q}+\mathbf{n})\mathbf{q}^2 - \left(\mathbf{q}^2 - \frac{1}{4}\right) + (-\mathbf{a}\mathbf{q}+\mathbf{n})\mathbf{q}^2 - \mathbf{a}\mathbf{q}^3 \\
&\quad + 2\mathbf{q}\left(\mathbf{q} - \frac{1}{2}\right) = \frac{2\mathbf{a}\mathbf{q}^3}{3} + 2\mathbf{a}\mathbf{q}^3 - 2\mathbf{n}\mathbf{q}^2 - \mathbf{q}^2 + \frac{1}{4} - \mathbf{a}\mathbf{q}^3 + \mathbf{n}\mathbf{q}^2 \\
&\quad - \mathbf{a}\mathbf{q}^3 + 2\mathbf{q}^2 - \mathbf{q} \\
&= \frac{2\mathbf{a}\mathbf{q}^3}{3} - \mathbf{n}\mathbf{q}^2 + \mathbf{q}^2 + \frac{1}{4} - \mathbf{q}, \\
\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a},\mathbf{n},\mathbf{q}) &= 2(-\mathbf{a}\mathbf{q}+\mathbf{n})\mathbf{q} + \mathbf{a}\mathbf{q}^2 + \frac{2}{3}(1-\mathbf{q})(\mathbf{n}-1) - 2\left(\mathbf{q} - \frac{1}{2}\right) \\
&= -2\mathbf{a}\mathbf{q}^2 + 2\mathbf{n}\mathbf{q} + \mathbf{a}\mathbf{q}^2 + \frac{2\mathbf{n}}{3} - \frac{2}{3} - \frac{2\mathbf{n}\mathbf{q}}{3} + \frac{2\mathbf{q}}{3} - 2\mathbf{q} + 1 \\
&= -\mathbf{a}\mathbf{q}^2 + \frac{4\mathbf{n}\mathbf{q}}{3} + \frac{2\mathbf{n}}{3} + \frac{1}{3} - \frac{4\mathbf{q}}{3}, \qquad \text{and} \\
\frac{\partial}{\partial \mathbf{q}} f(\mathbf{a},\mathbf{n},\mathbf{q}) &= \mathbf{a}^2\mathbf{q}^2 - 2\mathbf{a}\mathbf{q}(-\mathbf{a}\mathbf{q}+\mathbf{n}) + (-\mathbf{a}\mathbf{q}+\mathbf{n})^2 - 2\mathbf{a}\mathbf{q} - \mathbf{a}^2\mathbf{q}^2 + 2\mathbf{a}\mathbf{q}(-\mathbf{a}\mathbf{q}+\mathbf{n}) \\
&\quad - \frac{1}{3}(\mathbf{n}-1)^2 + 2\mathbf{a}\left(\mathbf{q} - \frac{1}{2}\right) + 1 + 2\mathbf{a}\mathbf{q} - 2\mathbf{n} \\
&= \mathbf{a}^2\mathbf{q}^2 + 2\mathbf{a}^2\mathbf{q}^2 - 2\mathbf{a}\mathbf{q}\mathbf{n} + \mathbf{a}^2\mathbf{q}^2 + \mathbf{n}^2 - 2\mathbf{a}\mathbf{q}\mathbf{n} - 2\mathbf{a}\mathbf{q} - \mathbf{a}^2\mathbf{q}^2 - 2\mathbf{a}^2\mathbf{q}^2 \\
&\quad + 2\mathbf{a}\mathbf{q}\mathbf{n} - \frac{\mathbf{n}^2}{3} - \frac{1}{3} + \frac{2\mathbf{n}}{3} + 2\mathbf{a}\mathbf{q} - \mathbf{a} + 1 + 2\mathbf{a}\mathbf{q} - 2\mathbf{n} \\
&= \mathbf{a}^2\mathbf{q}^2 - 2\mathbf{a}\mathbf{n}\mathbf{q} + \frac{2\mathbf{n}^2}{3} + 2\mathbf{a}\mathbf{q} + \frac{2}{3} - \frac{4\mathbf{n}}{3} - \mathbf{a}.
\end{aligned} \tag{2.110}$$

This implies that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [1/2, 1]$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = 0$ it holds that

$$0 = -\mathbf{a}\mathbf{q}^2 + \frac{4\mathbf{n}\mathbf{q}}{3} + \frac{2\mathbf{n}}{3} + \frac{1}{3} - \frac{4\mathbf{q}}{3} = \frac{2\mathbf{n}}{3}(2\mathbf{q}+1) - \mathbf{a}\mathbf{q}^2 + \frac{1}{3} - \frac{4\mathbf{q}}{3}. \tag{2.111}$$

Hence, we obtain that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [1/2, 1]$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = 0$ it holds that

$$\mathbf{n} = \frac{3}{2}\left(\mathbf{a}\mathbf{q}^2 - \frac{1}{3} + \frac{4\mathbf{q}}{3}\right)(2\mathbf{q}+1)^{-1}. \tag{2.112}$$

This and (2.110) assure that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$, $\mathbf{q} \in [0, 1/2]$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = \frac{\partial}{\partial \mathbf{q}} f(\mathbf{a},\mathbf{n},\mathbf{q}) = 0$ it holds that

$$\begin{aligned}
0 &= \mathbf{a}^2\mathbf{q}^2 - 3\mathbf{a}\mathbf{q}\left(\mathbf{a}\mathbf{q}^2 - \frac{1}{3} + \frac{4\mathbf{q}}{3}\right)(2\mathbf{q}+1)^{-1} + \frac{3}{2}\left(\mathbf{a}\mathbf{q}^2 - \frac{1}{3} + \frac{4\mathbf{q}}{3}\right)^2(2\mathbf{q}+1)^{-2} \\
&\quad + 2\mathbf{a}\mathbf{q} + \frac{2}{3} - 2\left(\mathbf{a}\mathbf{q}^2 - \frac{1}{3} + \frac{4\mathbf{q}}{3}\right)(2\mathbf{q}+1)^{-1} - \mathbf{a} \\
&= \mathbf{a}^2\mathbf{q}^2 + 2\mathbf{a}\mathbf{q} - \mathbf{a} + \frac{2}{3} + \left(-3\mathbf{a}^2\mathbf{q}^3 + \mathbf{a}\mathbf{q} - 4\mathbf{a}\mathbf{q}^2 - 2\mathbf{a}\mathbf{q}^2 + \frac{3}{2} - \frac{8\mathbf{q}}{3}\right)(2\mathbf{q}+1)^{-1} \\
&\quad + \frac{3}{2}\left(\mathbf{a}^2\mathbf{q}^4 + \frac{1}{9} + \frac{16\mathbf{q}^2}{9} - \frac{2\mathbf{a}\mathbf{q}^2}{3} + \frac{8\mathbf{a}\mathbf{q}^3}{3} - \frac{8\mathbf{q}}{9}\right)(2\mathbf{q}+1)^{-2}.
\end{aligned} \tag{2.113}$$

Therefore, we obtain that for all $\mathbf{a},\mathbf{n}\in\mathbb{R}$, $\mathbf{q}\in[1/2,1]$ such that $\frac{\partial}{\partial\mathbf{n}}f(\mathbf{a},\mathbf{n},\mathbf{q})=\frac{\partial}{\partial\mathbf{q}}f(\mathbf{a},\mathbf{n},\mathbf{q})=0$ it holds that

$$\begin{aligned}
0 &= \left(\mathbf{a}^2\mathbf{q}^2+2\mathbf{a}\mathbf{q}-\mathbf{a}+\frac{2}{3}\right)(2\mathbf{q}+1)^2+\left(-3\mathbf{a}^2\mathbf{q}^3+\mathbf{a}\mathbf{q}-4\mathbf{a}\mathbf{q}^2-2\mathbf{a}\mathbf{q}^2+\frac{2}{3}-\frac{8\mathbf{q}}{3}\right)\\
&\quad(2\mathbf{q}+1)+\frac{3}{2}\left(\mathbf{a}^2\mathbf{q}^4+\frac{1}{9}+\frac{16\mathbf{q}^2}{9}-\frac{2\mathbf{a}\mathbf{q}^2}{3}+\frac{8\mathbf{a}\mathbf{q}^3}{3}-\frac{8\mathbf{q}}{9}\right)\\
&= \left(\mathbf{a}^2\mathbf{q}^2+2\mathbf{a}\mathbf{q}-\mathbf{a}+\frac{2}{3}\right)(4\mathbf{q}^2+1+4\mathbf{q})-6\mathbf{a}^2\mathbf{q}^4+2\mathbf{a}\mathbf{q}^2-12\mathbf{a}\mathbf{q}^3+\frac{4\mathbf{q}}{3}\\
&\quad-\frac{16\mathbf{q}^2}{3}-3\mathbf{a}^2\mathbf{q}^3+\mathbf{a}\mathbf{q}-4\mathbf{a}\mathbf{q}^2-2\mathbf{a}\mathbf{q}^2+\frac{2}{3}-\frac{8\mathbf{q}}{3}+\frac{3\mathbf{a}^2\mathbf{q}^4}{2}+\frac{1}{6}+\frac{8\mathbf{q}^2}{3}-\mathbf{a}\mathbf{q}^2\\
&\quad+4\mathbf{a}\mathbf{q}^3-\frac{4\mathbf{q}}{3}=4\mathbf{a}^2\mathbf{q}^4+8\mathbf{a}\mathbf{q}^3-4\mathbf{a}\mathbf{q}^2+\frac{8\mathbf{q}^2}{3}+\mathbf{a}^2\mathbf{q}^2+2\mathbf{a}\mathbf{q}-\mathbf{a}+\frac{2}{3}+4\mathbf{a}^2\mathbf{q}^3\\
&\quad+8\mathbf{a}\mathbf{q}^2-4\mathbf{a}\mathbf{q}+\frac{8\mathbf{q}}{3}-\frac{9\mathbf{a}^2\mathbf{q}^4}{2}-5\mathbf{a}\mathbf{q}^2-8\mathbf{a}\mathbf{q}^3+\frac{5}{6}-\frac{8\mathbf{q}^2}{3}-3\mathbf{a}^2\mathbf{q}^3+\mathbf{a}\mathbf{q}-\frac{8\mathbf{q}}{3}\\
&= -\frac{\mathbf{a}^2\mathbf{q}^4}{2}-\mathbf{a}\mathbf{q}^2+\mathbf{a}^2\mathbf{q}^2-\mathbf{a}\mathbf{q}-\mathbf{a}+\frac{3}{2}+\mathbf{a}^2\mathbf{q}^3\\
&= \frac{\mathbf{a}\mathbf{q}^2}{2}(-\mathbf{a}\mathbf{q}^2+2\mathbf{a}\mathbf{q}+2\mathbf{a}-3)+\frac{\mathbf{a}\mathbf{q}^2}{2}-\mathbf{a}\mathbf{q}-\mathbf{a}+\frac{3}{2}\\
&= \frac{1}{2}(\mathbf{a}\mathbf{q}^2-1)(-\mathbf{a}\mathbf{q}^2+2\mathbf{a}\mathbf{q}+2\mathbf{a}-3)\\
&= \frac{1}{2}(\mathbf{a}\mathbf{q}^2-1)\left(\mathbf{q}-1-\sqrt{3-\frac{3}{\mathbf{a}}}\right)\left(\mathbf{q}-1+\sqrt{3-\frac{3}{\mathbf{a}}}\right).
\end{aligned} \tag{2.114}$$

In the following we distinguish between the case $\mathbf{q}=\sqrt{\mathbf{a}}/\mathbf{a}$, the case $\mathbf{q}=1-(3-3/\mathbf{a})^{1/2}$, the case $\mathbf{q}=1/2$, and the case $\mathbf{q}=1$. We first prove (2.108) in the case

$$\mathbf{q}=\frac{\sqrt{\mathbf{a}}}{\mathbf{a}}. \tag{2.115}$$

Note that (2.110), (2.112), and (2.115) imply that for all $\mathbf{a},\mathbf{n}\in\mathbb{R}$ such that $\nabla f(\mathbf{a},\mathbf{n},\mathbf{q})=0$ it holds that

$$\mathbf{n}=\frac{3}{2}\left(\frac{\mathbf{a}^2}{\mathbf{a}^2}-\frac{1}{3}+\frac{4\sqrt{\mathbf{a}}}{3\mathbf{a}}\right)\left(\frac{2\sqrt{\mathbf{a}}}{\mathbf{a}}+1\right)^{-1}=1. \tag{2.116}$$

and

$$\begin{aligned}
0=\frac{\partial}{\partial\mathbf{a}}f(\mathbf{a},\mathbf{n},\mathbf{q})&=\frac{2\mathbf{a}\mathbf{q}^3}{3}-\mathbf{n}\mathbf{q}^2+\mathbf{q}^2+\frac{1}{4}-\mathbf{q}=\frac{2\mathbf{a}(\sqrt{\mathbf{a}}/\mathbf{a})^3}{3}+\frac{1}{4}-\sqrt{\mathbf{a}}/\mathbf{a}\\
&=\frac{3\sqrt{\mathbf{a}}-4}{12\sqrt{\mathbf{a}}}.
\end{aligned} \tag{2.117}$$

Hence, we obtain that for all $\mathbf{a},\mathbf{n}\in\mathbb{R}$ such that $\nabla f(\mathbf{a},\mathbf{n},\mathbf{q})=0$ it holds that $\mathbf{a}=16/9$. This and the fact that $\mathbf{n}=1$ ensure that

$$f\left(16/9,1,\frac{4/3}{16/9}\right)=f(16/9,0,3/4)=\frac{1}{18}. \tag{2.118}$$

This implies (2.108) in the case $\mathbf{q}=\frac{\sqrt{\mathbf{a}}}{\mathbf{a}}$. We show (2.108) in the case

$$\mathbf{q}=1-(3-3/\mathbf{a})^{1/2}. \tag{2.119}$$

Observe that (2.112) and (2.119) show that for all $\mathbf{a},\mathbf{n}\in\mathbb{R}$ such that $\frac{\partial}{\partial\mathbf{n}}f(\mathbf{a},\mathbf{n},\mathbf{q})=0$ it holds that

$$\mathbf{n}=\frac{6\mathbf{a}^2-3\mathbf{a}\sqrt{3\mathbf{a}^2-3\mathbf{a}}-3\mathbf{a}-2\sqrt{3\mathbf{a}^2-3\mathbf{a}}}{3\mathbf{a}-2\sqrt{3\mathbf{a}^2-3\mathbf{a}}} \tag{2.120}$$

This and (2.110) ensure that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\nabla f(\mathbf{a}, \mathbf{n}, \mathbf{q}) = 0$ it holds that

$$\mathbf{a} = \frac{9}{8}, \qquad \mathbf{n} = 1 - \frac{3\sqrt{3}}{8}, \qquad \text{and} \qquad \mathbf{q} = 1 - \frac{1}{\sqrt{3}}. \tag{2.121}$$

Therefore, we obtain that

$$f(9/8, 1 - 3\sqrt{3}/8, 1 - 1/\sqrt{3}) = \frac{5}{64}. \tag{2.122}$$

This demonstrates (2.108) in the case $\mathbf{q} = (3 - 3/\mathbf{a})^{1/2}$. We now prove (2.108) in the case

$$\mathbf{q} = 1/2. \tag{2.123}$$

Note that (2.112) and (2.123) ensure that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, 0) = 0$ it holds that

$$\mathbf{n} = \frac{3}{2}\left(\frac{\mathbf{a}}{4} - \frac{1}{3} + \frac{2}{3}\right)(1+1)^{-1} = \frac{3\mathbf{a}}{16} + \frac{1}{4}. \tag{2.124}$$

Combining this and (2.110) shows that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, 1/2) = \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, 1/2) = 0$ it holds that

$$0 = \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, 1/2) = \frac{2\mathbf{a}}{24} - \frac{\mathbf{n}}{4} + \frac{1}{4} + \frac{1}{4} - \frac{1}{2} = \frac{\mathbf{a}}{12} - \frac{3\mathbf{a}}{64} - \frac{1}{16} = \frac{7\mathbf{a}}{192} - \frac{1}{16}. \tag{2.125}$$

This and (2.109) imply that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ it holds that

$$f(\mathbf{a}, \mathbf{n}, 1/2) \geq f(12/7, 4/7, 1/2) = \frac{1}{14}. \tag{2.126}$$

This assures (2.108) in the case $\mathbf{q} = 1/2$. We demonstrate (2.108) in the case

$$\mathbf{q} = 1. \tag{2.127}$$

Observe that (2.112) and (2.127) ensure that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, 1) = 0$ it holds that

$$\mathbf{n} = \frac{3}{2}\left(\mathbf{a} - \frac{1}{3} + \frac{4}{3}\right)(2+1)^{-1} = \frac{\mathbf{a}}{2} + \frac{1}{2}. \tag{2.128}$$

Combining this and (2.110) shows that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ such that $\frac{\partial}{\partial \mathbf{n}} f(\mathbf{a}, \mathbf{n}, 1) = \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, 1) = 0$ it holds that

$$0 = \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, \mathbf{n}, 1/2) = \frac{2\mathbf{a}}{3} - \mathbf{n} + 1 + \frac{1}{4} - 1 = \frac{2\mathbf{a}}{3} - \frac{\mathbf{a}}{2} - \frac{1}{2} + \frac{1}{4} = \frac{\mathbf{a}}{6} - \frac{1}{4}. \tag{2.129}$$

This and (2.109) imply that for all $\mathbf{a}, \mathbf{n} \in \mathbb{R}$ it holds that

$$f(\mathbf{a}, \mathbf{n}, 1) \geq f(3/2, 5/4, 1) = \frac{1}{16}. \tag{2.130}$$

This ensures (2.108) in the case $\mathbf{q} = 0$. The proof of Lemma 2.14 is thus complete. □

Lemma 2.15. *Let $H \in \mathbb{N}$, $c \in (0, \infty)$ and let $(\mathbf{q}_{0,n})_{n\in\mathbb{N}} \subseteq [0, 1/2]$, $(\mathbf{q}_{1,n})_{n\in\mathbb{N}} \subseteq [1/2, 1]$, $(\mathbf{m}_{0,n})_{n\in\mathbb{N}}, (\mathbf{m}_{1,n})_{n\in\mathbb{N}} \subseteq [1, H]$, $(\mathbf{a}_n)_{n\in\mathbb{N}}, (\mathbf{b}_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ satisfy for all $n \in \mathbb{N}$ that*

$$\max\{|\mathbf{a}_n|, |\mathbf{b}_n|\} \geq c \tag{2.131}$$

and

$$\begin{aligned} 0 &= \lim_{n\to\infty} \left(\mathbf{q}_{1,n} - \frac{1}{2}\right)^3 (\mathbf{m}_{0,n}\mathbf{a}_n + (H - \mathbf{m}_{1,n} + 1)\mathbf{b}_n)^2 \\ &= \lim_{n\to\infty} \left(\frac{1}{2} - \mathbf{q}_{0,n}\right)^3 (\mathbf{m}_{0,n}\mathbf{a}_n + (H - \mathbf{m}_{1,n} + 1)\mathbf{b}_n)^2 \\ &= \lim_{n\to\infty} (\mathbf{b}_n)^2 (1 - \mathbf{q}_{1,n})^3 \\ &= \lim_{n\to\infty} (\mathbf{a}_n)^2 (\mathbf{q}_{0,n})^3. \end{aligned} \tag{2.132}$$

Then there exists a strictly increasing $n\colon \mathbb{N} \to \mathbb{N}$ *such that*

$$\lim_{k\to\infty} \mathbf{q}_{1,n(k)} = 1 \qquad \textit{and} \qquad \lim_{k\to\infty} \mathbf{q}_{0,n(k)} = 0. \tag{2.133}$$

Proof of Lemma 2.15. Note that (2.131) demonstrates that there exists $C \in [c, \infty]$ which satisfies $C = \max\{\limsup_{n\to\infty} |\mathbf{a}_n|, \limsup_{n\to\infty} |\mathbf{b}_n|\}$. In the following we distinguish between the case $\limsup_{n\to\infty} |\mathbf{a}_n| = C$ and the case $\limsup_{n\to\infty} |\mathbf{b}_n| = C$. We first prove (2.133) in the case

$$\limsup_{n\to\infty} |\mathbf{a}_n| = C. \tag{2.134}$$

Observe that (2.134) shows that there exists a strictly increasing $n\colon \mathbb{N} \to \mathbb{N}$ such that $\lim_{k\to\infty} |\mathbf{a}_{n(k)}| = C$. Combining this with (2.132) implies that

$$0 = \lim_{k\to\infty} (\mathbf{a}_{n(k)})^2 (\mathbf{q}_{0,n(k)})^3 = \lim_{k\to\infty} C^2 (\mathbf{q}_{0,n(k)})^3. \tag{2.135}$$

Hence, we obtain that $\lim_{k\to\infty} \mathbf{q}_{0,n(k)} = 0$. This and (2.132) assure that

$$\begin{aligned} 0 &= \lim_{k\to\infty} (\mathbf{m}_{0,n(k)} \mathbf{a}_{n(k)} + (H - \mathbf{m}_{1,n(k)} + 1)\mathbf{b}_{n(k)})^2 \\ &= \lim_{k\to\infty} \mathbf{m}_{0,n(k)} \mathbf{a}_{n(k)} + (H - \mathbf{m}_{1,n(k)} + 1)\mathbf{b}_{n(k)}. \end{aligned} \tag{2.136}$$

Therefore, we obtain that

$$\begin{aligned} \lim_{k\to\infty} |\mathbf{b}_{n(k)}| &= \lim_{k\to\infty} |\mathbf{m}_{0,n(k)} (H - \mathbf{m}_{1,n(k)} + 1)^{-1} \mathbf{a}_{n(k)}| \\ &> \lim_{k\to\infty} |(H+1)^{-1} \mathbf{a}_{n(k)}| = (H+1)^{-1} C. \end{aligned} \tag{2.137}$$

Combining this with (2.132) demonstrates that

$$0 = \lim_{k\to\infty} (\mathbf{b}_{n(k)})^2 (1 - \mathbf{q}_{1,n(k)})^3 = \lim_{k\to\infty} (1 - \mathbf{q}_{1,n(k)})^3. \tag{2.138}$$

This and (2.135) show that $\lim_{k\to\infty} \mathbf{q}_{1,n(k)} = 1$ and $\lim_{k\to\infty} \mathbf{q}_{0,n(k)} = 0$. This implies (2.133) in the case $\limsup_{n\to\infty} |\mathbf{a}_n| = C$. In the next step we prove (2.133) in the case

$$\limsup_{n\to\infty} |\mathbf{b}_n| = C. \tag{2.139}$$

Note that (2.139) shows that there exists a strictly increasing $n\colon \mathbb{N} \to \mathbb{N}$ such that $\lim_{k\to\infty} |\mathbf{b}_{n(k)}| = C$. Combining this with (2.132) implies that

$$0 = \lim_{k\to\infty} (\mathbf{b}_{n(k)})^2 (1 - \mathbf{q}_{1,n(k)})^3 = \lim_{k\to\infty} C^2 (1 - \mathbf{q}_{1,n(k)})^3. \tag{2.140}$$

Hence, we obtain that $\lim_{k\to\infty} \mathbf{q}_{1,n(k)} = 1$. This and (2.132) assure that

$$\begin{aligned} 0 &= \lim_{k\to\infty} (\mathbf{m}_{0,n(k)} \mathbf{a}_{n(k)} + (H - \mathbf{m}_{1,n(k)} + 1)\mathbf{b}_{n(k)})^2 \\ &= \lim_{k\to\infty} \mathbf{m}_{0,n(k)} \mathbf{a}_{n(k)} + (H - \mathbf{m}_{1,n(k)} + 1)\mathbf{b}_{n(k)}. \end{aligned} \tag{2.141}$$

Therefore, we obtain that

$$\begin{aligned} \lim_{k\to\infty} |\mathbf{a}_{n(k)}| &= \lim_{k\to\infty} |(\mathbf{m}_{0,n(k)})^{-1} (H - \mathbf{m}_{1,n(k)} + 1)\mathbf{b}_{n(k)}| \\ &> \lim_{k\to\infty} |(H+1)^{-1} \mathbf{b}_{n(k)}| = (H+1)^{-1} C. \end{aligned} \tag{2.142}$$

Combining this with (2.132) demonstrates that

$$0 = \lim_{k\to\infty} (\mathbf{a}_{n(k)})^2 (\mathbf{q}_{0,n(k)})^3 = \lim_{k\to\infty} (\mathbf{q}_{0,n(k)})^3. \tag{2.143}$$

This and (2.140) show that $\lim_{k\to\infty} \mathbf{q}_{1,n(k)} = 1$ and $\lim_{k\to\infty} \mathbf{q}_{0,n(k)} = 0$. This implies (2.133) in the case $\limsup_{n\to\infty} |\mathbf{b}_n| = C$. The proof of Lemma 2.15 is thus complete. □

2.4 Properties of critical points of the risk function

Proposition 2.16. *Assume Setting 2.1 and let $h \in \mathbb{N}$. Then it holds for all $\theta \in \mathbb{R}^{\mathfrak{d}_h}$, $i \in \{1, 2, \ldots, h\}$ that*

$$\begin{aligned}
\mathcal{G}_i^h(\theta) &= 2\mathfrak{v}_i^\theta \int_{I_i^\theta} x(\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x, \\
\mathcal{G}_{h+i}^h(\theta) &= 2\mathfrak{v}_i^\theta \int_{I_i^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x, \\
\mathcal{G}_{2h+i}^h(\theta) &= 2\int_0^1 \big[\mathbb{A}_\infty(\mathfrak{w}_i^\theta x + \mathfrak{b}_i^\theta)\big](\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x, \\
\text{and} \qquad \mathcal{G}_{\mathfrak{d}_h}^h(\theta) &= 2\int_0^1 (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x.
\end{aligned} \tag{2.144}$$

Proof of Proposition 2.16. Observe that (2.3)–(2.5) imply (2.144) (cf., e.g., [22, Proposition 2.3]). The proof of Proposition 2.16 is thus complete. ☐

Corollary 2.17. *Assume Setting 2.1, let $h \in \mathbb{N}$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $x \in [0, \mathfrak{q}_{m_{0,1}^\theta}^\theta]$ that $\mathscr{N}_\infty^{h,\theta}(x) = 0$, $M_0^\theta \neq \varnothing$, and $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$. Then it holds for all $x \in [0, \mathfrak{q}_{m_{0,2}^\theta}^\theta]$ that $\mathscr{N}_\infty^{h,\theta}(x) = 0$.*

Proof of Corollary 2.17. Assume without loss of generality that $\mathfrak{q}_1^\theta \leq \ldots \leq \mathfrak{q}_h^\theta$. Note that the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, the assumption that $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$, and Proposition 2.16 imply that for all $j \in M_0^\theta$ it holds that

$$\int_0^{\mathfrak{q}_{m_{0,1}^\theta}^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}_{j-1}^\theta}^{\mathfrak{q}_j^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.145}$$

Combining this with Lemma 2.11 demonstrates that for all $j \in M_0^\theta$ it holds that $0 = \mathscr{N}_\infty^{h,\theta}(0) = \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_j^\theta)$. This and piecewise linearity of $\mathscr{N}_\infty^{h,\theta}$ assure that for all $x \in [0, \mathfrak{q}_{m_{0,2}^\theta}^\theta]$ it holds that $\mathscr{N}_\infty^{h,\theta}(x) = 0$. The proof of Corollary 2.17 is thus complete. ☐

Corollary 2.18. *Assume Setting 2.1, let $h \in \mathbb{N}$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $x \in [\mathfrak{q}_{m_{1,2}^\theta}^\theta, 1]$ that $\mathscr{N}_\infty^{h,\theta}(x) = 1$, $M_1^\theta \neq \varnothing$, and $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$. Then it holds for all $x \in [\mathfrak{q}_{m_{1,1}^\theta}^\theta, 1]$ that $\mathscr{N}_\infty^{h,\theta}(x) = 1$.*

Proof of Corollary 2.18. Assume without loss of generality that $\mathfrak{q}_1^\theta \leq \ldots \leq \mathfrak{q}_h^\theta$. Observe that the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, the assumption that $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$, and Proposition 2.16 imply that for all $j \in M_1^\theta$ it holds that

$$\int_{\mathfrak{q}_{m_{1,2}^\theta}^\theta}^1 (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}_j^\theta}^{\mathfrak{q}_{j+1}^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.146}$$

Combining this with Lemma 2.11 demonstrates that for all $j \in M_1^\theta$ it holds that $1 = \mathscr{N}_\infty^{h,\theta}(1) = \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_j^\theta)$. This and the fact that $\mathscr{N}_\infty^{h,\theta}$ is piecewise affine linear assure that for all $x \in [\mathfrak{q}_{m_{1,1}^\theta}^\theta, 1]$ it holds that $\mathscr{N}_\infty^{h,\theta}(x) = 1$. The proof of Corollary 2.18 is thus complete. ☐

Lemma 2.19. *Assume Setting 2.1, let $h \in \mathbb{N}$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $j \in \{0, 1, \ldots, h\}$ that $\mathfrak{q}_j^\theta < \mathfrak{q}_{j+1}^\theta$ and $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$, and let $i \in \{0, 1, \ldots, h\}$ satisfy $0 < \mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta$ and $\mathfrak{q}_{i+1}^\theta \leq 1/2$. Then it holds for all $x \in [0, \mathfrak{q}_{m_{0,2}^\theta}^\theta]$ that $\mathscr{N}_\infty^{h,\theta}(x) = 0$.*

Proof of Lemma 2.19. Note that the assumption that $\theta \in (\mathscr{G}^h)^{-1}(\{0\})$, the fact that $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0 < \mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta$, and Proposition 2.16 imply that

$$\int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} x(\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.147}$$

This, Lemma 2.7, and the assumption that $\mathfrak{q}_{i+1}^\theta \leq 1/2$ assure that for all $x \in [\mathfrak{q}_i^\theta, \mathfrak{q}_{i+1}^\theta]$ it holds that

$$\mathscr{N}_\infty^{h,\theta}(x) = 0. \tag{2.148}$$

Furthermore, observe that the assumption that $\theta \in (\mathscr{G}^h)^{-1}(\{0\})$, the assumption that $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$, and Proposition 2.16 imply that for all $j \in M_0^\theta$ it holds that

$$\int_{\mathfrak{q}_{j-1}^\theta}^{\mathfrak{q}_j^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.149}$$

Combining this with (2.148) and Lemma 2.11 demonstrates that for all $j \in M_0^\theta$ it holds that $\mathscr{N}_\infty^{h,\theta}(0) = \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_j^\theta) = 0$. This and the fact that $\mathscr{N}_\infty^{h,\theta}$ is piecewise affine linear ensure that for all $x \in [0, \mathfrak{q}_{m_{0,2}^\theta}^\theta]$ it holds that $\mathscr{N}_\infty^{h,\theta}(x) = 0$. The proof of Lemma 2.19 is thus complete. ◻

Lemma 2.20. *Assume Setting 2.1, let $h \in \mathbb{N}$, $\theta \in (\mathscr{G}^h)^{-1}(\{0\})$ satisfy for all $j \in \{0, 1, \ldots, h\}$ that $\mathfrak{q}_j^\theta < \mathfrak{q}_{j+1}^\theta$ and $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$, and let $i \in \{0, 1, \ldots, h\}$ satisfy $0 < \mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta$ and $1/2 \leq \mathfrak{q}_i^\theta$. Then it holds for all $x \in [\mathfrak{q}_{m_{1,1}^\theta}^\theta, 1]$ that $\mathscr{N}_\infty^{h,\theta}(x) = 1$.*

Proof of Lemma 2.20. Note that the assumption that $\theta \in (\mathscr{G}^h)^{-1}(\{0\})$, the fact that $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0 < \mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta$, and Proposition 2.16 imply that

$$\int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} x(\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.150}$$

This, Lemma 2.7, and the assumption that $1/2 \leq \mathfrak{q}_i^\theta$ assure that for all $x \in [\mathfrak{q}_i^\theta, \mathfrak{q}_{i+1}^\theta]$ it holds that

$$\mathscr{N}_\infty^{h,\theta}(x) = 1. \tag{2.151}$$

Furthermore, observe that the assumption that $\theta \in (\mathscr{G}^h)^{-1}(\{0\})$, the assumption that $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$, and Proposition 2.16 imply that for all $j \in M_1^\theta$ it holds that

$$\int_{\mathfrak{q}_j^\theta}^{\mathfrak{q}_{j+1}^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.152}$$

Combining this with (2.151) and Lemma 2.11 demonstrates that for all $j \in M_1^\theta$ it holds that $\mathscr{N}_\infty^{h,\theta}(1) = \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_j^\theta) = 1$. This and the fact that $\mathscr{N}_\infty^{h,\theta}$ is piecewise affine linear ensure that for all $x \in [\mathfrak{q}_{m_{1,1}^\theta}^\theta, 1]$ it holds that $\mathscr{N}_\infty^{h,\theta}(x) = 1$. The proof of Lemma 2.20 is thus complete. ◻

Lemma 2.21. *Assume Setting 2.1 and let $h \in \mathbb{N}$, $\theta \in \mathbb{R}^{\mathfrak{d}_h}$ satisfy $M_0^\theta \neq \varnothing \neq M_1^\theta$, $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$, for all $x \in [0, \mathfrak{q}_{m_{0,2}^\theta}^\theta]$ that $\mathscr{N}_\infty^{h,\theta}(x) = 0$, and for all $x \in [\mathfrak{q}_{m_{1,1}^\theta}^\theta, 1]$ that $\mathscr{N}_\infty^{h,\theta}(x) = 1$. Then*

$$\mathscr{G}^h(\theta) \neq 0. \tag{2.153}$$

Proof of Lemma 2.21. Note that continuity of $\mathscr{N}_\infty^{h,\theta}$ implies for all $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ that $\mathfrak{q}^\theta_{m^\theta_{0,2}} < \mathfrak{q}^\theta_{m^\theta_{1,1}}$ and

$$\mathscr{N}_\infty^{h,\theta}(x) = \frac{x}{\mathfrak{q}^\theta_{m^\theta_{1,1}} - \mathfrak{q}^\theta_{m^\theta_{0,2}}} - \frac{\mathfrak{q}^\theta_{m^\theta_{0,2}}}{\mathfrak{q}^\theta_{m^\theta_{1,1}} - \mathfrak{q}^\theta_{m^\theta_{0,2}}}. \tag{2.154}$$

This shows that there exists $i \in \{1, 2, \dots, h\}$ which satisfies that

$$I_i^\theta \cap (\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}) = (\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}). \tag{2.155}$$

We prove (2.153) by contradiction. Assume that

$$\mathcal{G}^h(\theta) = 0. \tag{2.156}$$

Observe that (2.155), (2.156), and Proposition 2.16 imply that

$$\int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} x(\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.157}$$

Combining this and the fact that $\mathscr{N}_\infty^{h,\theta}(\mathfrak{q}^\theta_{m^\theta_{0,2}}) = 0$ with Corollary 2.9 demonstrates that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ it holds that

$$\mathscr{N}_\infty^{h,\theta}(x) = \frac{16x}{9(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 1)} - \frac{4(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 3)}{9(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 1)}. \tag{2.158}$$

This establishes that $\mathscr{N}_\infty^{h,\theta}(\mathfrak{q}^\theta_{m^\theta_{1,1}}) = 4/3$ which is a contradiction. The proof of Lemma 2.21 is thus complete. ☐

Lemma 2.22. *Assume Setting 2.1 and let $h \in \mathbb{N}$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$ and*

$$\int_0^1 x(\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.159}$$

Then

$$\mathcal{L}_\infty^h(\theta) \geq 1/36. \tag{2.160}$$

Proof of Lemma 2.22. Assume without loss of generality that $\mathfrak{q}_1^\theta \leq \ldots \leq \mathfrak{q}_h^\theta$. In the following we distinguish between the case $M_0^\theta = \varnothing = M_1^\theta$, the case $M_0^\theta \neq \varnothing$, and the case $M_0^\theta = \varnothing \neq M_1^\theta$. We first prove (2.160) in the case

$$M_0^\theta = \varnothing = M_1^\theta. \tag{2.161}$$

Note that (2.161) implies that there exist $a, b \in \mathbb{R}$ wich satisfy for all $x \in [0,1]$ that $\mathscr{N}_\infty^{h,\theta}(x) = ax + b$. This and Proposition 2.2 establish that

$$\mathcal{L}_\infty^h(\theta) \geq \frac{1}{16}. \tag{2.162}$$

This establishes (2.160) in the case $M_0^\theta = \varnothing = M_1^\theta$. Next we prove (2.160) in the case

$$M_0^\theta \neq \varnothing. \tag{2.163}$$

Observe that the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, the assumption that $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$, (2.159), (2.163), and Proposition 2.16 demonstrate that

$$\int_0^{\mathfrak{q}^\theta_{m^\theta_{0,1}}} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_0^{\mathfrak{q}^\theta_{m^\theta_{0,1}}} x(\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.164}$$

Combining this with Lemma 2.7 and Corollary 2.17 proves that for all $x \in [0, \mathfrak{q}^\theta_{m^\theta_{0,2}}]$ it holds that $\mathscr{N}^\theta_\infty(x) = 0$. This and Lemma 2.21 imply that $M^\theta_1 = \varnothing$. Hence, we obtain that there exists $a \in \mathbb{R}$ which satisfy for all $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, 1]$ that $\mathscr{N}^{h,\theta}_\infty(x) = a(x - \mathfrak{q}^\theta_{m^\theta_{0,2}})$. This and Proposition 2.3 establish that

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{36}. \tag{2.165}$$

This demonstrates (2.160) in the case $M^\theta_0 \neq \varnothing$. Next we prove (2.160) in the case

$$M^\theta_0 = \varnothing \neq M^\theta_1. \tag{2.166}$$

Note that the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, the assumption that $\prod_{k=1}^h \mathfrak{v}^\theta_k \neq 0$, (2.159), (2.166), and Proposition 2.16 show that

$$\int_{\mathfrak{q}^\theta_{m^\theta_{1,2}}}^1 (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))\, \mathrm{d}x = \int_{\mathfrak{q}^\theta_{m^\theta_{1,2}}}^1 x(\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))\, \mathrm{d}x = 0. \tag{2.167}$$

Combining this with Lemma 2.7 and Corollary 2.17 proves that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{1,1}}, 1]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = 1$. Therefore, we obtain that there exists $a \in \mathbb{R}$ which satisfies for all $x \in [0, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ that $\mathscr{N}^{h,\theta}_\infty(x) = a(x - \mathfrak{q}^\theta_{m^\theta_{1,1}}) + 1$. This and Proposition 2.4 establish that

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{36}. \tag{2.168}$$

This demonstrates (2.160) in the case $M^\theta_0 = \varnothing \neq M^\theta_1$. The proof of Lemma 2.22 is thus complete. □

Lemma 2.23. *Assume Setting 2.1, let $h \in \mathbb{N}$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $j \in \{0, 1, \dots, h\}$, $i \in \{1, 2, \dots, m^\theta_{0,2}\}$ that $\mathfrak{q}^\theta_j < \mathfrak{q}^\theta_{j+1}$, $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i-1} < 0 \neq \mathfrak{v}^\theta_i$, $M^\theta_0 \neq \varnothing$, and $\alpha^\theta \in \mathbb{R}\backslash\{0\}$. Then*

(i) for all $j \in \{1, 2, \dots, m^\theta_{0,2}\}$ it holds that $\mathfrak{q}^\theta_j = j\mathfrak{q}^\theta_1$,

(ii) for all $j \in \{0, 1, \dots, m^\theta_{0,2}\}$ it holds that

$$-\frac{\alpha^\theta \mathfrak{q}^\theta_1}{2} = (-1)^j \mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_j), \tag{2.169}$$

(iii) for all $j \in \{1, 2, \dots, m^\theta_{0,2}\}$, $x \in [\mathfrak{q}^\theta_{j-1}, \mathfrak{q}^\theta_j]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = (-1)^{j+1}\alpha^\theta x + (-1)^j \left(j - \frac{1}{2}\right)\alpha^\theta \mathfrak{q}^\theta_1, \tag{2.170}$$

and

(iv) it holds that

$$\int_0^{\mathfrak{q}^\theta_{m^\theta_{0,1}}} (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2\, \mathrm{d}x = \frac{m^\theta_{0,2}}{12}(\alpha^\theta)^2(\mathfrak{q}^\theta_1)^3. \tag{2.171}$$

Proof of Lemma 2.23. Observe that the assumption that $M^\theta_0 \neq \varnothing$ ensures that $\mathfrak{q}^\theta_1 \leq 1/2$. This, the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, the fact that $\mathfrak{w}^\theta_0 \mathfrak{w}^\theta_1 < 0 \neq \mathfrak{v}^\theta_1$, and Proposition 2.16 assure that

$$0 = \int_0^{\mathfrak{q}^\theta_1} (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))\, \mathrm{d}x = \int_0^{\mathfrak{q}^\theta_1} (\mathscr{N}^{h,\theta}_\infty(x))\, \mathrm{d}x = \frac{\alpha^\theta}{2}(\mathfrak{q}^\theta_1)^2 + \mathscr{N}^{h,\theta}_\infty(0)\mathfrak{q}^\theta_1. \tag{2.172}$$

This demonstrates that for all $x \in [0, \mathfrak{q}_1^\theta]$ it holds that

$$\mathscr{N}_\infty^{h,\theta}(x) = \alpha^\theta x - \frac{\alpha^\theta \mathfrak{q}_1^\theta}{2}. \tag{2.173}$$

Hence, we obtain that

$$\mathscr{N}_\infty^{h,\theta}(0) = -\frac{\alpha^\theta \mathfrak{q}_1^\theta}{2} = -\mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_1^\theta) \tag{2.174}$$

and

$$\int_0^{\mathfrak{q}_1^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x = \frac{1}{12}(\alpha^\theta)^2(\mathfrak{q}_1^\theta)^3. \tag{2.175}$$

Note that (2.173), (2.174), and (2.175) establish items (i), (ii), (iii), and (iv) in the case $m_{0,2}^\theta = 1$. Assume now that

$$m_{0,2}^\theta > 1. \tag{2.176}$$

Observe that the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, the assumption that for all $i \in \{1, 2, \ldots, m_{0,2}^\theta\}$ it holds that $\mathfrak{w}_i^\theta \mathfrak{w}_{i-1}^\theta < 0 \neq \mathfrak{v}_i^\theta$, (2.176), and Proposition 2.16 imply that for all $i \in \{1, 2, \ldots, m_{0,2}^\theta\}$ it holds that

$$\int_{\mathfrak{q}_{i-1}^\theta}^{\mathfrak{q}_i^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x)) \, \mathrm{d}x = 0. \tag{2.177}$$

This, Lemma 2.11, and (2.174) show that for all $i \in \{0, 1, \ldots, m_{0,2}^\theta\}$ it holds that

$$-\frac{\alpha^\theta \mathfrak{q}_1^\theta}{2} = (-1)^i \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_j^\theta). \tag{2.178}$$

This establishes that for all $i \in \{1, 2, \ldots, m_{0,2}^\theta\}$, $x \in [\mathfrak{q}_{i-1}^\theta, \mathfrak{q}_i^\theta]$ it holds that

$$\mathscr{N}_\infty^{h,\theta}(x) = \frac{(-1)^{i+1}\alpha^\theta \mathfrak{q}_1^\theta x}{\mathfrak{q}_i^\theta - \mathfrak{q}_{i-1}^\theta} + (-1)^i \left(\frac{\alpha^\theta \mathfrak{q}_1^\theta}{\mathfrak{q}_i^\theta - \mathfrak{q}_{i-1}^\theta} \mathfrak{q}_{i-1}^\theta + \frac{\alpha^\theta \mathfrak{q}_1^\theta}{2} \right). \tag{2.179}$$

Furthermore, note that the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, the assumption that for all $i \in \{1, 2, \ldots, m_{0,2}^\theta\}$ it holds that $\mathfrak{w}_i^\theta \mathfrak{w}_{i-1}^\theta < 0 \neq \mathfrak{v}_i^\theta$, (2.176), and Proposition 2.16 imply that for all $i \in \{1, 2, \ldots, m_{0,2}^\theta - 1\}$ it holds that

$$\int_{\mathfrak{q}_{i-1}^\theta}^{\mathfrak{q}_{i+1}^\theta} x(\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x)) \, \mathrm{d}x = 0. \tag{2.180}$$

Combining this and (2.177) with Lemma 2.12 demonstrates that for all $i \in \{1, 2, \ldots, m_{0,2}^\theta - 1\}$ it holds that $\mathfrak{q}_i^\theta - \mathfrak{q}_{i-1}^\theta = \mathfrak{q}_{i+1}^\theta - \mathfrak{q}_i^\theta$. This and the fact that $\mathfrak{q}_0^\theta = 0$ show that for all $i \in \{1, 2, \ldots, m_{0,2}^\theta\}$ it holds that

$$\mathfrak{q}_i^\theta = i \mathfrak{q}_1^\theta. \tag{2.181}$$

Combining this with (2.179) assures that for all $i \in \{1, 2, \ldots, m_{0,2}^\theta\}$, $x \in [\mathfrak{q}_{i-1}^\theta, \mathfrak{q}_i^\theta]$ it holds that

$$\mathscr{N}_\infty^{h,\theta}(x) = (-1)^{i+1}\alpha^\theta x + (-1)^i \left(i - \frac{1}{2} \right) \alpha^\theta \mathfrak{q}_1^\theta. \tag{2.182}$$

This ensures that

$$\begin{aligned} \int_0^{\mathfrak{q}_{m_{0,2}^\theta}^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x &= m_{0,2}^\theta \int_0^{\mathfrak{q}_1^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \\ &= \frac{m_{0,2}^\theta}{12} (\alpha^\theta)^2 (\mathfrak{q}_1^\theta)^3. \end{aligned} \tag{2.183}$$

This, (2.178), (2.181), and (2.183) prove items (i), (ii), (iii), and (iv) in the case $m_{0,2}^\theta > 1$. The proof of Lemma 2.23 is thus complete. ☐

Lemma 2.24. *Assume Setting 2.1, let $h \in \mathbb{N}$, $\theta \in (\mathscr{G}^h)^{-1}(\{0\})$ satisfy for all $j \in \{0, 1, \ldots, h\}$, $i \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \ldots, h\}$ that $\mathfrak{q}^\theta_j < \mathfrak{q}^\theta_{j+1}$, $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0 \neq \mathfrak{v}^\theta_i$, $M^\theta_1 \neq \varnothing$, and $\beta^\theta \in \mathbb{R}\backslash\{0\}$. Then*

(i) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \ldots, h\}$ it holds that $\mathfrak{q}^\theta_j = 1 - (h + 1 - j)(1 - \mathfrak{q}^\theta_h)$,

(ii) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \ldots, h + 1\}$ it holds that

$$\frac{\beta^\theta}{2}(1 - \mathfrak{q}^\theta_h) = (-1)^{h+1-j}(\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_j) - 1), \tag{2.184}$$

(iii) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \ldots, h\}$, $x \in [\mathfrak{q}^\theta_j, \mathfrak{q}^\theta_{j+1}]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = (-1)^{h-j}\beta^\theta x + 1 + (-1)^{h+1-j}\beta^\theta\left(1 - \left(h + \frac{1}{2} - j\right)(1 - \mathfrak{q}^\theta_h)\right), \tag{2.185}$$

and

(iv) it holds that

$$\int_{\mathfrak{q}^\theta_{m^\theta_{1,1}}}^1 (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, dx = \frac{1}{12}(h + 1 - m^\theta_{1,1})(\beta^\theta)^2(1 - \mathfrak{q}^\theta_h)^3. \tag{2.186}$$

Proof of Lemma 2.24. Observe that the assumption that $M^\theta_1 \neq \varnothing$ ensures that $\mathfrak{q}^\theta_h \geq 1/2$. This, the assumption that $\theta \in (\mathscr{G}^h)^{-1}(\{0\})$, the assumption that $\mathfrak{w}^\theta_h \mathfrak{w}^\theta_{h+1} < 0 \neq \mathfrak{v}^\theta_1$, and Proposition 2.16 assure that

$$\begin{aligned} 0 &= \int_{\mathfrak{q}^\theta_h}^1 (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x)) \, dx = \int_{\mathfrak{q}^\theta_h}^1 (\mathscr{N}^{h,\theta}_\infty(x) - 1) \, dx \\ &= \frac{\beta^\theta}{2}(1 - (\mathfrak{q}^\theta_h)^2) + (\mathscr{N}^{h,\theta}_\infty(1) - 1 - \beta^\theta)(1 - \mathfrak{q}^\theta_h). \end{aligned} \tag{2.187}$$

This demonstrates that for all $x \in [\mathfrak{q}^\theta_h, 1]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = \beta^\theta x + 1 - \frac{\beta^\theta}{2}(1 + \mathfrak{q}^\theta_h). \tag{2.188}$$

Therefore, we obtain that

$$\frac{\beta^\theta}{2}(1 - \mathfrak{q}^\theta_h) = \mathscr{N}^{h,\theta}_\infty(1) - 1 = -\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_h) + 1 \tag{2.189}$$

and

$$\int_{\mathfrak{q}^\theta_h}^1 (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, dx = \frac{1}{12}(\beta^\theta)^2(1 - \mathfrak{q}^\theta_h)^3. \tag{2.190}$$

Note that (2.188), (2.189), and (2.190) establish items (i), (ii), (iii), and (iv) in the case $m^\theta_{1,1} = h$. Assume now

$$m^\theta_{1,1} < h. \tag{2.191}$$

Observe that the assumption that $\theta \in (\mathscr{G}^h)^{-1}(\{0\})$, the assumption that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \ldots, h\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0 \neq \mathfrak{v}^\theta_i$, (2.191), and Proposition 2.16 imply that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \ldots, h\}$ it holds that

$$\int_{\mathfrak{q}^\theta_i}^{\mathfrak{q}^\theta_{i+1}} (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x)) \, dx = 0. \tag{2.192}$$

This, Lemma 2.11, and (2.189) show that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \dots, h+1\}$ it holds that

$$\frac{\beta^\theta}{2}(1-\mathfrak{q}^\theta_h) = \mathscr{N}^{h,\theta}_\infty(1) - 1 = (-1)^{h+1-i}(\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_i) - 1). \tag{2.193}$$

This establishes that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \dots, h\}$, $x \in [\mathfrak{q}^\theta_i, \mathfrak{q}^\theta_{i+1}]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = \frac{(-1)^{h-i}\beta^\theta(1-\mathfrak{q}^\theta_h)x}{\mathfrak{q}^\theta_{i+1}-\mathfrak{q}^\theta_i} + 1 + (-1)^{h+1-i}\left(\frac{\beta^\theta(1-\mathfrak{q}^\theta_h)}{\mathfrak{q}^\theta_{i+1}-\mathfrak{q}^\theta_i}\mathfrak{q}^\theta_i + \frac{\beta^\theta}{2}(1-\mathfrak{q}^\theta_h)\right). \tag{2.194}$$

Furthermore, note that the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, the assumption that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \dots, h\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0 \neq \mathfrak{v}^\theta_i$, (2.191), and Proposition 2.16 imply that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \dots, h-1\}$ it holds that

$$\int_{\mathfrak{q}^\theta_i}^{\mathfrak{q}^\theta_{i+2}} x(\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.195}$$

Combining this and (2.192) with Lemma 2.12 demonstrates that for all $i \in \{m^\theta_{1,1}+1, m^\theta_{1,1}+2, \dots, h\}$ it holds that $\mathfrak{q}^\theta_i - \mathfrak{q}^\theta_{i-1} = \mathfrak{q}^\theta_{i+1} - \mathfrak{q}^\theta_i$. This and the fact that $\mathfrak{q}^\theta_{h+1} = 1$ show that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \dots, h\}$ it holds that

$$\mathfrak{q}^\theta_i = 1 - (h+1-i)(1-\mathfrak{q}^\theta_h). \tag{2.196}$$

Combining this with (2.194) assures that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \dots, h\}$, $x \in [\mathfrak{q}^\theta_i, \mathfrak{q}^\theta_{i+1}]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = (-1)^{h-i}\beta^\theta x + 1 + (-1)^{h+1-i}\beta^\theta\left(1 - \left(h + \frac{1}{2} - i\right)(1-\mathfrak{q}^\theta_h)\right). \tag{2.197}$$

This ensures that

$$\begin{aligned} \int_{\mathfrak{q}^\theta_{m^\theta_{1,1}}}^1 (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2\,\mathrm{d}x &= (h+1-m^\theta_{1,1})\int_{\mathfrak{q}^\theta_h}^1 (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2\,\mathrm{d}x \\ &= \frac{1}{12}(h+1-m^\theta_{1,1})(\beta^\theta)^2(1-\mathfrak{q}^\theta_h)^3. \end{aligned} \tag{2.198}$$

This, (2.193), (2.196), and (2.197) prove items (i), (ii), (iii), and (iv) in the case $m^\theta_{1,1} < h$. The proof of Lemma 2.24 is thus complete. ☐

Lemma 2.25. *Assume Setting 2.1, let $h \in \mathbb{N}$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $j \in \{0, 1, \dots, h\}$, $i \in \{1, 2, \dots, m^\theta_{0,2}\}$, $x \in [\mathfrak{q}^\theta_{i-1}, \mathfrak{q}^\theta_i]$ that $\mathfrak{q}^\theta_j < \mathfrak{q}^\theta_{j+1}$, $\mathfrak{w}^\theta_i\mathfrak{w}^\theta_{i-1} < 0 \neq \prod_{k=1}^h \mathfrak{v}^\theta_k$, $M^\theta_0 \neq \varnothing$, and*

$$\mathscr{N}^{h,\theta}_\infty(x) = (-1)^{i+1}\alpha^\theta x + (-1)^i\left(i - \frac{1}{2}\right)\alpha^\theta \mathfrak{q}^\theta_1. \tag{2.199}$$

Then it holds for all $j \in \{1, 2, \dots, m^\theta_{0,2}\}\backslash\{m^\theta_{0,2}\}$ that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\alpha^\theta. \tag{2.200}$$

Proof of Lemma 2.25. Observe that the assumption that for all $i \in \{1, 2, \dots, m^\theta_{0,2}\}$ it holds that $\mathfrak{w}^\theta_i\mathfrak{w}^\theta_{i-1} < 0$ and the fact that $\mathfrak{w}^\theta_0 = -1$ implies that for all $j \in \{1, 2, \dots, m^\theta_{0,2}\}\backslash\{m^\theta_{0,2}\}$ with $\{n \in \mathbb{N} \colon j = 2n-1\} \neq \varnothing$ it holds that

$$I^\theta_j = (\mathfrak{q}^\theta_j, 1] \qquad \text{and} \qquad \mathfrak{w}^\theta_j > 0. \tag{2.201}$$

This and (2.199) show that for all $j \in \{1,2,\dots,m^\theta_{0,2}\}\backslash\{m^\theta_{0,2}\}$ with $\{n \in \mathbb{N} \colon j = 2n-1\} \neq \varnothing$ it holds that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\alpha^\theta. \tag{2.202}$$

Furthermore, note that the assumption that for all $i \in \{1,2,\dots,m^\theta_{0,2}\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i-1} < 0$ and the fact that $\mathfrak{w}^\theta_0 = -1$ ensures that for all $j \in \{1,2,\dots,m^\theta_{0,2}\}\backslash\{m^\theta_{0,2}\}$ with $\{n \in \mathbb{N} \colon j = 2n\} \neq \varnothing$ it holds that

$$I^\theta_j = [0, \mathfrak{q}^\theta_j) \qquad \text{and} \qquad \mathfrak{w}^\theta_j < 0. \tag{2.203}$$

This and (2.199) show that for all $j \in \{1,2,\dots,m^\theta_{0,2}\}\backslash\{m^\theta_{0,2}\}$ with $\{n \in \mathbb{N} \colon j = 2n\} \neq \varnothing$ it holds that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\alpha^\theta. \tag{2.204}$$

Combining this and (2.202) demonstrates that for all $j \in \{1,2,\dots,m^\theta_{0,2}\}\backslash\{m^\theta_{0,2}\}$ it holds that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\alpha^\theta. \tag{2.205}$$

The proof of Lemma 2.25 is thus complete. ☐

Lemma 2.26. *Assume Setting 2.1, let $h \in \mathbb{N}$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $j \in \{0,1,\dots,h\}$, $i \in \{m^\theta_{1,1}, m^\theta_{1,1}+1,\dots,h\}$, $x \in [\mathfrak{q}^\theta_i, \mathfrak{q}^\theta_{i+1}]$ that $\mathfrak{q}^\theta_j < \mathfrak{q}^\theta_{j+1}$, $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0 \neq \prod_{k=1}^h \mathfrak{v}^\theta_k$, $M^\theta_1 \neq \varnothing$, and*

$$\mathscr{N}^{h,\theta}_\infty(x) = (-1)^{h-i}\beta^\theta x + 1 + (-1)^{h+1-i}\beta^\theta\left(1 - \left(h + \frac{1}{2} - i\right)(1 - \mathfrak{q}^\theta_h)\right). \tag{2.206}$$

Then it holds for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1,\dots,h\}\backslash\{m^\theta_{1,1}\}$ that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\beta^\theta. \tag{2.207}$$

Proof of Lemma 2.26. Observe that the assumption that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1}+1,\dots,h\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0$ and the fact that $\mathfrak{w}^\theta_{h+1} = 1$ implies that for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1,\dots,h\}\backslash\{m^\theta_{1,1}\}$ with $\{n \in \mathbb{N} \colon j = h-2n+1\} \neq \varnothing$ it holds that

$$I^\theta_j = (\mathfrak{q}^\theta_j, 1] \qquad \text{and} \qquad \mathfrak{w}^\theta_j > 0. \tag{2.208}$$

This and (2.206) show that for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1,\dots,h\}\backslash\{m^\theta_{1,1}\}$ with $\{n \in \mathbb{N} \colon j = h-2n+1\} \neq \varnothing$ it holds that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\beta^\theta. \tag{2.209}$$

Furthermore, note that the assumption that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1}+1,\dots,h\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0$ and the fact that $\mathfrak{w}^\theta_{h+1} = -1$ ensures that for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1,\dots,h\}\backslash\{m^\theta_{1,1}\}$ with $\{n \in \mathbb{N} \colon j = h-2n+2\} \neq \varnothing$ it holds that

$$I^\theta_j = [0, \mathfrak{q}^\theta_j) \qquad \text{and} \qquad \mathfrak{w}^\theta_j < 0. \tag{2.210}$$

This and (2.206) establish that for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1,\dots,h\}\backslash\{m^\theta_{1,1}\}$ with $\{n \in \mathbb{N} \colon j = h-2n+2\} \neq \varnothing$ it holds that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\beta^\theta. \tag{2.211}$$

Combining this and (2.209) demonstrates that for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1,\dots,h\}\backslash\{m^\theta_{1,1}\}$ it holds that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\beta^\theta. \tag{2.212}$$

The proof of Lemma 2.26 is thus complete. ☐

Lemma 2.27. *Assume Setting 2.1, let $h \in \mathbb{N} \cap (1, \infty)$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $j \in \{0, 1, \ldots, h\}$, $x \in [0, \mathfrak{q}^\theta_{m^\theta_{0,2}}]$ that $\mathfrak{q}^\theta_j < \mathfrak{q}^\theta_{j+1}$, $\mathcal{N}^{h,\theta}_\infty(x) = 0$, $M^\theta_0 \neq \varnothing \neq M^\theta_1$, and $0 \neq \prod_{k=1}^h \mathfrak{v}^\theta_k$. Then*

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.213}$$

Proof of Lemma 2.27. Observe that the assumption that $M^\theta_0 \neq \varnothing \neq M^\theta_1$, the assumption that for all $x \in [0, \mathfrak{q}^\theta_{m^\theta_{0,2}}]$ it holds that $\mathcal{N}^{h,\theta}_\infty(x) = 0$, Corollary 2.17, Lemma 2.20, and Lemma 2.21 assure that for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \ldots, h\}$ it holds that $\mathfrak{w}^\theta_j \mathfrak{w}^\theta_{j+1} < 0$ and $\beta^\theta \neq 0$. This and Lemma 2.24 demonstrate that

(i) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \ldots, h\}$ it holds that $\mathfrak{q}^\theta_j = 1 - (h+1-j)(1-\mathfrak{q}^\theta_h)$,

(ii) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \ldots, h+1\}$ it holds that

$$\frac{\beta^\theta}{2}(1-\mathfrak{q}^\theta_h) = (-1)^{h+1-j}(\mathcal{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_j) - 1), \tag{2.214}$$

(iii) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \ldots, h\}$, $x \in [\mathfrak{q}^\theta_j, \mathfrak{q}^\theta_{j+1}]$ it holds that

$$\mathcal{N}^{h,\theta}_\infty(x) = (-1)^{h-j}\beta^\theta x + 1 + (-1)^{h+1-j}\beta^\theta\left(1 - \left(h + \frac{1}{2} - j\right)(1-\mathfrak{q}^\theta_h)\right), \tag{2.215}$$

and

(iv) it holds that

$$\int_{\mathfrak{q}^\theta_{m^\theta_{1,1}}}^1 (\mathcal{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, dx = \frac{1}{12}(h+1-m^\theta_{1,1})(\beta^\theta)^2(1-\mathfrak{q}^\theta_h)^3. \tag{2.216}$$

Note that (2.214) assures that in the case $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$ it holds that

$$1 = |\mathcal{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}}) - 1| = |\mathcal{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{1,1}}) - 1| = |1/2\beta^\theta(1-\mathfrak{q}^\theta_h)|. \tag{2.217}$$

Furthermore, observe that in the case $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$ it holds that $(h+1-m^\theta_{1,1})(1-\mathfrak{q}^\theta_h) = 1/2$. Combining this and (2.217) with (2.216) implies that in the case $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$ it holds that

$$\mathcal{L}^h_\infty(\theta) = \frac{1}{6}. \tag{2.218}$$

This establishes (2.213) in the case $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$. Assume now that $\mathfrak{q}^\theta_{m^\theta_{0,2}} < 1/2 < \mathfrak{q}^\theta_{m^\theta_{1,2}}$. In the following we distinguish between the case $\mathfrak{w}^\theta_{m^\theta_{0,2}} < 0 < \mathfrak{w}^\theta_{m^\theta_{1,1}}$ and the case $\max\{\mathfrak{w}^\theta_{m^\theta_{0,2}}, \mathfrak{w}^\theta_{m^\theta_{1,1}}, -\mathfrak{w}^\theta_{m^\theta_{1,1}}\} > 0$. We first prove (2.213) in the case

$$\max\{\mathfrak{w}^\theta_{m^\theta_{0,2}}, \mathfrak{w}^\theta_{m^\theta_{1,1}}, -\mathfrak{w}^\theta_{m^\theta_{1,1}}\} > 0. \tag{2.219}$$

Note that (2.219), the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, and Proposition 2.16 assure that

$$\int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} x(\mathcal{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x)) \, dx = \int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} (\mathcal{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x)) \, dx = 0. \tag{2.220}$$

Combining this with Corollary 2.9 demonstrates that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = \frac{16x}{9(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 1)} + \frac{4(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 3)}{9(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 1)}. \tag{2.221}$$

Hence, we obtain that $\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{1,1}}) = 4/3$ and $\mathfrak{q}^\theta_{m^\theta_{0,2}} = 3/4 - \mathfrak{q}^\theta_{m^\theta_{1,1}}/2$. Combining this with (2.214) and (2.216) shows that $|\beta^\theta/2(1 - \mathfrak{q}^\theta_h)| = |\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{1,1}}) - 1| = 1/3$ and

$$\begin{aligned} \mathcal{L}^h_\infty(\theta) &= \int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} \left(\frac{16x}{9(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 1)} + \frac{4(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 3)}{9(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 1)} - \mathbb{1}_{(1/2,\infty)}(x) \right)^2 \mathrm{d}x \\ &\quad + \frac{h + 1 - m^\theta_{1,1}}{12} (\beta^\theta)^2 (1 - \mathfrak{q}^\theta_h)^3 = \frac{7}{36}(2\mathfrak{q}^\theta_{m^\theta_{1,1}} - 1) + \frac{1}{27}(1 - \mathfrak{q}^\theta_{m^\theta_{1,1}}) \\ &= \frac{19\mathfrak{q}^\theta_{m^\theta_{1,1}}}{54} - \frac{17}{108} \geq \frac{19}{108} - \frac{17}{108} = \frac{1}{54}. \end{aligned} \tag{2.222}$$

This implies (2.213) in the case $\max\{\mathfrak{w}^\theta_{m^\theta_{0,2}} \mathfrak{w}^\theta_{m^\theta_{1,1}}, -\mathfrak{w}^\theta_{m^\theta_{1,1}}\} > 0$. Next we demonstrate (2.213) in the case

$$\mathfrak{w}^\theta_{m^\theta_{0,2}} < 0 < \mathfrak{w}^\theta_{m^\theta_{1,1}}. \tag{2.223}$$

Observe that (2.223) and the fact that for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \ldots, h\}$ it holds that $\mathfrak{w}^\theta_j \mathfrak{w}^\theta_{j+1} < 0$ prove that there exists $k^* \in \mathbb{N}$ which satisfies that $h = m^\theta_{1,1} + 2k^* - 1$. Combining this with the fact that for all $x \in [0, \mathfrak{q}^\theta_{m^\theta_{0,2}}]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = 0$ and the fact that for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \ldots, h\}$ it holds that $\mathfrak{w}^\theta_j \mathfrak{w}^\theta_{j+1} < 0$ establishes that

$$\{k \in \{1, 2, \ldots, h\} \colon I^\theta_k \cup (\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}) \neq \varnothing\} = \cup_{k=1}^{k^*} \{m^\theta_{1,1} + 2k - 1\}. \tag{2.224}$$

Note that Lemma 2.26 ensures that for all $j \in \cup_{k=1}^{k^*} \{m^\theta_{1,1} + 2k - 1\}$ it holds that $\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\beta^\theta$. Combining this with (2.224) and the fact that $\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}}) = 0$ assures that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = -2k^* \beta^\theta x + 2k^* \beta^\theta \mathfrak{q}^\theta_{m^\theta_{0,2}}. \tag{2.225}$$

This and (2.214) demonstrate that

$$-2k^* \beta^\theta \mathfrak{q}^\theta_{m^\theta_{1,1}} + 2k^* \beta^\theta \mathfrak{q}^\theta_{m^\theta_{0,2}} = 1 + \frac{\beta^\theta}{2}(1 - \mathfrak{q}^\theta_h). \tag{2.226}$$

Therefore, we obtain that

$$|\beta^\theta| = \frac{1}{|\frac{1}{2}(1 - \mathfrak{q}^\theta_h) + 2k^*(\mathfrak{q}^\theta_{m^\theta_{1,1}} - \mathfrak{q}^\theta_{m^\theta_{0,2}})|} \geq \frac{1}{1 + 2k^*}. \tag{2.227}$$

Combining this, Proposition 2.3, (2.216), and (2.225) shows that

$$\begin{aligned} \mathcal{L}^h_\infty(\theta) &\geq \begin{cases} \int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} (-2k^* \beta^\theta x + 2k^* \beta^\theta \mathfrak{q}^\theta_{m^\theta_{0,2}} - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x & : \mathfrak{q}^\theta_{m^\theta_{1,1}} \geq \frac{3}{4} \\ \int_{\mathfrak{q}^\theta_{m^\theta_{1,1}}}^1 (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x = \frac{2k^*}{12} (\beta^\theta)^2 (1 - \mathfrak{q}^\theta_h)^3 & : \mathfrak{q}^\theta_{m^\theta_{1,1}} \leq \frac{3}{4} \end{cases} \\ &\geq \begin{cases} \frac{1}{36} & : \mathfrak{q}^\theta_{m^\theta_{1,1}} \geq \frac{3}{4} \\ \frac{k^*}{6} \frac{1}{(1+2k^*)^2} \frac{1}{64} \geq \frac{1}{384(1+h)^2} & : \mathfrak{q}^\theta_{m^\theta_{1,1}} \leq \frac{3}{4}. \end{cases} \end{aligned} \tag{2.228}$$

This, (2.218), and (2.222) assure that

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.229}$$

This shows (2.213) in the case $\mathfrak{w}^\theta_{m^\theta_{0,2}} < 0 < \mathfrak{w}^\theta_{m^\theta_{1,1}}$. The proof of Lemma 2.27 is thus complete. □

Lemma 2.28. *Assume Setting 2.1, let* $h \in \mathbb{N} \cap (1,\infty)$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ *satisfy for all* $j \in \{0,1,\ldots,h\}$, $x \in [\mathfrak{q}^\theta_{m^\theta_{1,2}}, 1]$ *that* $\mathfrak{q}^\theta_j < \mathfrak{q}^\theta_{j+1}$, $\mathcal{N}^{h,\theta}_\infty(x) = 1$, $M^\theta_0 \neq \varnothing \neq M^\theta_1$, *and* $0 \neq \prod_{k=1}^h \mathfrak{v}^\theta_k$. *Then*

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.230}$$

Proof of Lemma 2.28. Observe that the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, the assumption that $M^\theta_0 \neq \varnothing \neq M^\theta_1$, the assumption that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{1,1}}, 1]$ it holds that $\mathcal{N}^\theta_\infty(x) = 1$, Corollary 2.18, Lemma 2.19, and Lemma 2.21 assure that for all $j \in \{1,2,\ldots,m^\theta_{0,2}\}$ it holds that $\mathfrak{w}^\theta_j \mathfrak{w}^\theta_{j-1} < 0$ and $\alpha^\theta \neq 0$. This and Lemma 2.23 demonstrate that

(i) for all $j \in \{1,2,\ldots,m^\theta_{0,2}\}$ it holds that $\mathfrak{q}^\theta_j = j\mathfrak{q}^\theta_1$,

(ii) for all $j \in \{0,1,\ldots,m^\theta_{0,2}\}$ it holds that

$$-\frac{\alpha^\theta \mathfrak{q}^\theta_1}{2} = (-1)^j \mathcal{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_j), \tag{2.231}$$

(iii) for all $j \in \{1,2,\ldots,m^\theta_{0,2}\}$, $x \in [\mathfrak{q}^\theta_{j-1}, \mathfrak{q}^\theta_j]$ it holds that

$$\mathcal{N}^{h,\theta}_\infty(x) = (-1)^{j+1}\alpha^\theta x + (-1)^j\left(j - \frac{1}{2}\right)\alpha^\theta \mathfrak{q}^\theta_1, \tag{2.232}$$

and

(iv) it holds that

$$\int_0^{\mathfrak{q}^\theta_{m^\theta_{0,1}}} (\mathcal{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, dx = \frac{m^\theta_{0,2}}{12}(\alpha^\theta)^2(\mathfrak{q}^\theta_1)^3. \tag{2.233}$$

Note that (2.231) assures that in the case $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$ it holds that

$$1 = |\mathcal{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{1,1}})| = |\mathcal{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}})| = \left|\frac{\alpha^\theta \mathfrak{q}^\theta_1}{2}\right|. \tag{2.234}$$

Furthermore, observe that in the case $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$ it holds that $m^\theta_{0,2}\mathfrak{q}^\theta_1 = 1/2$. Combining this and (2.234) with (2.233) implies that in the case $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$ it holds that

$$\mathcal{L}^h_\infty(\theta) = \frac{1}{6}. \tag{2.235}$$

This demonstrates (2.230) in the case $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$. Assume now that $\mathfrak{q}^\theta_{m^\theta_{0,2}} < 1/2 < \mathfrak{q}^\theta_{m^\theta_{1,2}}$. In the following we distinguish between the case $\mathfrak{w}^\theta_{m^\theta_{0,2}} < 0 < \mathfrak{w}^\theta_{m^\theta_{1,1}}$ and the case $\max\{\mathfrak{w}^\theta_{m^\theta_{0,2}} \mathfrak{w}^\theta_{m^\theta_{1,1}}, \mathfrak{w}^\theta_{m^\theta_{0,2}}\} > 0$. We first demonstrate (2.230) in the case

$$\max\{\mathfrak{w}^\theta_{m^\theta_{0,2}} \mathfrak{w}^\theta_{m^\theta_{1,1}}, \mathfrak{w}^\theta_{m^\theta_{0,2}}\} > 0. \tag{2.236}$$

Note that (2.236), the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, and Proposition 2.16 assure that

$$\int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} x(\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.237}$$

Combining this with Corollary 2.10 demonstrates that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = -\frac{16x}{9(2\mathfrak{q}^\theta_{m^\theta_{0,2}} - 1)} + \frac{10\mathfrak{q}^\theta_{m^\theta_{0,2}} + 3}{9(2\mathfrak{q}^\theta_{m^\theta_{0,2}} - 1)}. \tag{2.238}$$

Hence, we obtain that $\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}}) = -1/3$ and $\mathfrak{q}^\theta_{m^\theta_{1,1}} = 3/4 - 1/2\mathfrak{q}^\theta_{m^\theta_{0,2}}$. Combining this with (2.231) and (2.233) shows that $|1/2\alpha^\theta \mathfrak{q}^\theta_1| = |\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}})| = 1/3$ and

$$\begin{aligned}\mathcal{L}^h_\infty(\theta) &= \int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} \left(-\frac{16x}{9(2\mathfrak{q}^\theta_{m^\theta_{0,2}} - 1)} + \frac{10\mathfrak{q}^\theta_{m^\theta_{0,2}} + 3}{9(2\mathfrak{q}^\theta_{m^\theta_{0,2}} - 1)} - \mathbb{1}_{(1/2,\infty)}(x) \right)^2 \mathrm{d}x \\ &\quad + \frac{m^\theta_{0,2}}{12}(\alpha^\theta)^2(\mathfrak{q}^\theta_1)^3 = \frac{1}{18}(1 - 2\mathfrak{q}^\theta_{m^\theta_{0,2}}) + \frac{1}{27}\mathfrak{q}^\theta_{m^\theta_{0,2}} \\ &= -\frac{2\mathfrak{q}^\theta_{m^\theta_{0,2}}}{27} + \frac{1}{18} \geq -\frac{1}{27} + \frac{1}{18} = \frac{1}{54}.\end{aligned} \tag{2.239}$$

This implies (2.230) in the case $\max\{\mathfrak{w}^\theta_{m^\theta_{0,2}}, \mathfrak{w}^\theta_{m^\theta_{1,1}}, \mathfrak{w}^\theta_{m^\theta_{0,2}}\} > 0$. Next we demonstrate (2.230) in the case

$$\mathfrak{w}^\theta_{m^\theta_{0,2}} < 0 < \mathfrak{w}^\theta_{m^\theta_{1,1}}. \tag{2.240}$$

Observe that (2.240) and the fact that for all $i \in \{0, 1, \ldots, m^\theta_{0,2} - 1\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0$ prove that there exists $k^* \in \mathbb{N}$ which satisfies that $m^\theta_{0,2} = 2k^*$. Combining this with the fact that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{1,1}}, 1]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = 1$ and the fact that for all $j \in \{0, 1, \ldots, m^\theta_{0,2} - 1\}$ it holds that $\mathfrak{w}^\theta_j \mathfrak{w}^\theta_{j+1} < 0$ establishes that

$$\{k \in \{1, 2, \ldots, h\} \colon I^\theta_k \cap (\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}) \neq \varnothing\} = \cup_{k=1}^{k^*}\{2k - 1\}. \tag{2.241}$$

Note that Lemma 2.26 ensures that for all $j \in \cup_{k=1}^{k^*}\{2k-1\}$ it holds that $\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = -2\alpha^\theta$. Combining this with (2.241) and the fact that $\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{1,1}}) = 1$ assures that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = -2k^*\alpha^\theta x + 2k^*\alpha^\theta \mathfrak{q}^\theta_{m^\theta_{1,1}} + 1. \tag{2.242}$$

This and (2.231) demonstrate that

$$-2k^*\alpha^\theta \mathfrak{q}^\theta_{m^\theta_{0,2}} + 2k^*\alpha^\theta \mathfrak{q}^\theta_{m^\theta_{1,1}} + 1 = -\frac{\alpha^\theta \mathfrak{q}^\theta_1}{2}. \tag{2.243}$$

Therefore, we obtain that

$$|\alpha^\theta| = \frac{1}{|\frac{\mathfrak{q}^\theta_1}{2} + 2k^*(\mathfrak{q}^\theta_{m^\theta_{1,1}} - \mathfrak{q}^\theta_{m^\theta_{0,2}})|} \geq \frac{1}{1 + 2k^*}. \tag{2.244}$$

Combining this, Proposition 2.4, (2.233), and (2.242) shows that

$$\begin{aligned}\mathcal{L}^h_\infty(\theta) &\geq \begin{cases}\int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}}(-2k^*\alpha^\theta x + 2k^*\alpha^\theta \mathfrak{q}^\theta_{m^\theta_{1,1}} + 1 - \mathbb{1}_{(1/2,\infty)}(x))^2\,\mathrm{d}x & : \mathfrak{q}^\theta_{m^\theta_{0,2}} \leq \frac{1}{4}\\ \int_{\mathfrak{q}^\theta_{m^\theta_{1,1}}}^1(\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2\,\mathrm{d}x = \frac{2k^*}{12}(\alpha^\theta)^2(\mathfrak{q}^\theta_1)^3 & : \mathfrak{q}^\theta_{m^\theta_{0,2}} \geq \frac{1}{4}\end{cases}\\ &\geq \begin{cases}\frac{1}{36} & : \mathfrak{q}^\theta_{m^\theta_{0,2}} \leq \frac{1}{4}\\ \frac{k^*}{6}\frac{1}{(1+2k^*)^2}\frac{1}{64} \geq \frac{1}{384(1+h)^2} & : \mathfrak{q}^\theta_{m^\theta_{0,2}} \geq \frac{1}{4}.\end{cases}\end{aligned} \tag{2.245}$$

This, (2.235), and (2.239) assure that

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.246}$$

This shows (2.230) in the case $\mathfrak{w}^\theta_{m^\theta_{0,2}} < 0 < \mathfrak{w}^\theta_{m^\theta_{1,1}}$. The proof of Lemma 2.28 is thus complete. □

Corollary 2.29. *Assume Setting 2.1 and let $h \in \mathbb{N} \cap (1,\infty)$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $i \in \{0,\dots,h\}$ that $\mathfrak{w}^\theta_i\mathfrak{w}^\theta_{i+1} < 0 \neq \alpha^\theta\beta^\theta$, $\prod_{k=1}^h \mathfrak{v}^\theta_k \neq 0 < \mathfrak{q}^\theta_1 < \mathfrak{q}^\theta_2 < \dots < \mathfrak{q}^\theta_h < 1$, and $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$. Then $\mathcal{L}^h_\infty(\theta) \geq 1/12$.*

Proof of Corollary 2.29. Observe that Lemma 2.23, Lemma 2.24, and the assumption that $\mathfrak{q}^\theta_{m^\theta_{0,2}} = \mathfrak{q}^\theta_{m^\theta_{1,1}} = 1/2$ assure that $|1+1/2(-1)^{h+1-m^\theta_{1,1}}\beta^\theta(1-\mathfrak{q}^\theta_h)| = |\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{1,1}})| = |\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}})| = |\alpha^\theta\mathfrak{q}^\theta_1/2|$ and

$$\int_0^1(\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2\,\mathrm{d}x = \frac{1}{24}(\alpha^\theta)^2(\mathfrak{q}^\theta_1)^2 + \frac{1}{24}(\beta^\theta)^2(1-\mathfrak{q}^\theta_h)^2. \tag{2.247}$$

Hence, we obtain that

$$\begin{aligned}\mathcal{L}^h_\infty(\theta) &= \frac{1}{6}\left(1 + \frac{1}{2}(-1)^{h+1-m^\theta_{1,1}}\beta^\theta(1-\mathfrak{q}^\theta_h)\right)^2 + \frac{1}{24}(\beta^\theta)^2(1-\mathfrak{q}^\theta_h)^2\\ &= \frac{1}{12}(\beta^\theta(1-\mathfrak{q}^\theta_h))^2 + \frac{1}{6}(-1)^{h+1-m^\theta_{1,1}}\beta^\theta(1-\mathfrak{q}^\theta_h) + \frac{1}{6}\\ &\geq \frac{1}{12}(-1)^2 + \frac{1}{6}(-1) + \frac{1}{6} = \frac{1}{12}.\end{aligned} \tag{2.248}$$

This demonstrates that $\mathcal{L}^h_\infty(\theta) \geq 1/12$. The proof of Corollary 2.29 is thus complete. □

Lemma 2.30. *Assume Setting 2.1 and let $h \in \mathbb{N}\cap(1,\infty)$, $\theta \in \mathbb{R}^{\mathfrak{d}_h}$ satisfy for all $i \in \{0,1,\dots,h\} \backslash\{m^\theta_{0,2}\}$ that $\mathfrak{w}^\theta_i\mathfrak{w}^\theta_{i+1} < 0$, $\prod_{k=1}^h \mathfrak{v}^\theta_k \neq 0 < \mathfrak{q}^\theta_1 < \mathfrak{q}^\theta_2 < \dots < \mathfrak{q}^\theta_h < 1$, $\alpha^\theta\beta^\theta \neq 0 < m^\theta_{0,2} < m^\theta_{1,1} < h+1$, and $0 < \mathfrak{w}^\theta_{m^\theta_{0,2}}\mathfrak{w}^\theta_{m^\theta_{1,1}}$. Then*

$$\mathcal{G}^h(\theta) \neq 0. \tag{2.249}$$

Proof of Lemma 2.30. We prove (2.249) by contradiction. Assume that

$$\mathcal{G}^h(\theta) = 0. \tag{2.250}$$

Note that (2.250), Lemma 2.23, and Lemma 2.24 ensure that

(i) for all $j \in \{1,2\dots,m^\theta_{0,2}\}$ it holds that

$$\mathfrak{q}^\theta_j = j\mathfrak{q}^\theta_1, \tag{2.251}$$

(ii) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \ldots, h\}$ it holds that

$$\mathfrak{q}^\theta_j = 1 - (h+1-j)(1-\mathfrak{q}^\theta_h), \tag{2.252}$$

and

(iii) for all $j \in \{0,1,\ldots,h\}\backslash\{m^\theta_{0,2}\}$, $x \in [\mathfrak{q}^\theta_j, \mathfrak{q}^\theta_{j+1}]$ it holds that

$$\begin{aligned} &\mathscr{N}^{h,\theta}_\infty(x) \\ &= \begin{cases} (-1)^j \alpha^\theta x + (-1)^{j+1}(j+{}^1\!/\!_2)\alpha^\theta \mathfrak{q}^\theta_1 & : x \le \mathfrak{q}^\theta_{m^\theta_{0,2}} \\ (-1)^{h-j}\beta^\theta x + 1 + (-1)^{h+1-j}\beta^\theta(1-(h+{}^1\!/\!_2-j)(1-\mathfrak{q}^\theta_h)) & : x \ge \mathfrak{q}^\theta_{m^\theta_{1,1}}. \end{cases} \end{aligned} \tag{2.253}$$

This, Lemma 2.25, and Lemma 2.26 assure that for all $j \in \{1,2,\ldots,h\}\backslash\{m^\theta_{0,2}, m^\theta_{1,1}\}$ it holds that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = \begin{cases} -2\alpha^\theta & : j < m^\theta_{0,2} \\ -2\beta^\theta & : j > m^\theta_{1,1}. \end{cases} \tag{2.254}$$

In the following we distinguish between the case $\max\{\mathfrak{w}^\theta_{m^\theta_{0,2}}, \mathfrak{w}^\theta_{m^\theta_{1,1}}\} < 0$ and the case $\min\{\mathfrak{w}^\theta_{m^\theta_{0,2}}, \mathfrak{w}^\theta_{m^\theta_{1,1}}\} > 0$. We first establish the contradiction in the case

$$\max\{\mathfrak{w}^\theta_{m^\theta_{0,2}}, \mathfrak{w}^\theta_{m^\theta_{1,1}}\} < 0. \tag{2.255}$$

Observe that (2.255) and the assumption that for all $i \in \{0,1,\ldots,h\}\backslash\{m^\theta_{0,2}\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0$ prove that there exist $k_1, k_2 \in \mathbb{N}$ which satisfy that $m^\theta_{0,2} = 2k_1$, $h - m^\theta_{1,1} = 2(k_2-1)$, and

$$\{k \in \{1,2,\ldots,h\} \colon I^\theta_k \cap (\mathfrak{q}^\theta_{m^\theta_{1,1}}, \mathfrak{q}^\theta_{m^\theta_{1,1}+1}) \neq \varnothing\} = (\cup_{k=0}^{1/2(h-1)}\{1+2k\})\backslash\{m^\theta_{1,1}\}. \tag{2.256}$$

This, (2.254), and the fact that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{1,1}}, \mathfrak{q}^\theta_{m^\theta_{1,1}+1}]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = \beta^\theta x + 1 - \beta^\theta\big(1-(h+{}^1\!/\!_2-m^\theta_{1,1})(1-\mathfrak{q}^\theta_h)\big)$ show that

$$\beta^\theta = -m^\theta_{0,2}\alpha^\theta - (h-m^\theta_{1,1})\beta^\theta. \tag{2.257}$$

Combining this, the fact that $\max\{\mathfrak{w}^\theta_{m^\theta_{0,2}}, \mathfrak{w}^\theta_{m^\theta_{1,1}}\} < 0 < \mathfrak{q}^\theta_{m^\theta_{0,2}} < \mathfrak{q}^\theta_{m^\theta_{1,1}} < 1$, (2.251), (2.252), (2.253), the assumption that $\theta \in \mathcal{G}^{-1}(\{0\})$, and Proposition 2.16 assures that

$$\begin{aligned} 0 &= \int_{\mathfrak{q}^\theta_{m^\theta_{0,2}-1}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}+1}} x(\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x \\ &= \int_{\mathfrak{q}^\theta_{m^\theta_{0,2}-1}}^{\mathfrak{q}^\theta_{m^\theta_{0,2}}} x(\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x + \int_{\mathfrak{q}^\theta_{m^\theta_{1,1}}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}+1}} x(\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x \\ &= -\frac{\alpha^\theta}{12}(\mathfrak{q}^\theta_1)^3 + \frac{\beta^\theta}{12}(1-\mathfrak{q}^\theta_h)^3 = -\frac{\alpha^\theta}{12}\bigg((\mathfrak{q}^\theta_1)^3 + \frac{m^\theta_{0,2}}{h-m^\theta_{1,1}+1}(1-\mathfrak{q}^\theta_h)^3\bigg). \end{aligned} \tag{2.258}$$

This implies that

$$\mathfrak{q}^\theta_1 = -\sqrt[3]{\frac{m^\theta_{0,2}}{h-m^\theta_{1,1}+1}}(1-\mathfrak{q}^\theta_h) < 0 \tag{2.259}$$

which is a contradiction. In the next step we establish the contradiction in the case

$$\min\{\mathfrak{w}^{\theta}_{m^{\theta}_{0,2}}, \mathfrak{w}^{\theta}_{m^{\theta}_{1,1}}\} > 0. \tag{2.260}$$

Note that (2.260) and the assumption that for all $i \in \{0, 1, \ldots, h\}\backslash\{m^{\theta}_{0,2}\}$ it holds that $\mathfrak{w}^{\theta}_i \mathfrak{w}^{\theta}_{i+1} < 0$ prove that there exist $k_1, k_2 \in \mathbb{N}$ which satisfy that $m^{\theta}_{0,2} = 2k_1 - 1$, $h - m^{\theta}_{1,1} = 2k_2 - 1$, and

$$\{k \in \{1, 2, \ldots, h\} \colon I^{\theta}_k \cap (\mathfrak{q}^{\theta}_{m^{\theta}_{0,2}-1}, \mathfrak{q}^{\theta}_{m^{\theta}_{0,2}}) \neq \varnothing\} = (\cup_{k=0}^{1/2(h-1)}\{1 + 2k\})\backslash\{m^{\theta}_{0,2}\}. \tag{2.261}$$

This, (2.254), and the fact that for all $x \in [\mathfrak{q}^{\theta}_{m^{\theta}_{0,2}-1}, \mathfrak{q}^{\theta}_{m^{\theta}_{0,2}}]$ it holds that $\mathscr{N}^{h,\theta}_{\infty}(x) = \alpha^{\theta}x - (m^{\theta}_{0,2} - 1/2)\alpha^{\theta}\mathfrak{q}^{\theta}_1$ show that

$$\alpha^{\theta} = -(m^{\theta}_{0,2} - 1)\alpha^{\theta} - (h - m^{\theta}_{1,1} + 1)\beta^{\theta}. \tag{2.262}$$

Combining this, the fact that $-\min\{\mathfrak{w}^{\theta}_{m^{\theta}_{0,2}}, \mathfrak{w}^{\theta}_{m^{\theta}_{1,1}}\} < 0 < \mathfrak{q}^{\theta}_{m^{\theta}_{0,2}} < \mathfrak{q}^{\theta}_{m^{\theta}_{1,1}} < 1$, (2.251), (2.252), (2.253), the assumption that $\theta \in \mathcal{G}^{-1}(\{0\})$, and Proposition 2.16 assures that

$$\begin{aligned} 0 &= \int_{\mathfrak{q}^{\theta}_{m^{\theta}_{0,2}-1}}^{\mathfrak{q}^{\theta}_{m^{\theta}_{1,1}+1}} x(\mathscr{N}^{h,\theta}_{\infty}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x \\ &= \int_{\mathfrak{q}^{\theta}_{m^{\theta}_{0,2}-1}}^{\mathfrak{q}^{\theta}_{m^{\theta}_{0,2}}} x(\mathscr{N}^{h,\theta}_{\infty}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x + \int_{\mathfrak{q}^{\theta}_{m^{\theta}_{1,1}}}^{\mathfrak{q}^{\theta}_{m^{\theta}_{1,1}+1}} x(\mathscr{N}^{h,\theta}_{\infty}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x \\ &= \frac{\alpha^{\theta}}{12}(\mathfrak{q}^{\theta}_1)^3 - \frac{\beta^{\theta}}{12}(1 - \mathfrak{q}^{\theta}_h)^3 = \frac{\alpha^{\theta}}{12}\left((\mathfrak{q}^{\theta}_1)^3 + \frac{m^{\theta}_{0,2}}{h - m^{\theta}_{1,1} + 1}(1 - \mathfrak{q}^{\theta}_h)^3\right). \end{aligned} \tag{2.263}$$

This implies that

$$\mathfrak{q}^{\theta}_1 = -\sqrt[3]{\frac{m^{\theta}_{0,2}}{h - m^{\theta}_{1,1} + 1}}(1 - \mathfrak{q}^{\theta}_h) < 0 \tag{2.264}$$

which is a contradiction. The proof of Lemma 2.30 is thus complete. □

Lemma 2.31. *Assume Setting 2.1, let $h \in \mathbb{N} \cap (1, \infty)$ and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}_h}$ satisfy for all $n \in \mathbb{N}$ that $0 < m^{\theta_n}_{0,2} < m^{\theta_n}_{1,1} < h + 1$ and*

$$(m^{\theta_n}_{0,2}\alpha^{\theta_n} + (h - m^{\theta_n}_{1,1} + 1)\beta^{\theta_n})(\mathfrak{q}^{\theta_n}_{m^{\theta_n}_{1,1}} - \mathfrak{q}^{\theta_n}_{m^{\theta_n}_{0,2}}) + \frac{1}{2}\alpha^{\theta_n}\mathfrak{q}^{\theta_n}_1 = 1 - \frac{1}{2}\beta^{\theta_n}(1 - \mathfrak{q}^{\theta_n}_h). \tag{2.265}$$

Then there exists $c \in (0, \infty)$ such that for all $n \in \mathbb{N}$ it holds that

$$\max\{|\alpha^{\theta_n}|, |\beta^{\theta_n}|\} \geq c. \tag{2.266}$$

Proof of Lemma 2.31. We prove (2.266) by contradiction. We thus assume that for every $n \in \mathbb{N}$ it holds that

$$\max\{|\alpha^{\theta_n}|, |\beta^{\theta_n}|\} \leq \frac{1}{n}. \tag{2.267}$$

Observe that (2.267) implies that $\lim_{n\to\infty} \alpha^{\theta_n} = \lim_{n\to\infty} \beta^{\theta_n} = 0$. Combining this with (2.265) demonstrates that

$$\begin{aligned} 0 &= \lim_{n\to\infty}(m^{\theta_n}_{0,2}\alpha^{\theta_n} + (h - m^{\theta_n}_{1,1} + 1)\beta^{\theta_n})(\mathfrak{q}^{\theta_n}_{m^{\theta_n}_{1,1}} - \mathfrak{q}^{\theta_n}_{m^{\theta_n}_{0,2}}) + \frac{1}{2}\alpha^{\theta_n}\mathfrak{q}^{\theta_n}_1 \\ &= \lim_{n\to\infty} 1 - \frac{1}{2}\beta^{\theta_n}(1 - \mathfrak{q}^{\theta_n}_h) = 1. \end{aligned} \tag{2.268}$$

This is a contradiction. The proof of Lemma 2.31 is thus complete. □

Lemma 2.32. *Assume Setting 2.1 and let $h \in \mathbb{N} \cap (1, \infty)$. Then there exists $\varepsilon \in (0, \infty)$ such that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\forall\, i \in \{0, 1, \ldots, h\} \backslash \{m^\theta_{0,2}\} \colon \mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0$, $\prod_{k=1}^h \mathfrak{v}^\theta_k \neq 0 < \mathfrak{q}^\theta_1 < \mathfrak{q}^\theta_2 < \ldots < \mathfrak{q}^\theta_h < 1$, $\alpha^\theta \beta^\theta \neq 0 < m^\theta_{0,2} < m^\theta_{1,1} < h + 1$, and $\mathfrak{w}^\theta_{m^\theta_{0,2}} < 0 < \mathfrak{w}^\theta_{m^\theta_{1,1}}$ it holds that*

$$\mathcal{L}^h_\infty(\theta) \geq \varepsilon. \tag{2.269}$$

Proof of Lemma 2.32. Throughout this proof let $R \subseteq \mathbb{R}^{\mathfrak{d}_h}$ satisfy

$$\begin{aligned} R = \{\theta \in (\mathcal{G}^h)^{-1}(\{0\}) \colon [\forall\, i \in \{0, 1, \ldots, h\} \backslash \{m^\theta_{0,2}\} \colon \mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0], \mathfrak{w}^\theta_{m^\theta_{0,2}} < 0 < \mathfrak{w}^\theta_{m^\theta_{1,1}}, \\ \textstyle\prod_{k=1}^h \mathfrak{v}^\theta_k \neq 0 < \mathfrak{q}^\theta_1 < \mathfrak{q}^\theta_2 < \ldots < \mathfrak{q}^\theta_h < 1, \alpha^\theta \beta^\theta \neq 0 < m^\theta_{0,2} < m^\theta_{1,1} < h + 1\}. \end{aligned} \tag{2.270}$$

Note that Lemma 2.23 and Lemma 2.24 ensure for all $\theta \in R$ that

(i) for all $j \in \{0, 1, \ldots, m^\theta_{0,2}\}$ it holds that

$$-\frac{\alpha^\theta \mathfrak{q}^\theta_1}{2} = (-1)^j \mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_j), \tag{2.271}$$

(ii) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \ldots, h + 1\}$ it holds that

$$\frac{\beta^\theta}{2}(1 - \mathfrak{q}^\theta_h) = (-1)^{h+1-j}(\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_j) - 1), \tag{2.272}$$

(iii) it holds that

$$\begin{aligned} &\int_{[0,1]\backslash[\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]} (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \\ &= \frac{m^\theta_{0,2}}{12}(\alpha^\theta)^2(\mathfrak{q}^\theta_1)^3 + \frac{1}{12}(h + 1 - m^\theta_{1,1})(\beta^\theta)^2(1 - \mathfrak{q}^\theta_h)^3, \end{aligned} \tag{2.273}$$

and

(iv) for all $j \in \{0, 1, \ldots, h\} \backslash \{m^\theta_{0,2}\}$, $x \in [\mathfrak{q}^\theta_j, \mathfrak{q}^\theta_{j+1}]$ it holds that

$$\begin{aligned} &\mathscr{N}^{h,\theta}_\infty(x) \\ &= \begin{cases} (-1)^j \alpha^\theta x + (-1)^{j+1}(j + 1/2)\alpha^\theta \mathfrak{q}^\theta_1 & : x \leq \mathfrak{q}^\theta_{m^\theta_{0,2}} \\ (-1)^{h-j}\beta^\theta x + 1 + (-1)^{h+1-j}\beta^\theta \left(1 - (h + 1/2 - j)(1 - \mathfrak{q}^\theta_h)\right) & : x \geq \mathfrak{q}^\theta_{m^\theta_{1,1}}. \end{cases} \end{aligned} \tag{2.274}$$

This, Lemma 2.25, and Lemma 2.26 assure that for all $\theta \in R$, $j \in \{1, 2, \ldots, h\} \backslash \{m^\theta_{0,2}, m^\theta_{1,1}\}$ it holds that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = \begin{cases} -2\alpha^\theta & : j < m^\theta_{0,2} \\ -2\beta^\theta & : j > m^\theta_{1,1}. \end{cases} \tag{2.275}$$

Observe that the fact that for all $\theta \in R$ it holds that $\mathfrak{w}^\theta_{m^\theta_{0,2}} < 0 < \mathfrak{w}^\theta_{m^\theta_{1,1}}$ and the fact that for all $\theta \in R$, $i \in \{0, 1, \ldots, h\} \backslash \{m^\theta_{0,2}\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0$ prove that there exist $k_1, k_2 \in \mathbb{N}$ such that for all $\theta \in R$ it holds that $m^\theta_{0,2} = 2k_1$, $h - m^\theta_{1,1} = 2k_2 - 1$, and

$$\{k \in \{1, 2, \ldots, h\} \colon I^\theta_k \cap (\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}) \neq \varnothing\} = (\cup_{k=0}^{k_1-1}\{2k+1\}) \cup (\cup_{k=0}^{k_2-1}\{m^\theta_{1,1} + 1 + 2k\}). \tag{2.276}$$

This, (2.271), and (2.275) assure that for all $\theta \in R$, $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ it holds that

$$\mathscr{N}^{h,\theta}_\infty(x) = -(m^\theta_{0,2}\alpha^\theta + (h - m^\theta_{1,1} + 1)\beta^\theta)(x - \mathfrak{q}^\theta_{m^\theta_{0,2}}) - \frac{\alpha^\theta \mathfrak{q}^\theta_1}{2}. \tag{2.277}$$

Combining this, (2.272), and the fact that for all $\theta \in R$ the function $\mathscr{N}_\infty^{h,\theta}$ is continuous proves for all $\theta \in R$ that

$$-(m_{0,2}^\theta \alpha^\theta + (h - m_{1,1}^\theta + 1)\beta^\theta)(\mathfrak{q}_{m_{1,1}^\theta}^\theta - \mathfrak{q}_{m_{0,2}^\theta}^\theta) - \frac{\alpha^\theta \mathfrak{q}_1^\theta}{2} = 1 + \frac{\beta^\theta}{2}(1 - \mathfrak{q}_h^\theta). \tag{2.278}$$

Furthermore, note that (2.277) and Lemma 2.6 imply for all $\theta \in R$ that

$$\int_{\mathfrak{q}_{m_{0,2}^\theta}^\theta}^{\frac{1}{2}} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \geq \frac{1}{12}\left(\frac{1}{2} - \mathfrak{q}_{m_{0,2}^\theta}^\theta\right)^3 (m_{0,2}^\theta \alpha^\theta + (h - m_{1,1}^\theta + 1)\beta^\theta)^2 \tag{2.279}$$

and

$$\int_{\frac{1}{2}}^{\mathfrak{q}_{m_{1,1}^\theta}^\theta} (\mathscr{N}_\infty^{h,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \geq \frac{1}{12}\left(\mathfrak{q}_{m_{1,1}^\theta}^\theta - \frac{1}{2}\right)^3 (m_{0,2}^\theta \alpha^\theta + (h - m_{1,1}^\theta + 1)\beta^\theta)^2. \tag{2.280}$$

This and (2.273) establish for all $\theta \in R$ that

$$\begin{aligned} \mathcal{L}_\infty^h(\theta) \geq \frac{1}{12}\left(\left(\mathfrak{q}_{m_{1,1}^\theta}^\theta - \frac{1}{2}\right)^3 + \left(\frac{1}{2} - \mathfrak{q}_{m_{0,2}^\theta}^\theta\right)^3\right)&(m_{0,2}^\theta \alpha^\theta + (h - m_{1,1}^\theta + 1)\beta^\theta)^2 \\ + \frac{1}{12}(\beta^\theta)^2(1 - \mathfrak{q}_{m_{1,1}^\theta}^\theta)^3 + \frac{1}{12}(\alpha^\theta)^2(\mathfrak{q}_{m_{0,2}^\theta}^\theta)^3.& \end{aligned} \tag{2.281}$$

We prove (2.269) by contradiction. Assume that for every $n \in \mathbb{N}$ there exists $\theta_n \in (\mathcal{G}^h)^{-1}(\{0\})$ with $0 < \mathfrak{q}_1^{\theta_n} < \mathfrak{q}_2^{\theta_n} < \ldots < \mathfrak{q}_h^{\theta_n} < 1$, $\alpha^{\theta_n}\beta^{\theta_n} \neq 0 < m_{0,2}^{\theta_n} < m_{1,1}^{\theta_n} < h+1$, $\mathfrak{w}_{m_{0,2}^{\theta_n}}^{\theta_n} < 0 < \mathfrak{w}_{m_{1,1}^{\theta_n}}^{\theta_n}$, and $\forall\, i \in \{0, 1, \ldots, h\} \backslash \{m_{0,2}^{\theta_n}\}\colon \mathfrak{w}_i^{\theta_n}\mathfrak{w}_{i+1}^{\theta_n} < 0$ which satisfies that

$$\mathcal{L}_\infty^h(\theta_n) \leq \frac{1}{n}. \tag{2.282}$$

Observe that (2.281) and (2.282) assure that

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{12}\left(\left(\mathfrak{q}_{m_{1,1}^{\theta_n}}^{\theta_n} - \frac{1}{2}\right)^3 + \left(\frac{1}{2} - \mathfrak{q}_{m_{0,2}^{\theta_n}}^{\theta_n}\right)^3\right)&(m_{0,2}^{\theta_n} \alpha^{\theta_n} + (h - m_{1,1}^{\theta_n} + 1)\beta^{\theta_n})^2 \\ + \frac{1}{12}(\beta^{\theta_n})^2(1 - \mathfrak{q}_{m_{1,1}^{\theta_n}}^{\theta_n})^3 + \frac{1}{12}(\alpha^{\theta_n})^2(\mathfrak{q}_{m_{0,2}^{\theta_n}}^{\theta_n})^3 = 0.& \end{aligned} \tag{2.283}$$

Therefore, we obtain that

$$\begin{aligned} &\lim_{n\to\infty} \left(\mathfrak{q}_{m_{1,1}^{\theta_n}}^{\theta_n} - \frac{1}{2}\right)^3 (m_{0,2}^{\theta_n} \alpha^{\theta_n} + (h - m_{1,1}^{\theta_n} + 1)\beta^{\theta_n})^2 = 0, \\ &\lim_{n\to\infty} \left(\frac{1}{2} - \mathfrak{q}_{m_{0,2}^{\theta_n}}^{\theta_n}\right)^3 (m_{0,2}^{\theta_n} \alpha^{\theta_n} + (h - m_{1,1}^{\theta_n} + 1)\beta^{\theta_n})^2 = 0, \\ &\lim_{n\to\infty} (\beta^{\theta_n})^2(1 - \mathfrak{q}_{m_{1,1}^{\theta_n}}^{\theta_n})^3 = 0, \qquad \text{and} \\ &\lim_{n\to\infty} (\alpha^{\theta_n})^2(\mathfrak{q}_{m_{0,2}^{\theta_n}}^{\theta_n})^3 = 0. \end{aligned} \tag{2.284}$$

Note that (2.278) and Lemma 2.31 demonstrate that there exists $c \in (0, \infty)$ such that for all $n \in \mathbb{N}$ it holds that $\max\{|\alpha^{\theta_n}|, |\beta^{\theta_n}|\} \geq c$. Combining this and (2.284) with Lemma 2.15 assure that there exists a strictly increasing $n\colon \mathbb{N} \to \mathbb{N}$ such that

$$\lim_{k\to\infty} \mathfrak{q}_{m_{1,1}^{\theta_{n(k)}}}^{\theta_{n(k)}} = 1 \qquad \text{and} \qquad \lim_{k\to\infty} \mathfrak{q}_{m_{0,2}^{\theta_{n(k)}}}^{\theta_{n(k)}} = 0. \tag{2.285}$$

This shows that there exists $k^* \in \mathbb{N}$ such that for all $k \in \mathbb{N} \cap [k^*, \infty)$ it holds that $\mathfrak{q}_{m_{1,1}^\theta}^{\theta_{n(k)}} \geq 3/4$ and $\mathfrak{q}_{m_{0,2}^\theta}^{\theta_{n(k)}} \leq 1/4$. Combining this and Proposition 2.5 implies that $\lim_{k\to\infty} \mathcal{L}_\infty^h(\theta_{n(k)}) \geq 1/32$ which is a contradiction. The proof of Lemma 2.32 is thus complete. □

Lemma 2.33. *Assume Setting 2.1 and let $h \in \mathbb{N} \cap (1, \infty)$. Then there exists $\varepsilon \in (0, \infty)$ such that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\forall\, i \in \{0, \ldots, h\} \backslash \{m^\theta_{0,2}\} \colon \mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0$, $\prod_{k=1}^h \mathfrak{v}^\theta_k \neq 0 < \mathfrak{q}^\theta_1 < \mathfrak{q}^\theta_2 < \ldots < \mathfrak{q}^\theta_h < 1$, $\alpha^\theta \beta^\theta \neq 0 < m^\theta_{0,2} < m^\theta_{1,1} < h+1$, and $\mathfrak{w}^\theta_{m^\theta_{1,1}} < 0 < \mathfrak{w}^\theta_{m^\theta_{0,2}}$ it holds that*

$$\mathcal{L}^h_\infty(\theta) \geq \varepsilon. \tag{2.286}$$

Proof of Lemma 2.33. Throughout this proof let $R \subseteq \mathbb{R}^{\mathfrak{d}_h}$ satisfy

$$\begin{aligned} R = \{&\theta \in (\mathcal{G}^h)^{-1}(\{0\}) \colon [\forall\, i \in \{0, 1, \ldots, h\} \backslash \{m^\theta_{0,2}\} \colon \mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0], \mathfrak{w}^\theta_{m^\theta_{1,1}} < 0 < \mathfrak{w}^\theta_{m^\theta_{0,2}}, \\ &\textstyle\prod_{k=1}^h \mathfrak{v}^\theta_k \neq 0 < \mathfrak{q}^\theta_1 < \mathfrak{q}^\theta_2 < \ldots < \mathfrak{q}^\theta_h < 1, \alpha^\theta \beta^\theta \neq 0 < m^\theta_{0,2} < m^\theta_{1,1} < h+1\}. \end{aligned} \tag{2.287}$$

Observe that Lemma 2.23 and Lemma 2.24 ensure for all $\theta \in R$ that

(i) for all $j \in \{0, 1, \ldots, m^\theta_{0,2}\}$ it holds that

$$-\frac{\alpha^\theta \mathfrak{q}^\theta_1}{2} = (-1)^j \mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_j), \tag{2.288}$$

(ii) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1}+1, \ldots, h+1\}$ it holds that

$$\frac{\beta^\theta}{2}(1 - \mathfrak{q}^\theta_h) = (-1)^{h+1-j}(\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_j) - 1), \tag{2.289}$$

(iii) it holds that

$$\begin{aligned} &\int_{[0,1] \backslash [\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]} (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2, \infty)}(x))^2 \, \mathrm{d}x \\ &= \frac{m^\theta_{0,2}}{12}(\alpha^\theta)^2 (\mathfrak{q}^\theta_1)^3 + \frac{1}{12}(h + 1 - m^\theta_{1,1})(\beta^\theta)^2 (1 - \mathfrak{q}^\theta_h)^3, \end{aligned} \tag{2.290}$$

and

(iv) for all $j \in \{0, 1, \ldots, h\} \backslash \{m^\theta_{0,2}\}$, $x \in [\mathfrak{q}^\theta_j, \mathfrak{q}^\theta_{j+1}]$ it holds that

$$\begin{aligned} &\mathscr{N}^{h,\theta}_\infty(x) \\ &= \begin{cases} (-1)^j \alpha^\theta x + (-1)^{j+1}(j + 1/2)\alpha^\theta \mathfrak{q}^\theta_1 & \colon x \leq \mathfrak{q}^\theta_{m^\theta_{0,2}} \\ (-1)^{h-j}\beta^\theta x + 1 + (-1)^{h+1-j}\beta^\theta \left(1 - (h + 1/2 - j)(1 - \mathfrak{q}^\theta_h)\right) & \colon x \geq \mathfrak{q}^\theta_{m^\theta_{1,1}}. \end{cases} \end{aligned} \tag{2.291}$$

This, Lemma 2.25, and Lemma 2.26 assure that for all $\theta \in R$, $j \in \{1, 2, \ldots, h\} \backslash \{m^\theta_{0,2}, m^\theta_{1,1}\}$ it holds that

$$\mathfrak{w}^\theta_j \mathfrak{v}^\theta_j = \begin{cases} -2\alpha^\theta & \colon j < m^\theta_{0,2} \\ -2\beta^\theta & \colon j > m^\theta_{1,1}. \end{cases} \tag{2.292}$$

Furthermore, note that the fact that for all $\theta \in R$ it holds that $\mathfrak{w}^\theta_{m^\theta_{1,1}} < 0 < \mathfrak{w}^\theta_{m^\theta_{0,2}}$ and the fact that for all $\theta \in R$, $i \in \{0, 1, \ldots, h\} \backslash \{m^\theta_{0,2}\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0$ prove that there exist $k_1, k_2 \in \mathbb{N}$ such that for all $\theta \in R$ it holds that $m^\theta_{0,2} = 2k_1 - 1$, $h - m^\theta_{1,1} = 2k_2 - 2$,

$$\begin{aligned} \{k \in \{1, \ldots, h\} \colon I^\theta_k \cap (\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}) \neq \varnothing\} &= (\cup_{k=0}^{k_1-1}\{2k+1\}) \cup (\cup_{k=0}^{k_2-1}\{m^\theta_{1,1} + 2k\}), \\ \{k \in \{1, \ldots, h\} \colon I^\theta_k \cap (\mathfrak{q}^\theta_{m^\theta_{0,2}-1}, \mathfrak{q}^\theta_{m^\theta_{0,2}}) \neq \varnothing\} &= \big((\cup_{k=0}^{k_1-1}\{2k+1\}) \backslash \{m^\theta_{0,2}\} \\ &\qquad \cup (\cup_{k=0}^{k_2-1}\{m^\theta_{1,1} + 2k\})\big), \qquad \text{and} \\ \{k \in \{1, \ldots, h\} \colon I^\theta_k \cap (\mathfrak{q}^\theta_{m^\theta_{1,1}}, \mathfrak{q}^\theta_{m^\theta_{1,1}+1}) \neq \varnothing\} &= \big((\cup_{k=0}^{k_1-1}\{2k+1\}) \\ &\qquad \cup (\cup_{k=0}^{k_2-1}\{m^\theta_{1,1} + 2k\}) \backslash \{m^\theta_{1,1}\}\big). \end{aligned} \tag{2.293}$$

This, (2.292), the fact that for all $\theta \in R$, $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}-1}, \mathfrak{q}^\theta_{m^\theta_{0,2}}]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = \alpha^\theta x - (m^\theta_{0,2} - 1/2)\alpha^\theta \mathfrak{q}^\theta_1$, and the fact that for all $\theta \in R$, $x \in [\mathfrak{q}^\theta_{m^\theta_{1,1}}, \mathfrak{q}^\theta_{m^\theta_{1,1}+1}]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = \beta^\theta x + 1 - \beta^\theta \big(1 - (h + 1/2 - m^\theta_{1,1})(1 - \mathfrak{q}^\theta_h)\big)$ demonstrate for all $\theta \in R$ that

$$\begin{aligned} &\alpha^\theta = -((m^\theta_{0,2} - 1)\alpha^\theta + (h - m^\theta_{1,1})\beta^\theta) + \mathfrak{v}^\theta_{m^\theta_{1,1}} \mathfrak{w}^\theta_{m^\theta_{1,1}} \\ \text{and} \quad &\beta^\theta = -((m^\theta_{0,2} - 1)\alpha^\theta + (h - m^\theta_{1,1})\beta^\theta) + \mathfrak{v}^\theta_{m^\theta_{0,2}} \mathfrak{w}^\theta_{m^\theta_{0,2}}. \end{aligned} \tag{2.294}$$

Combining this, (2.288), and (2.293) ensures that for all $\theta \in R$, $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ it holds that

$$\begin{aligned} \mathscr{N}^{h,\theta}_\infty(x) &= \big(- ((m^\theta_{0,2} - 1)\alpha^\theta + (h - m^\theta_{1,1})\beta^\theta) + \mathfrak{v}^\theta_{m^\theta_{0,2}} \mathfrak{w}^\theta_{m^\theta_{0,2}} \\ &\quad + \mathfrak{v}^\theta_{m^\theta_{1,1}} \mathfrak{w}^\theta_{m^\theta_{1,1}} \big)(x - \mathfrak{q}^\theta_{m^\theta_{0,2}}) + \frac{\alpha^\theta \mathfrak{q}^\theta_1}{2} \\ &= (m^\theta_{0,2}\alpha^\theta + (h - m^\theta_{1,1} + 1)\beta^\theta)(x - \mathfrak{q}^\theta_{m^\theta_{0,2}}) + \frac{\alpha^\theta \mathfrak{q}^\theta_1}{2}. \end{aligned} \tag{2.295}$$

This, (2.289), and the fact that for all $\theta \in R$ it holds that the function $\mathscr{N}^{h,\theta}_\infty$ is continuous prove for all $\theta \in R$ that

$$(m^\theta_{0,2}\alpha^\theta + (h - m^\theta_{1,1} + 1)\beta^\theta)(\mathfrak{q}^\theta_{m^\theta_{1,1}} - \mathfrak{q}^\theta_{m^\theta_{0,2}}) + \frac{\alpha^\theta \mathfrak{q}^\theta_1}{2} = 1 - \frac{\beta^\theta}{2}(1 - \mathfrak{q}^\theta_h). \tag{2.296}$$

Moreover, observe that (2.295) and Lemma 2.6 imply for all $\theta \in R$ that

$$\int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^{\frac{1}{2}} (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \geq \frac{1}{12}\left(\frac{1}{2} - \mathfrak{q}^\theta_{m^\theta_{0,2}}\right)^3 (m^\theta_{0,2}\alpha^\theta + (h - m^\theta_{1,1} + 1)\beta^\theta)^2 \tag{2.297}$$

and

$$\int_{\frac{1}{2}}^{\mathfrak{q}^\theta_{m^\theta_{1,1}}} (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \geq \frac{1}{12}\left(\mathfrak{q}^\theta_{m^\theta_{1,1}} - \frac{1}{2}\right)^3 (m^\theta_{0,2}\alpha^\theta + (h - m^\theta_{1,1} + 1)\beta^\theta)^2. \tag{2.298}$$

This and (2.290) establish for all $\theta \in R$ that

$$\begin{aligned} \mathcal{L}^h_\infty(\theta) \geq \frac{1}{12}\left(\left(\mathfrak{q}^\theta_{m^\theta_{1,1}} - \frac{1}{2}\right)^3 + \left(\frac{1}{2} - \mathfrak{q}^\theta_{m^\theta_{0,2}}\right)^3\right)(m^\theta_{0,2}\alpha^\theta + (h - m^\theta_{1,1} + 1)\beta^\theta)^2 \\ + \frac{1}{12}(\beta^\theta)^2(1 - \mathfrak{q}^\theta_{m^\theta_{1,1}})^3 + \frac{1}{12}(\alpha^\theta)^2(\mathfrak{q}^\theta_{m^\theta_{0,2}})^3. \end{aligned} \tag{2.299}$$

We prove (2.286) by contradiction. Assume that for every $n \in \mathbb{N}$ there exists $\theta_n \in (\mathcal{G}^h)^{-1}(\{0\})$ with $0 < \mathfrak{q}^{\theta_n}_1 < \mathfrak{q}^{\theta_n}_2 < \ldots < \mathfrak{q}^{\theta_n}_h < 1$, $\alpha^{\theta_n}\beta^{\theta_n} \neq 0 < m^{\theta_n}_{0,2} < m^{\theta_n}_{1,1} < h + 1$, $\mathfrak{w}^{\theta_n}_{m^{\theta_n}_{1,1}} < 0 < \mathfrak{w}^{\theta_n}_{m^{\theta_n}_{0,2}}$, and $\forall\, i \in \{0, 1, \ldots, h\} \backslash \{m^{\theta_n}_{0,2}\}\colon \mathfrak{w}^{\theta_n}_i \mathfrak{w}^{\theta_n}_{i+1} < 0$ which satisfies that

$$\mathcal{L}^h_\infty(\theta_n) \leq \frac{1}{n}. \tag{2.300}$$

Note that (2.299) and (2.300) assure that

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{12}\left(\left(\mathfrak{q}^{\theta_n}_{m^{\theta_n}_{1,1}} - \frac{1}{2}\right)^3 + \left(\frac{1}{2} - \mathfrak{q}^{\theta_n}_{m^{\theta_n}_{0,2}}\right)^3\right)(m^{\theta_n}_{0,2}\alpha^{\theta_n} + (h - m^{\theta_n}_{1,1} + 1)\beta^{\theta_n})^2 \\ + \frac{1}{12}(\beta^{\theta_n})^2(1 - \mathfrak{q}^{\theta_n}_{m^{\theta_n}_{1,1}})^3 + \frac{1}{12}(\alpha^{\theta_n})^2(\mathfrak{q}^{\theta_n}_{m^{\theta_n}_{0,2}})^3 = 0. \end{aligned} \tag{2.301}$$

Hence, we obtain that

$$\begin{aligned}
&\lim_{n\to\infty}\left(\mathfrak{q}^{\theta_n}_{m^{\theta_n}_{1,1}}-\frac{1}{2}\right)^3(m^{\theta_n}_{0,2}\alpha^{\theta_n}+(h-m^{\theta_n}_{1,1}+1)\beta^{\theta_n})^2=0,\\
&\lim_{n\to\infty}\left(\frac{1}{2}-\mathfrak{q}^{\theta_n}_{m^{\theta_n}_{0,2}}\right)^3(m^{\theta}_{0,2}\alpha^{\theta_n}+(h-m^{\theta_n}_{1,1}+1)\beta^{\theta_n})^2=0,\\
&\lim_{n\to\infty}(\beta^{\theta_n})^2(1-\mathfrak{q}^{\theta_n}_{m^{\theta_n}_{1,1}})^3=0, \qquad \text{and}\\
&\lim_{n\to\infty}(\alpha^{\theta_n})^2(\mathfrak{q}^{\theta_n}_{m^{\theta_n}_{0,2}})^3=0.
\end{aligned} \tag{2.302}$$

Observe that (2.296) and Lemma 2.31 demonstrate that there exists $c\in(0,\infty)$ such that for all $n\in\mathbb{N}$ it holds that $\max\{|\alpha^{\theta_n}|,|\beta^{\theta_n}|\}\geq c$. Combining this and (2.302) with Lemma 2.15 assure that there exists a strictly increasing $n\colon\mathbb{N}\to\mathbb{N}$ such that

$$\lim_{k\to\infty}\mathfrak{q}^{\theta_{n(k)}}_{m^{\theta_{n(k)}}_{1,1}}=1 \qquad \text{and} \qquad \lim_{k\to\infty}\mathfrak{q}^{\theta_{n(k)}}_{m^{\theta_{n(k)}}_{0,2}}=0. \tag{2.303}$$

This shows that there exists $k^*\in\mathbb{N}$ such that for all $k\in\mathbb{N}\cap[k^*,\infty)$ it holds that $\mathfrak{q}^{\theta_{n(k)}}_{m^{\theta}_{1,1}}\geq 3/4$ and $\mathfrak{q}^{\theta_{n(k)}}_{m^{\theta}_{0,2}}\leq 1/4$. Combining this and Proposition 2.5 assures that $\lim_{k\to\infty}\mathcal{L}^h_\infty(\theta_{n(k)})\geq 1/32$ which is a contradiction. The proof of Lemma 2.33 is thus complete. □

2.5 Estimates for the risk of critical points

Corollary 2.34. *Assume Setting 2.1 and let $h\in\mathbb{N}$, $j\in\{1,2,\ldots,h\}$, $\theta\in(\mathcal{G}^h)^{-1}(\{0\})$ satisfy $(0,1)\subseteq I^\theta_j$ and $\prod_{k=1}^h\mathfrak{v}^\theta_k\neq 0$. Then $\mathcal{L}^h_\infty(\theta)\geq 1/36$.*

Proof of Corollary 2.34. Note that the assumption that $\theta\in(\mathcal{G}^h)^{-1}(\{0\})$, the assumption that $\prod_{k=1}^h\mathfrak{v}^\theta_k\neq 0$, the assumption that $(0,1)\subseteq I^\theta_j$, and Proposition 2.16 demonstrate that

$$\int_0^1 x(\mathscr{N}^{h,\theta}_\infty(x)-\mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x=0. \tag{2.304}$$

Combining this with Lemma 2.22 assures that $\mathcal{L}^h_\infty(\theta)\geq 1/36$. The proof of Corollary 2.34 is thus complete. □

Corollary 2.35. *Assume Setting 2.1 and let $h\in\mathbb{N}\cap(1,\infty)$, $i,j\in\{1,2,\ldots,h\}$, $\theta\in(\mathcal{G}^h)^{-1}(\{0\})$ satisfy $i\neq j$, $\mathfrak{q}^\theta_i=\mathfrak{q}^\theta_j\in(0,1)$, and $\mathfrak{w}^\theta_j\mathfrak{w}^\theta_i<0\neq\prod_{k=1}^h\mathfrak{v}^\theta_k$. Then $\mathcal{L}^h_\infty(\theta)\geq 1/36$.*

Proof of Corollary 2.35. Observe that the assumption that $\theta\in(\mathcal{G}^h)^{-1}(\{0\})$, the assumption that $\mathfrak{w}^\theta_j\mathfrak{w}^\theta_i<0\neq\prod_{k=1}^h\mathfrak{v}^\theta_k$, and Proposition 2.16 assure that

$$\int_0^{\mathfrak{q}^\theta_i} x(\mathscr{N}^{h,\theta}_\infty(x)-\mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x=\int_{\mathfrak{q}^\theta_i}^1 x(\mathscr{N}^{h,\theta}_\infty(x)-\mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x=0. \tag{2.305}$$

This implies that

$$\int_0^1 x(\mathscr{N}^{h,\theta}_\infty(x)-\mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x=0. \tag{2.306}$$

Combing this with Lemma 2.22 demonstrates that $\mathcal{L}^h_\infty(\theta)\geq 1/36$. The proof of Corollary 2.35 is thus complete. □

Proposition 2.36. *Assume Setting 2.1 and let $h\in\mathbb{N}\cap(1,\infty)$, $i,j\in\{1,2,\ldots,h\}$, $\theta\in(\mathcal{G}^h)^{-1}(\{0\})$ satisfy $i\neq j$, $\mathfrak{q}^\theta_i=\mathfrak{q}^\theta_j\in(0,1)$, and $\prod_{k=1}^h\mathfrak{v}^\theta_k\neq 0<\mathfrak{w}^\theta_j\mathfrak{w}^\theta_i$. Then there exist $\vartheta\in(\mathcal{G}^h)^{-1}(\{0\})$, $k\in\{1,2,\ldots,h\}$ such that $I^\vartheta_k=\varnothing$ and $\mathscr{N}^{h,\theta}_\infty|_{[0,1]}=\mathscr{N}^{h,\vartheta}_\infty|_{[0,1]}$.*

Proof of Proposition 2.36. Note that the assumption that $\mathfrak{q}_i^\theta = \mathfrak{q}_j^\theta$ and the assumption that $0 < \mathfrak{w}_j^\theta \mathfrak{w}_i^\theta$ demonstrate that

$$\frac{\mathfrak{b}_i^\theta}{|\mathfrak{w}_i^\theta|} = \frac{\mathfrak{b}_j^\theta}{|\mathfrak{w}_j^\theta|}. \tag{2.307}$$

Assume without loss of generality that $i = 1$ and $j = 2$. This and (2.307) ensure for all $x \in [0, 1]$ that

$$\begin{aligned}\mathfrak{v}_1^\theta\big[\mathbb{A}_\infty(\mathfrak{w}_1^\theta x + \mathfrak{b}_1^\theta)\big] + \mathfrak{v}_2^\theta\big[\mathbb{A}_\infty(\mathfrak{w}_2^\theta x + \mathfrak{b}_2^\theta)\big] &= \mathfrak{v}_1^\theta|\mathfrak{w}_1^\theta|\mathbb{A}_\infty\bigg(\frac{\mathfrak{w}_1^\theta}{|\mathfrak{w}_1^\theta|}x + \frac{\mathfrak{b}_1^\theta}{|\mathfrak{w}_1^\theta|}\bigg) \\ &\quad + \mathfrak{v}_2^\theta|\mathfrak{w}_2^\theta|\mathbb{A}_\infty\bigg(\frac{\mathfrak{w}_2^\theta}{|\mathfrak{w}_2^\theta|}x + \frac{\mathfrak{b}_2^\theta}{|\mathfrak{w}_2^\theta|}\bigg) \\ &= \big(\mathfrak{v}_1^\theta|\mathfrak{w}_1^\theta| + \mathfrak{v}_2^\theta|\mathfrak{w}_2^\theta|\big)\mathbb{A}_\infty\bigg(\frac{\mathfrak{w}_1^\theta}{|\mathfrak{w}_1^\theta|}x + \frac{\mathfrak{b}_1^\theta}{|\mathfrak{w}_1^\theta|}\bigg).\end{aligned} \tag{2.308}$$

Let $\vartheta \in \mathbb{R}^{\mathfrak{d}_h}$ satisfy for all $m \in \{1, 2, \dots, h\}\backslash\{1, 2\}$ that $\mathfrak{v}_1^\vartheta = \mathfrak{v}_1^\theta|\mathfrak{w}_1^\theta| + \mathfrak{v}_2^\theta|\mathfrak{w}_2^\theta|$, $\mathfrak{w}_1^\vartheta = \mathfrak{w}_1^\theta/|\mathfrak{w}_1^\theta|$, $\mathfrak{b}_1^\vartheta = \mathfrak{b}_1^\theta/|\mathfrak{w}_1^\theta|$, $\mathfrak{v}_2^\vartheta = \mathfrak{w}_2^\vartheta = \mathfrak{b}_2^\vartheta = 0$, $\mathfrak{c}^\vartheta = \mathfrak{c}^\theta$, $\mathfrak{v}_m^\vartheta = \mathfrak{v}_m^\theta$, $\mathfrak{w}_m^\vartheta = \mathfrak{w}_m^\theta$, and $\mathfrak{b}_m^\vartheta = \mathfrak{b}_m^\theta$. This and (2.308) imply for all $x \in [0, 1]$ that $I_2^\vartheta = \varnothing$, $\mathcal{G}^h(\vartheta) = 0$, and

$$\mathscr{N}_\infty^{h,\theta}(x) = \mathscr{N}_\infty^{h,\vartheta}(x). \tag{2.309}$$

The proof of Proposition 2.36 is thus complete. □

Lemma 2.37. *Assume Setting 2.1 and let $h \in \mathbb{N}$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $i \in \{1, 2, \dots, h\}$ that $\mathfrak{q}_{i-1}^\theta < \mathfrak{q}_i^\theta \le 1/2$ and $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$. Then*

$$\mathcal{L}_\infty^h(\theta) \ge 1/36. \tag{2.310}$$

Proof of Lemma 2.37. In the following we distinguish between the case $\alpha^\theta = 0$ and the case $\alpha^\theta \neq 0$. We first show (2.310) in the case

$$\alpha^\theta = 0. \tag{2.311}$$

Observe that (2.311), the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, and Corollary 2.17 establish that for all $x \in [0, \mathfrak{q}_{m_{0,2}^\theta}^\theta]$ it holds that $\mathscr{N}_\infty^{h,\theta}(x) = 0$. Therefore, we obtain that there exists $a^\theta \in \mathbb{R}$ such that for all $x \in [\mathfrak{q}_{m_{0,2}^\theta}^\theta, 1]$ it holds that $\mathscr{N}_\infty^{h,\theta}(x) = a^\theta(x - \mathfrak{q}_{m_{0,2}^\theta}^\theta)$. Combining this with Proposition 2.3 ensures that

$$\mathcal{L}_\infty^h(\theta) \ge \frac{1}{36}. \tag{2.312}$$

This establishes (2.310) in the case $\alpha^\theta = 0$. In the next step we demonstrate (2.310) in the case

$$\alpha^\theta \neq 0. \tag{2.313}$$

Note that (2.313) and Lemma 2.19 prove that for all $i \in \{1, 2, \dots, m_{0,2}^\theta\}$ it holds that $\mathfrak{w}_i^\theta \mathfrak{w}_{i-1}^\theta < 0$. Combining this with Lemma 2.23 demonstrates that

(i) for all $j \in \{1, 2, \dots, m_{0,2}^\theta\}$ it holds that $\mathfrak{q}_j^\theta = j\mathfrak{q}_1^\theta$,

(ii) for all $j \in \{0, 1, \dots, m_{0,2}^\theta\}$ it holds that

$$-\frac{\alpha^\theta \mathfrak{q}_1^\theta}{2} = (-1)^j \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_j^\theta), \tag{2.314}$$

and

(iii) it holds that

$$\int_0^{\mathfrak{q}^\theta_{m^\theta_{0,2}}} (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x = \frac{m^\theta_{0,2}}{12} (\alpha^\theta)^2 (\mathfrak{q}^\theta_1)^3. \tag{2.315}$$

Furthermore, observe that the assumption that for all $i \in \{1, 2, \dots, h\}$ it holds that $\mathfrak{q}^\theta_i \le 1/2$ implies that there exists $a^\theta \in \mathbb{R}$ such that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{0,2}}, 1]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = a^\theta(x - \mathfrak{q}^\theta_{m^\theta_{0,2}}) + \mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}})$. This, (2.314), and (2.315) establish that

$$\begin{aligned} \mathcal{L}^h_\infty(\theta) &= \frac{\mathfrak{q}^\theta_{m^\theta_{0,2}}}{3} (\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}}))^2 + \int_{\mathfrak{q}^\theta_{m^\theta_{0,2}}}^1 (a^\theta(x - \mathfrak{q}^\theta_{m^\theta_{0,2}}) + \mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}}) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \\ &= \frac{\mathfrak{q}^\theta_{m^\theta_{0,2}}}{3} (\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}}))^2 + \frac{(a^\theta)^2}{3}(1 - (\mathfrak{q}^\theta_{m^\theta_{0,2}})^3) + \frac{1}{2} - \frac{3a^\theta}{4} \\ &\quad + (-a^\theta \mathfrak{q}^\theta_{m^\theta_{0,2}} + \mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}}))^2 (1 - \mathfrak{q}^\theta_{m^\theta_{0,2}}) + (a^\theta \mathfrak{q}^\theta_{m^\theta_{0,2}} - \mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}})) \\ &\quad + a^\theta(-a^\theta \mathfrak{q}^\theta_{m^\theta_{0,2}} + \mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_{m^\theta_{0,2}}))(1 - (\mathfrak{q}^\theta_{m^\theta_{0,2}})^2). \end{aligned} \tag{2.316}$$

Combining this and Lemma 2.13 ensures that $\mathcal{L}^h_\infty(\theta) \ge 1/18$ This shows (2.310) in the case $\alpha^\theta \ne 0$. The proof of Lemma 2.37 is thus complete. ☐

Lemma 2.38. *Assume Setting 2.1 and let $h \in \mathbb{N}$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy for all $i \in \{1, 2, \dots, h\}$ that $1/2 \le \mathfrak{q}^\theta_i < \mathfrak{q}^\theta_{i+1}$ and $\prod_{k=1}^h \mathfrak{v}^\theta_k \ne 0$. Then*

$$\mathcal{L}^h_\infty(\theta) \ge 1/36. \tag{2.317}$$

Proof of Lemma 2.38. In the following we distinguish between the case $\beta^\theta = 0$ and the case $\beta^\theta \ne 0$. We first demonstrate (2.317) in the case

$$\beta^\theta = 0. \tag{2.318}$$

Note that (2.318), the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, and Corollary 2.18 establish that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{1,1}}, 1]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = 1$. Hence, we obtain that there exists $a^\theta \in \mathbb{R}$ such that for all $x \in [0, \mathfrak{q}^\theta_{m^\theta_{1,1}}]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = a^\theta(x - \mathfrak{q}^\theta_{m^\theta_{1,1}}) + 1$. Combining this with Proposition 2.4 ensures that

$$\mathcal{L}^h_\infty(\theta) \ge \frac{1}{36}. \tag{2.319}$$

This establishes (2.317) in the case $\beta^\theta = 0$. In the next step we demonstrate (2.317) in the case

$$\beta^\theta \ne 0. \tag{2.320}$$

Observe that (2.320) and Lemma 2.20 prove that for all $i \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \dots, h\}$ it holds that $\mathfrak{w}^\theta_i \mathfrak{w}^\theta_{i+1} < 0$. Combining this with Lemma 2.24 demonstrates that

(i) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \dots, h\}$ it holds that $\mathfrak{q}^\theta_j = 1 - (h + 1 - j)(1 - \mathfrak{q}^\theta_h)$,

(ii) for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \dots, h + 1\}$ it holds that

$$\frac{\beta^\theta}{2}(1 - \mathfrak{q}^\theta_h) = (-1)^{h+1-j} (\mathscr{N}^{h,\theta}_\infty(\mathfrak{q}^\theta_j) - 1), \tag{2.321}$$

(iii) and for all $j \in \{m^\theta_{1,1}, m^\theta_{1,1} + 1, \dots, h\}$, $x \in [\mathfrak{q}^\theta_j, \mathfrak{q}^\theta_{j+1}]$ it holds that

$$\int_{\mathfrak{q}^\theta_{m^\theta_{1,1}}}^1 (\mathscr{N}^{h,\theta}_\infty(x) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x = \frac{1}{12}(h + 1 - m^\theta_{1,1})(\beta^\theta)^2 (1 - \mathfrak{q}^\theta_h)^3. \tag{2.322}$$

Furthermore, note that the assumption that for all $i \in \{1, 2, \ldots, h\}$ it holds that $1/2 \leq \mathfrak{q}_i^\theta$ implies that there exists $a^\theta \in \mathbb{R}$ such that for all $x \in [0, \mathfrak{q}_{m_{1,1}^\theta}^\theta]$ it holds that $\mathscr{N}_\infty^{h,\theta}(x) = a^\theta(x - \mathfrak{q}_{m_{1,1}^\theta}^\theta) + \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_{m_{1,1}^\theta}^\theta)$. This, (2.321), and (2.322) establish that

$$\begin{aligned} \mathcal{L}_\infty^h(\theta) &= \int_0^{\mathfrak{q}_{m_{1,1}^\theta}^\theta} (a^\theta(x - \mathfrak{q}_{m_{1,1}^\theta}^\theta) + \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_{m_{1,1}^\theta}^\theta) - \mathbb{1}_{(1/2,\infty)}(x))^2 \, \mathrm{d}x \\ &\quad + \frac{1}{3}(1 - \mathfrak{q}_{m_{1,1}^\theta}^\theta)(\mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_{m_{1,1}^\theta}^\theta) - 1)^2 \\ &= \frac{(a^\theta)^2}{3}(\mathfrak{q}_{m_{1,1}^\theta}^\theta)^3 + (-a^\theta \mathfrak{q}_{m_{1,1}^\theta}^\theta + \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_{m_{1,1}^\theta}^\theta))^2 \mathfrak{q}_{m_{1,1}^\theta}^\theta - a^\theta \left((\mathfrak{q}_{m_{1,1}^\theta}^\theta)^2 - \frac{1}{4} \right) \\ &\quad + a^\theta(-a^\theta \mathfrak{q}_{m_{1,1}^\theta}^\theta + \mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_{m_{1,1}^\theta}^\theta))(\mathfrak{q}_{m_{1,1}^\theta}^\theta)^2 + \frac{1}{3}(1 - \mathfrak{q}_{m_{1,1}^\theta}^\theta)(\mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_{m_{1,1}^\theta}^\theta) - 1)^2 \\ &\quad + (1 + 2a^\theta \mathfrak{q}_{m_{1,1}^\theta}^\theta - 2\mathscr{N}_\infty^{h,\theta}(\mathfrak{q}_{m_{1,1}^\theta}^\theta)) \left(\mathfrak{q}_{m_{1,1}^\theta}^\theta - \frac{1}{2} \right). \end{aligned} \tag{2.323}$$

Combining this and Lemma 2.14 ensures that $\mathcal{L}_\infty^h(\theta) \geq 1/18$. This shows (2.317) in the case $\beta^\theta \neq 0$. The proof of Lemma 2.38 is thus complete. □

Corollary 2.39. *Assume Setting 2.1 and let $h \in \mathbb{N} \cap (1, \infty)$. Then there exists $\varepsilon \in (0, \infty)$ such that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\forall\, i \in \{0, 1, \ldots, h\} \backslash \{m_{0,2}^\theta\} \colon \mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta < 0 \neq \alpha^\theta \beta^\theta$, $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$, and $M_0^\theta \neq \varnothing \neq M_1^\theta$ it holds that $\mathcal{L}_\infty^h(\theta) \geq \varepsilon$.*

Proof of Corollary 2.39. Observe that Corollary 2.29 implies that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\forall\, i \in \{0, 1, \ldots, h\} \colon \mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta < 0 \neq \alpha^\theta \beta^\theta$, $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$, and $\mathfrak{q}_{m_{0,2}^\theta}^\theta = \mathfrak{q}_{m_{1,1}^\theta}^\theta = 1/2$ it holds that

$$\mathcal{L}_\infty^h(\theta) \geq \frac{1}{12}. \tag{2.324}$$

Furthermore, note that Lemma 2.30 assures that for all $\theta \in \mathbb{R}^{\mathfrak{d}_h}$ with $\forall\, i \in \{0, 1, \ldots, h\} \backslash \{m_{0,2}^\theta\} \colon \mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta < 0$, $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$, $\alpha^\theta \beta^\theta \neq 0 < m_{0,2}^\theta < m_{1,1}^\theta < h + 1$, and $0 < \mathfrak{w}_{m_{0,2}^\theta}^\theta \mathfrak{w}_{m_{1,1}^\theta}^\theta$ it holds that

$$\mathcal{G}^h(\theta) \neq 0. \tag{2.325}$$

Moreover, observe that Lemma 2.32 and Lemma 2.33 demonstrate that there exists $\delta \in (0, \infty)$ such that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\forall\, i \in \{0, 1, \ldots, h\} \backslash \{m_{0,2}^\theta\} \colon \mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta < 0$, $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$, $\alpha^\theta \beta^\theta \neq 0 < m_{0,2}^\theta < m_{1,1}^\theta < h + 1$, and $\mathfrak{w}_{m_{0,2}^\theta}^\theta \mathfrak{w}_{m_{1,1}^\theta}^\theta < 0$ it holds that $\mathcal{L}_\infty^h(\theta) \geq \delta$. Combining this with (2.324) and (2.325) shows that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\forall\, i \in \{0, 1, \ldots, h\} \backslash \{m_{0,2}^\theta\} \colon \mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta < 0 \neq \alpha^\theta \beta^\theta$, $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$, and $M_0^\theta \neq \varnothing \neq M_1^\theta$ it holds that $\mathcal{L}_\infty^h(\theta) \geq \min\{1/12, \delta\}$. The proof of Corollary 2.39 is thus complete. □

Lemma 2.40. *Assume Setting 2.1, let $h \in \mathbb{N} \cap (1, \infty)$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, $i \in \{0, 1, \ldots, h\} \backslash \{m_{0,2}^\theta\}$ satisfy $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$, $-\mathfrak{w}_i^\theta \mathfrak{w}_{i+1}^\theta < 0$, and $M_0^\theta \neq \varnothing \neq M_1^\theta$. Then*

$$\mathcal{L}_\infty^h(\theta) \geq \frac{1}{384(1 + h)^2}. \tag{2.326}$$

Proof of Lemma 2.40. Note that the fact that $i \in \{0, 1, \ldots, h\} \backslash \{m_{0,2}^\theta\}$ demonstrates that $1/2 \notin (\mathfrak{q}_i^\theta, \mathfrak{q}_{i+1}^\theta)$. In the following we distinguish between the case $\mathfrak{q}_{i+1}^\theta \leq 1/2$ and the case $\mathfrak{q}_i^\theta \geq 1/2$. We first prove (2.326) in the case

$$\mathfrak{q}_{i+1}^\theta \leq \frac{1}{2}. \tag{2.327}$$

Observe that (2.327) and Lemma 2.19 ensure that for all $x \in [0, \mathfrak{q}^\theta_{m^\theta_{0,2}}]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = 0$. This and Lemma 2.27 imply that

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.328}$$

This establishes (2.326) in the case $\mathfrak{q}^\theta_{i+1} \leq 1/2$. In the next step we demonstrate (2.326) in the case

$$\mathfrak{q}^\theta_i \geq \frac{1}{2}. \tag{2.329}$$

Note that (2.329) and Lemma 2.20 show that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{1,1}}, 1]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = 1$. This and Lemma 2.28 establish that

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.330}$$

This establishes (2.326) in the case $\mathfrak{q}^\theta_i \geq 1/2$. The proof of Lemma 2.40 is thus complete. □

Lemma 2.41. *Assume Setting 2.1, let $h \in \mathbb{N} \cap (1, \infty)$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy $\prod_{k=1}^h \mathfrak{v}^\theta_k \neq 0 < \mathfrak{q}^\theta_1 < \mathfrak{q}^\theta_2 < \ldots < \mathfrak{q}^\theta_h < 1$, $\alpha^\theta \beta^\theta = 0$, and $M^\theta_0 \neq \varnothing \neq M^\theta_1$. Then*

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.331}$$

Proof of Lemma 2.41. Observe that Corollary 2.17, Corollary 2.18, Lemma 2.21, and the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ prove that $\max\{|\alpha^\theta|, |\beta^\theta|\} \neq 0$. In the following we distinguish between the case $\alpha^\theta = 0 \neq \beta^\theta$ and the case $\alpha^\theta \neq 0 = \beta^\theta$. We first demonstrate (2.331) in the case

$$\alpha^\theta = 0 \neq \beta^\theta. \tag{2.332}$$

Note that (2.332), the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, and Corollary 2.17 ensure that for all $x \in [0, \mathfrak{q}^\theta_{m^\theta_{0,2}}]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = 0$. This and Lemma 2.27 imply that

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.333}$$

This establishes (2.331) in the case $\alpha^\theta = 0 \neq \beta^\theta$. In the next step we prove (2.331) in the case

$$\alpha^\theta \neq 0 = \beta^\theta. \tag{2.334}$$

Observe that (2.334), the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$, and Corollary 2.18 show that for all $x \in [\mathfrak{q}^\theta_{m^\theta_{1,1}}, 1]$ it holds that $\mathscr{N}^{h,\theta}_\infty(x) = 1$. This and Lemma 2.28 establish that

$$\mathcal{L}^h_\infty(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.335}$$

This demonstrates (2.331) in the case $\alpha^\theta \neq 0 = \beta^\theta$. The proof of Lemma 2.41 is thus complete. □

Proposition 2.42. *Assume Setting 2.1 and let $\theta \in (\mathcal{G}^1)^{-1}(\{0\})$. Then*

$$\mathcal{L}^1_\infty(\theta) \geq 1/36. \tag{2.336}$$

Proof of Proposition 2.42. Note that in the case $\mathfrak{v}_1^\theta = 0$ there exists $b \in \mathbb{R}$ such that for all $x \in \mathbb{R}$ it holds that $\mathscr{N}_\infty^{1,\theta}(x) = b$. This and Proposition 2.2 establish that in the case $\mathfrak{v}_1^\theta = 0$ it holds that

$$\mathcal{L}_\infty^1(\theta) \geq \frac{1}{16}. \tag{2.337}$$

Assume now that $\mathfrak{v}_1^\theta \neq 0$. In the following we distinguish between the case $\mathfrak{q}_1^\theta \in (0,1)$ and the case $\mathfrak{q}_1^\theta \notin (0,1)$. We first show (2.336) in the case

$$\mathfrak{q}_1^\theta \in (0,1). \tag{2.338}$$

Observe that (2.338), Lemma 2.37, and Lemma 2.38 prove that

$$\mathcal{L}_\infty^1(\theta) \geq \frac{1}{36}. \tag{2.339}$$

This establishes (2.336) in the case $\mathfrak{q}_1^\theta \in (0,1)$. In the next step we demonstrate (2.336) in the case

$$\mathfrak{q}_1^\theta \notin (0,1). \tag{2.340}$$

Note that (2.340) ensures that there exist $a, b \in \mathbb{R}$ such that for all $x \in \mathbb{R}$ it holds that $\mathscr{N}_\infty^{1,\theta}(x) = ax + b$. This and Proposition 2.2 demonstrate that

$$\mathcal{L}_\infty^1(\theta) \geq \frac{1}{16}. \tag{2.341}$$

This establishes (2.336) in the case $\mathfrak{q}_1^\theta \notin (0,1)$. The proof of Proposition 2.42 is thus complete. □

Lemma 2.43. *Assume Setting 2.1 and let $h \in \mathbb{N} \cap (1,\infty)$, $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ satisfy $\{\vartheta \in (\mathcal{G}^{h-1})^{-1}(\{0\}) \colon \mathscr{N}_\infty^{h,\theta}|_{[0,1]} = \mathscr{N}_\infty^{h-1,\vartheta}|_{[0,1]}\} = \varnothing$. Then it holds that*

$$\textstyle\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0. \tag{2.342}$$

Proof of Lemma 2.43. We prove (2.342) by contradiction. Assume that $\prod_{k=1}^h \mathfrak{v}_k^\theta = 0$ and assume without loss of generality that $\mathfrak{v}_h^\theta = 0$. Throughout this proof let $\vartheta \in \mathbb{R}^{\mathfrak{d}_{h-1}}$ satisfy for all $i \in \{1, 2, \ldots, h-1\}$ that

$$\vartheta_i = \theta_i, \qquad \vartheta_{h-1+i} = \theta_{h+i}, \qquad \vartheta_{2h-2+i} = \theta_{2h+i}, \qquad \text{and} \qquad \vartheta_{3h-2} = \theta_{3h+1}. \tag{2.343}$$

Observe that (2.343) assures that

$$\mathscr{N}_\infty^{h,\theta}|_{[0,1]} = \mathscr{N}_\infty^{h-1,\vartheta}|_{[0,1]}. \tag{2.344}$$

Furthermore, note that that (2.343) and Proposition 2.16 show that for all $i \in \{1, 2, \ldots, 3(h-1)\}$ it holds that $\mathcal{G}_i^{h-1}(\vartheta) = \mathcal{G}_i^h(\theta)$ and $\mathcal{G}_{\mathfrak{d}_{h-1}}^{h-1}(\vartheta) = \mathcal{G}_{\mathfrak{d}_h}^h(\theta)$. This, (2.344), and the assumption that $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ imply that $\vartheta \in \{v \in (\mathcal{G}^{h-1})^{-1}(\{0\}) \colon \mathscr{N}_\infty^{h,\theta}|_{[0,1]} = \mathscr{N}_\infty^{h-1,v}|_{[0,1]}\}$ which is a contradiction. The proof of Lemma 2.43 is thus complete. □

Lemma 2.44. *Assume Setting 2.1 and let $h \in \mathbb{N} \cap (1,\infty)$. Then there exists $\varepsilon \in (0,\infty)$ such that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\{\vartheta \in (\mathcal{G}^{h-1})^{-1}(\{0\}) \colon \mathscr{N}_\infty^{h,\theta}|_{[0,1]} = \mathscr{N}_\infty^{h-1,\vartheta}|_{[0,1]}\} = \varnothing$ it holds that $\mathcal{L}_\infty^h(\theta) \geq \varepsilon$.*

Proof of Lemma 2.44. Throughout this proof for every $\theta \in \mathbb{R}^{\mathfrak{d}_h}$ let $R^\theta \subseteq \mathbb{R}^{\mathfrak{d}_{h-1}}$ satisfy $R^\theta = \{\vartheta \in (\mathcal{G}^{h-1})^{-1}(\{0\}) \colon \mathscr{N}_\infty^{h,\theta}|_{[0,1]} = \mathscr{N}_\infty^{h-1,\vartheta}|_{[0,1]}\}$. Observe that Lemma 2.43 ensures that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $R^\theta = \varnothing$ it holds that $\prod_{k=1}^h \mathfrak{v}_k^\theta \neq 0$. Furthermore, note that Corollary 2.34

demonstrates that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\prod_{k=1}^{h} \mathfrak{v}_k^\theta \neq 0$ and $\{k \in \{1,2,\ldots,h\}\colon (0,1) \subseteq I_k^\theta\} \neq \varnothing$ it holds that

$$\mathcal{L}_\infty^h(\theta) \geq \frac{1}{36}. \tag{2.345}$$

Moreover, observe that Corollary 2.35 and Proposition 2.36 assure that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\prod_{k=1}^{h} \mathfrak{v}_k^\theta \neq 0$, $R^\theta = \varnothing$, and $\{k \in \{1,2,\ldots,h\}\colon [\exists\, j \in \{1,2,\ldots,h\}\backslash\{k\}\colon \mathfrak{q}_k^\theta = \mathfrak{q}_j^\theta \in (0,1)]\} \neq \varnothing$ it holds that

$$\mathcal{L}_\infty^h(\theta) \geq \frac{1}{36}. \tag{2.346}$$

In addition, note that Lemma 2.37 and Lemma 2.38 prove that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\prod_{k=1}^{h} \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$ and $\{M_0^\theta, M_1^\theta\} \supseteq \{\varnothing\}$ it holds that

$$\mathcal{L}_\infty^h(\theta) \geq \frac{1}{36}. \tag{2.347}$$

Furthermore, observe that Corollary 2.39 shows that there exists $\delta \in (0,\infty)$ such that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\prod_{k=1}^{h} \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$, $M_0^\theta \neq \varnothing \neq M_1^\theta$, and $\forall\, k \in \{0,1,\ldots,h\}\backslash\{m_{0,2}^\theta\}\colon \mathfrak{w}_k^\theta \mathfrak{w}_{k+1}^\theta < 0 \neq \alpha^\theta \beta^\theta$ we have that

$$\mathcal{L}_\infty^h(\theta) \geq \delta. \tag{2.348}$$

Moreover, note that Lemma 2.40 establishes that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\prod_{k=1}^{h} \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$, $M_0^\theta \neq \varnothing \neq M_1^\theta$, and $\{k \in \{0,1,\ldots,h\}\backslash\{m_{0,2}^\theta\}\colon \mathfrak{w}_k^\theta \mathfrak{w}_{k+1}^\theta > 0\} \neq \varnothing$ it holds that

$$\mathcal{L}_\infty^h(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.349}$$

In addition, observe that Lemma 2.41 proves that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\prod_{k=1}^{h} \mathfrak{v}_k^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < \ldots < \mathfrak{q}_h^\theta < 1$, $\alpha^\theta \beta^\theta = 0$, and $M_0^\theta \neq \varnothing \neq M_1^\theta$ it holds that

$$\mathcal{L}_\infty^h(\theta) \geq \frac{1}{384(1+h)^2}. \tag{2.350}$$

Next note that for every $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $R^\theta = \varnothing$ there exists $\vartheta \in (\mathcal{G}^h)^{-1}(\{0\})$ which satisfies that $\mathscr{N}_\infty^{h,\vartheta} = \mathscr{N}_\infty^{h,\vartheta}$, $R^\vartheta = \varnothing$, and $\mathfrak{q}_1^\vartheta \leq \mathfrak{q}_2^\vartheta \leq \ldots \leq \mathfrak{q}_h^\vartheta$. Combining this, (2.345), (2.346), (2.347), (2.348), (2.349), and (2.350) establishes that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $R^\theta = \varnothing$ it holds that

$$\mathcal{L}_\infty^h(\theta) \geq \min\left\{\frac{1}{384(1+h)^2}, \delta\right\}. \tag{2.351}$$

The proof of Lemma 2.44 is thus complete. □

2.6 Blow up phenomena for GFs in the training of ANNs

Proposition 2.45. *Let $d, h, \mathfrak{d}, n \in \mathbb{N}$, $\mathscr{a} \in \mathbb{R}$, $\mathscr{b} \in (\mathscr{a}, \infty)$ satisfy $\mathfrak{d} = dh+2h+1$, let $f\colon [a,b]^d \to \mathbb{R}$ be a function, for every $i \in \{1,2,\ldots,n\}$, $k \in \{0,1\}$ let $\alpha_i^k \in \mathbb{R}^{n\times d}$, let $\beta_i^k \in \mathbb{R}^n$, and let $P_i^k\colon \mathbb{R}^d \to \mathbb{R}$ be a polynomial, let $\mathfrak{p}\colon [\mathscr{a}, \mathscr{b}]^d \to [0,\infty)$ satisfy for all $k \in \{0,1\}$, $x \in [\mathscr{a}, \mathscr{b}]^d$ that*

$$k f(x) + (1-k)\mathfrak{p}(x) = \textstyle\sum_{i=1}^{n} \big[P_i^k(x) \mathbb{1}_{[0,\infty)^n}(\alpha_i^k x + \beta_i^k)\big], \tag{2.352}$$

let $\mathbb{A}_r \in C(\mathbb{R}, \mathbb{R})$, $r \in \mathbb{N} \cup \{\infty\}$, satisfy for all $x \in \mathbb{R}$ that $(\bigcup_{r\in\mathbb{N}}\{\mathbb{A}_r\}) \subseteq C^1(\mathbb{R},\mathbb{R})$, $\mathbb{A}_\infty(x) = \max\{x, 0\}$, $\sup_{r\in\mathbb{N}} \sup_{y \in [-|x|,|x|]} |(\mathbb{A}_r)'(y)| < \infty$, and

$$\limsup_{r\to\infty} \big(|\mathbb{A}_r(x) - \mathbb{A}_\infty(x)| + |(\mathbb{A}_r)'(x) - \mathbb{1}_{(0,\infty)}(x)|\big) = 0, \tag{2.353}$$

let $\mathcal{L}_r \colon \mathbb{R}^{\mathfrak{d}} \to \mathbb{R}$, $r \in \mathbb{N} \cup \{\infty\}$, *satisfy for all* $r \in \mathbb{N} \cup \{\infty\}$, $\theta = (\theta_1, \dots, \theta_{\mathfrak{d}}) \in \mathbb{R}^{\mathfrak{d}}$ *that*

$$\mathcal{L}_r(\theta) = \int_{[\mathscr{a},\mathscr{b}]^d} \big(f(x_1, \dots, x_d) \\ - \theta_{\mathfrak{d}} - \textstyle\sum_{i=1}^{h} \theta_{h(d+1)+i}\big[\mathbb{A}_r(\theta_{hd+i} + \sum_{j=1}^{d} \theta_{(i-1)d+j} x_j)\big]\big)^2 \mathfrak{p}(x) \, \mathrm{d}(x_1, \dots, x_d), \tag{2.354}$$

let $\mathcal{G} \colon \mathbb{R}^{\mathfrak{d}} \to \mathbb{R}^{\mathfrak{d}}$ *satisfy for all* $\theta \in \{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon ((\nabla \mathcal{L}_r)(\vartheta))_{r \in \mathbb{N}}$ *is convergent*$\}$ *that* $\mathcal{G}(\theta) = \lim_{r \to \infty} (\nabla \mathcal{L}_r)(\theta)$, *and let* $\Theta \in C([0,\infty), \mathbb{R}^{\mathfrak{d}})$ *satisfy* $\liminf_{t \to \infty} \|\Theta_t\| < \infty$ *and* $\forall\, t \in [0,\infty) \colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}(\Theta_s) \, \mathrm{d}s$. *Then there exist* $\vartheta \in \mathcal{G}^{-1}(\{0\})$, $\mathfrak{C}, \beta \in (0,\infty)$ *which satisfy for all* $t \in [0,\infty)$ *that*

$$\|\Theta_t - \vartheta\| \le \mathfrak{C}(1+t)^{-\beta} \qquad \text{and} \qquad |\mathcal{L}_\infty(\Theta_t) - \mathcal{L}_\infty(\vartheta)| \le \mathfrak{C}(1+t)^{-1}. \tag{2.355}$$

Proof of Proposition 2.45. The assertion is verified analogously to the proof of [20, Theorem 1.3] (compare, e.g., [15, Theorem 1.2]). The proof of Proposition 2.45 is thus complete. □

Lemma 2.46. *Assume Setting 2.1. Then it holds for all* $h \in \mathbb{N}$ *that there exists* $\varepsilon \in (0,\infty)$ *such that for all* $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ *it holds that*

$$\mathcal{L}_\infty^h(\theta) \ge \varepsilon. \tag{2.356}$$

Proof of Lemma 2.46. We prove (2.356) by induction. Note that Proposition 2.42 assures that for all $\theta \in (\mathcal{G}^1)^{-1}(\{0\})$ it holds that

$$\mathcal{L}_\infty^1(\theta) \ge \frac{1}{36}. \tag{2.357}$$

For the induction step let $h \in \mathbb{N} \cap (1,\infty)$ and assume that there exists $\epsilon \in (0,\infty)$ which satisfies for all $\vartheta \in (\mathcal{G}^{h-1})^{-1}(\{0\})$ that

$$\mathcal{L}_\infty^{h-1}(\vartheta) \ge \epsilon. \tag{2.358}$$

Observe that (2.358) shows that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\{\vartheta \in (\mathcal{G}^{h-1})^{-1}(\{0\}) \colon \mathscr{N}_\infty^{h,\theta}|_{[0,1]} = \mathscr{N}_\infty^{h-1,\vartheta}|_{[0,1]}\} \neq \varnothing$ there exists $\vartheta \in (\mathcal{G}^{h-1})^{-1}(\{0\})$ such that

$$\mathcal{L}_\infty^h(\theta) = \mathcal{L}_\infty^{h-1}(\vartheta) \ge \epsilon. \tag{2.359}$$

Note that Lemma 2.44 demonstrates that there exists $\delta \in (0,\infty)$ which satisfies for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ with $\{\vartheta \in (\mathcal{G}^{h-1})^{-1}(\{0\}) \colon \mathscr{N}_\infty^{h,\theta}|_{[0,1]} = \mathscr{N}_\infty^{h-1,\vartheta}|_{[0,1]}\} = \varnothing$ that

$$\mathcal{L}_\infty^h(\theta) \ge \delta. \tag{2.360}$$

Observe that (2.359) and (2.360) ensure that for all $\theta \in (\mathcal{G}^h)^{-1}(\{0\})$ it holds that $\mathcal{L}_\infty^h(\theta) \ge \min\{\epsilon, \delta\}$. Induction thus establishes (2.356). The proof of Lemma 2.46 is thus complete. □

Theorem 2.47. *Let* $\mathscr{a} \in \mathbb{R}$, $\mathscr{b} \in (\mathscr{a},\infty)$, $h, \mathfrak{d} \in \mathbb{N}$ *satisfy* $\mathfrak{d} = 3h+1$, *let* $\mathbb{A}_r \in C(\mathbb{R},\mathbb{R})$, $r \in \mathbb{N} \cup \{\infty\}$, *satisfy for all* $x \in \mathbb{R}$ *that* $(\bigcup_{r \in \mathbb{N}} \{\mathbb{A}_r\}) \subseteq C^1(\mathbb{R},\mathbb{R})$, $\mathbb{A}_\infty(x) = \max\{x, 0\}$, $\sup_{r \in \mathbb{N}} \sup_{y \in [-|x|,|x|]} |(\mathbb{A}_r)'(y)| < \infty$, *and*

$$\limsup_{r \to \infty} \big(|\mathbb{A}_r(x) - \mathbb{A}_\infty(x)| + |(\mathbb{A}_r)'(x) - \mathbb{1}_{(0,\infty)}(x)|\big) = 0, \tag{2.361}$$

for every $r \in \mathbb{N} \cup \{\infty\}$, $\theta = (\theta_1, \dots, \theta_{\mathfrak{d}}) \in \mathbb{R}^{\mathfrak{d}}$ *let* $\mathscr{N}_r^\theta \colon \mathbb{R} \to \mathbb{R}$ *satisfy for all* $x \in \mathbb{R}$ *that*

$$\mathscr{N}_r^\theta(x) = \theta_{\mathfrak{d}} + \textstyle\sum_{i=1}^{h} \theta_{2h+i}\big[\mathbb{A}_r(\theta_{h+i} + \theta_i x)\big], \tag{2.362}$$

for every $r \in \mathbb{N} \cup \{\infty\}$ *let* $\mathcal{L}_r \colon \mathbb{R}^{\mathfrak{d}} \to \mathbb{R}$ *satisfy for all* $\theta \in \mathbb{R}^{\mathfrak{d}}$ *that*

$$\mathcal{L}_r(\theta) = \int_{\mathscr{a}}^{\mathscr{b}} \big(\mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x) - \mathscr{N}_r^\theta(x)\big)^2 \, \mathrm{d}x, \tag{2.363}$$

and let $\mathcal{G} \colon \mathbb{R}^{\mathfrak{d}} \to \mathbb{R}^{\mathfrak{d}}$ *satisfy for all* $\theta \in \{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon ((\nabla \mathcal{L}_r)(\vartheta))_{r \in \mathbb{N}}$ *is convergent*$\}$ *that* $\mathcal{G}(\theta) = \lim_{r \to \infty} (\nabla \mathcal{L}_r)(\theta)$. *Then there exists* $\varepsilon \in (0,\infty)$ *such that for all* $\theta \in \mathcal{G}^{-1}(\{0\})$ *it holds that* $\mathcal{L}_\infty(\theta) \ge \varepsilon$.

Proof of Theorem 2.47. Throughout this proof for every $\theta \in \mathbb{R}^{\mathfrak{d}}$ let $v^\theta \in \mathbb{R}^{\mathfrak{d}}$ satisfy for all $i \in \{1, 2, \dots, h\}$ that $v_i^\theta = \theta_i(\mathscr{b} - \mathscr{a})$, $v_{h+i}^\theta = \theta_{h+i} + \theta_i \mathscr{a}$, $v_{2h+i}^\theta = \theta_{2h+i}$, and $v_{\mathfrak{d}}^\theta = \theta_{\mathfrak{d}}$ and for every $\theta \in \mathbb{R}^{\mathfrak{d}}$, $i \in \{1, 2, \dots, h\}$, $\mathscr{c} \in \mathbb{R}$, $\mathscr{d} \in (\mathscr{c}, \infty)$ let $I_{i,\mathscr{d}}^{\theta,\mathscr{c}} \subseteq \mathbb{R}$ satisfy $I_{i,\mathscr{d}}^{\theta,\mathscr{c}} = \{x \in [\mathscr{c}, \mathscr{d}] \colon \theta_i x + \theta_{h+i} > 0\}$. Note that for all $\theta \in \mathbb{R}^{\mathfrak{d}}$ it holds that

$$\begin{aligned} \mathcal{L}_\infty(\theta) &= \int_{\mathscr{a}}^{\mathscr{b}} \big(\mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x) - \mathscr{N}_\infty^\theta(x)\big)^2 \, \mathrm{d}x \\ &= (\mathscr{b} - \mathscr{a}) \int_0^1 \big(\mathbb{1}_{(1/2,\infty)}(y) - \mathscr{N}_\infty^\theta((\mathscr{b} - \mathscr{a})y + \mathscr{a})\big)^2 \, \mathrm{d}y \\ &= (\mathscr{b} - \mathscr{a}) \int_0^1 \big(\mathbb{1}_{(1/2,\infty)}(y) - \mathscr{N}_\infty^{v^\theta}(y)\big)^2 \, \mathrm{d}y. \end{aligned} \tag{2.364}$$

Furthermore, observe that [23, Proposition 2.2] establishes that for all $\theta \in \mathcal{G}^{-1}(\{0\})$, $i \in \{1, 2, \dots, h\}$ it holds that

$$\begin{aligned} 0 = \mathcal{G}_i(\theta) &= 2\theta_{2h+i} \int_{I_{i,\mathscr{b}}^{\theta,\mathscr{a}}} x(\mathscr{N}_\infty^\theta(x) - \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)) \, \mathrm{d}x \\ &= 2v_{2h+i}^\theta (\mathscr{b} - \mathscr{a}) \int_{I_{i,1}^{v^\theta,0}} ((\mathscr{b} - \mathscr{a})y + \mathscr{a})(\mathscr{N}_\infty^{v^\theta}(y) - \mathbb{1}_{(1/2,\infty)}(y)) \, \mathrm{d}y, \\ 0 = \mathcal{G}_{h+i}(\theta) &= 2\theta_{2h+i} \int_{I_{i,\mathscr{b}}^{\theta,\mathscr{a}}} (\mathscr{N}_\infty^\theta(x) - \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)) \, \mathrm{d}x \\ &= 2v_{2h+i}^\theta (\mathscr{b} - \mathscr{a}) \int_{I_{i,1}^{v^\theta,0}} (\mathscr{N}_\infty^{v^\theta}(y) - \mathbb{1}_{(1/2,\infty)}(y)) \, \mathrm{d}y, \\ 0 = \mathcal{G}_{2h+i}(\theta) &= 2 \int_{\mathscr{a}}^{\mathscr{b}} \big[\mathbb{A}_\infty(\theta_i x + \theta_{h+i})\big] (\mathscr{N}_\infty^\theta(x) - \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)) \, \mathrm{d}x \\ &= 2(\mathscr{b} - \mathscr{a}) \int_0^1 \big[\mathbb{A}_\infty(v_i^\theta y + v_{h+i}^\theta)\big] (\mathscr{N}_\infty^{v^\theta}(y) - \mathbb{1}_{(1/2,\infty)}(y)) \, \mathrm{d}y, \\ \text{and} \qquad 0 = \mathcal{G}_{\mathfrak{d}}(\theta) &= 2 \int_{\mathscr{a}}^{\mathscr{b}} (\mathscr{N}_\infty^\theta(x) - \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)) \, \mathrm{d}x \\ &= 2(\mathscr{b} - \mathscr{a}) \int_0^1 (\mathscr{N}_\infty^{v^\theta}(y) - \mathbb{1}_{(1/2,\infty)}(y)) \, \mathrm{d}y. \end{aligned} \tag{2.365}$$

Combining this and (2.364) with Lemma 2.46 demonstrates that there exists $\varepsilon \in (0, \infty)$ such that for all $\theta \in \mathcal{G}^{-1}(\{0\})$ it holds that

$$\mathcal{L}_\infty(\theta) = (\mathscr{b} - \mathscr{a}) \int_0^1 \big(\mathbb{1}_{(1/2,\infty)}(y) - \mathscr{N}_\infty^{v^\theta}(y)\big)^2 \, \mathrm{d}y \geq (\mathscr{b} - \mathscr{a})\varepsilon. \tag{2.366}$$

The proof of Theorem 2.47 is thus complete. ∎

Theorem 2.48. *Let $\mathscr{a} \in \mathbb{R}$, $\mathscr{b} \in (\mathscr{a}, \infty)$, $h, \mathfrak{d} \in \mathbb{N}$ satisfy $\mathfrak{d} = 3h + 1$, let $\mathbb{A}_r \in C(\mathbb{R}, \mathbb{R})$, $r \in \mathbb{N} \cup \{\infty\}$, satisfy for all $x \in \mathbb{R}$ that $(\bigcup_{r\in\mathbb{N}} \{\mathbb{A}_r\}) \subseteq C^1(\mathbb{R}, \mathbb{R})$, $\mathbb{A}_\infty(x) = \max\{x, 0\}$, $\sup_{r\in\mathbb{N}} \sup_{y\in[-|x|,|x|]} |(\mathbb{A}_r)'(y)| < \infty$, and*

$$\limsup_{r\to\infty} \big(|\mathbb{A}_r(x) - \mathbb{A}_\infty(x)| + |(\mathbb{A}_r)'(x) - \mathbb{1}_{(0,\infty)}(x)|\big) = 0, \tag{2.367}$$

for every $r \in \mathbb{N} \cup \{\infty\}$, $\theta = (\theta_1, \dots, \theta_{\mathfrak{d}}) \in \mathbb{R}^{\mathfrak{d}}$ let $\mathscr{N}_r^\theta \colon \mathbb{R} \to \mathbb{R}$ satisfy for all $x \in \mathbb{R}$ that

$$\mathscr{N}_r^\theta(x) = \theta_{\mathfrak{d}} + \textstyle\sum_{i=1}^h \theta_{2h+i} \big[\mathbb{A}_r(\theta_{h+i} + \theta_i x)\big], \tag{2.368}$$

for every $r \in \mathbb{N} \cup \{\infty\}$ let $\mathcal{L}_r \colon \mathbb{R}^{\mathfrak{d}} \to \mathbb{R}$ satisfy for all $\theta \in \mathbb{R}^{\mathfrak{d}}$ that

$$\mathcal{L}_r(\theta) = \int_{\mathscr{a}}^{\mathscr{b}} \big(\mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x) - \mathscr{N}_r^{\theta}(x)\big)^2 \, \mathrm{d}x, \tag{2.369}$$

and let $\mathcal{G} \colon \mathbb{R}^{\mathfrak{d}} \to \mathbb{R}^{\mathfrak{d}}$ satisfy for all $\theta \in \{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon ((\nabla \mathcal{L}_r)(\vartheta))_{r \in \mathbb{N}}$ is convergent$\}$ that $\mathcal{G}(\theta) = \lim_{r \to \infty} (\nabla \mathcal{L}_r)(\theta)$. Then there exists $\varepsilon \in (0,\infty)$ such that for all $\Theta \in C([0,\infty), \mathbb{R}^{\mathfrak{d}})$ with $\forall\, t \in [0,\infty) \colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}(\Theta_s)\, \mathrm{d}s$ and $\mathcal{L}_\infty(\Theta_0) < \varepsilon$ it holds that $\liminf_{t \to \infty} \|\Theta_t\| = \infty$.

Proof of Theorem 2.48. Note that Theorem 2.47 ensures that there exists $\varepsilon \in (0,\infty)$ which satisfies for all $\theta \in \mathcal{G}^{-1}(\{0\})$ that

$$\mathcal{L}_\infty(\theta) \geq \varepsilon. \tag{2.370}$$

Furthermore, observe that, e.g., [23, Lemma 3.1] implies that for all $\Theta \in C([0,\infty), \mathbb{R}^{\mathfrak{d}})$ with $\forall\, t \in [0,\infty) \colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}(\Theta_s)\, \mathrm{d}s$ it holds that $[0,\infty) \ni t \mapsto \mathcal{L}_\infty(\Theta_t) \in \mathbb{R}$ is non-increasing. Combining this with (2.370) and Proposition 2.45 assures that for all $\Theta \in C([0,\infty), \mathbb{R}^{\mathfrak{d}})$ with $\forall\, t \in [0,\infty) \colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}(\Theta_s)\, \mathrm{d}s$ and $\mathcal{L}_\infty(\Theta_0) < \varepsilon$ it holds that $\liminf_{t \to \infty} \|\Theta_t\| = \infty$. The proof of Theorem 2.48 is thus complete. □

2.7 Blow up phenomena for GFs in the training of ANNs with two hidden neurons

Lemma 2.49. *Assume Setting 2.1 and let $\theta \in \mathbb{R}^7$ satisfy $\mathfrak{w}_1^\theta \mathfrak{w}_2^\theta > 0 \neq \mathfrak{v}_1^\theta \mathfrak{v}_2^\theta$, $0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < 1$, and $1/2 \in (\mathfrak{q}_1^\theta, \mathfrak{q}_2^\theta)$. Then*

$$\mathcal{G}^2(\theta) \neq 0. \tag{2.371}$$

Proof of Lemma 2.49. We prove (2.371) by contradiction. Assume that $\mathcal{G}^2(\theta) = 0$. In the following we distinguish between the case $\min\{\mathfrak{w}_1^\theta, \mathfrak{w}_2^\theta\} > 0$ and the case $\max\{\mathfrak{w}_1^\theta, \mathfrak{w}_2^\theta\} < 0$. We first establish the contradiction in the case

$$\min\{\mathfrak{w}_1^\theta, \mathfrak{w}_2^\theta\} > 0. \tag{2.372}$$

Observe that (2.372), the fact that $\mathcal{G}^2(\theta) = 0$, and Proposition 2.16 imply that for all $i \in \{1,2\}$ it holds that

$$\int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} x(\mathscr{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\, \mathrm{d}x = \int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} (\mathscr{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\, \mathrm{d}x = 0. \tag{2.373}$$

This, Lemma 2.7, and Corollary 2.10 demonstrate that for all $x \in [\mathfrak{q}_1^\theta, 1]$ it holds that $\mathfrak{q}_2^\theta = 3/4 - 1/2\mathfrak{q}_1^\theta$ and

$$\mathscr{N}_\infty^{2,\theta}(x) = \begin{cases} 1 & \colon x \in [\mathfrak{q}_2^\theta, 1] \\ -\frac{16x}{9(2\mathfrak{q}_1^\theta - 1)} + \frac{10\mathfrak{q}_1^\theta + 3}{9(2\mathfrak{q}_1^\theta - 1)} & \colon x \in [\mathfrak{q}_1^\theta, \mathfrak{q}_2^\theta). \end{cases} \tag{2.374}$$

Combining this with the fact that $I_1^\theta = (\mathfrak{q}_1^\theta, 1]$, the fact that $I_2^\theta = (\mathfrak{q}_2^\theta, 1]$, and continuity of $\mathscr{N}_\infty^{2,\theta}$ shows that for all $x \in [0, \mathfrak{q}_1^\theta)$ it holds that

$$\mathscr{N}_\infty^{2,\theta}(x) = -\frac{1}{3}. \tag{2.375}$$

This and Proposition 2.16 imply that

$$0 = \int_0^1 (\mathscr{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\, \mathrm{d}x = \int_0^{\mathfrak{q}_1^\theta} (\mathscr{N}_\infty^{2,\theta}(x))\, \mathrm{d}x = \int_0^{\mathfrak{q}_1^\theta} -\frac{1}{3}\, \mathrm{d}x = -\frac{\mathfrak{q}_1^\theta}{3}. \tag{2.376}$$

This is a contradiction. In the next step we establish the contradiction in the case

$$\max\{\mathfrak{w}_1^\theta, \mathfrak{w}_2^\theta\} < 0. \tag{2.377}$$

Note that (2.377), the fact that $\mathcal{G}^2(\theta) = 0$, and Proposition 2.16 imply that for all $i \in \{0, 1\}$ it holds that

$$\int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} x(\mathcal{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} (\mathcal{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.378}$$

This, Lemma 2.7, and Corollary 2.9 demonstrate that for all $x \in [0, \mathfrak{q}_2^\theta]$ it holds that $\mathfrak{q}_1^\theta = 3/4 - 1/2\mathfrak{q}_2^\theta$ and

$$\mathcal{N}_\infty^{2,\theta}(x) = \begin{cases} 0 & : x \in [0, \mathfrak{q}_1^\theta] \\ \frac{16x}{9(2\mathfrak{q}_2^\theta - 1)} + \frac{4(2\mathfrak{q}_2^\theta - 3)}{9(2\mathfrak{q}_2^\theta - 1)} & : x \in (\mathfrak{q}_1^\theta, \mathfrak{q}_2^\theta]. \end{cases} \tag{2.379}$$

Combining this with the fact that $I_1^\theta = [0, \mathfrak{q}_1^\theta)$, the fact that $I_2^\theta = [0, \mathfrak{q}_2^\theta)$, and the fact that $\mathcal{N}_\infty^{2,\theta}$ is continuous shows that for all $x \in (\mathfrak{q}_2^\theta, 1]$ it holds that

$$\mathcal{N}_\infty^{2,\theta}(x) = \frac{4}{3}. \tag{2.380}$$

This and Proposition 2.16 imply that

$$0 = \int_0^1 (\mathcal{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}_2^\theta}^1 (\mathcal{N}_\infty^{2,\theta}(x) - 1)\,\mathrm{d}x = \int_{\mathfrak{q}_2^\theta}^1 \frac{1}{3}\,\mathrm{d}x = \frac{1}{3}(1 - \mathfrak{q}_2^\theta). \tag{2.381}$$

This is a contradiction. The proof of Lemma 2.49 is thus complete. □

Lemma 2.50. *Assume Setting 2.1 and let $\theta \in \mathbb{R}^7$ satisfy $\mathfrak{w}_1^\theta < 0 < \mathfrak{w}_2^\theta$, $\mathfrak{v}_1^\theta \mathfrak{v}_2^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < 1$, and $1/2 \in (\mathfrak{q}_1^\theta, \mathfrak{q}_2^\theta)$. Then*

$$\mathcal{G}^2(\theta) \neq 0. \tag{2.382}$$

Proof of Lemma 2.50. We prove (2.382) by contradiction. Assume that $\mathcal{G}^2(\theta) = 0$. This, the assumption that $\mathfrak{w}_1^\theta < 0 < \mathfrak{w}_2^\theta$, and Proposition 2.16 imply that for all $i \in \{0, 2\}$ it holds that

$$\int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} x(\mathcal{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+1}^\theta} (\mathcal{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.383}$$

This and Lemma 2.7 demonstrate that for all $x \in [0, 1] \backslash (\mathfrak{q}_1^\theta, \mathfrak{q}_2^\theta)$ it holds that

$$\mathcal{N}_\infty^{2,\theta}(x) = \begin{cases} 0 & : x \in [0, \mathfrak{q}_1^\theta] \\ 1 & : x \in [\mathfrak{q}_2^\theta, 1]. \end{cases} \tag{2.384}$$

Combining this with the fact that $I_1^\theta = [0, \mathfrak{q}_1^\theta)$, the fact that $I_2^\theta = (\mathfrak{q}_2^\theta, 1]$, and the fact that $\mathcal{N}_\infty^{2,\theta}$ is continuous shows that for all $x \in (\mathfrak{q}_1^\theta, \mathfrak{q}_2^\theta)$ it holds that $\mathcal{N}_\infty^{2,\theta}(x) = 0$. This is a contradiction. The proof of Lemma 2.50 is thus complete. □

Lemma 2.51. *Assume Setting 2.1 and let $\theta \in (\mathcal{G}^2)^{-1}(\{0\})$ satisfy $\mathfrak{w}_2^\theta < 0 < \mathfrak{w}_1^\theta$, $\mathfrak{v}_1^\theta \mathfrak{v}_2^\theta \neq 0 < \mathfrak{q}_1^\theta < \mathfrak{q}_2^\theta < 1$, and $1/2 \in (\mathfrak{q}_1^\theta, \mathfrak{q}_2^\theta)$. Then*

$$\mathcal{L}_\infty^2(\theta) \geq \frac{1}{864}. \tag{2.385}$$

Proof of Lemma 2.51. Observe that Lemma 2.23, Lemma 2.24, and the fact that $\mathcal{N}_\infty^\theta$ is continuous imply that

$$\mathcal{N}_\infty^{2,\theta}(x) = \begin{cases} \alpha^\theta x - \frac{\alpha^\theta \mathfrak{q}_1^\theta}{2} & : x \in [0, \mathfrak{q}_1^\theta] \\ (\alpha^\theta + \beta^\theta)(x - \mathfrak{q}_1^\theta) + \frac{\alpha^\theta \mathfrak{q}_1^\theta}{2} & : x \in (\mathfrak{q}_1^\theta, \mathfrak{q}_2^\theta] \\ \beta^\theta x + 1 - \frac{\beta^\theta}{2}(1 + \mathfrak{q}_2^\theta) & : x \in (\mathfrak{q}_2^\theta, 1] \end{cases} \tag{2.386}$$

and

$$(\alpha^\theta + \beta^\theta)(\mathfrak{q}_2^\theta - \mathfrak{q}_1^\theta) + \frac{\alpha^\theta \mathfrak{q}_1^\theta}{2} = 1 - \frac{\beta^\theta}{2}(1 - \mathfrak{q}_2^\theta). \tag{2.387}$$

Furthermore, note that the assumption that $\mathfrak{w}_2^\theta < 0 < \mathfrak{w}_1^\theta$ and Proposition 2.16 show that for all $i \in \{0, 1\}$ it holds that

$$\int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+2}^\theta} x(\mathscr{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = \int_{\mathfrak{q}_i^\theta}^{\mathfrak{q}_{i+2}^\theta} (\mathscr{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x = 0. \tag{2.388}$$

Combining this and (2.386) ensures that

$$\begin{aligned} 0 &= \int_0^{\mathfrak{q}_1^\theta} x(\mathscr{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x - \int_{\mathfrak{q}_2^\theta}^1 x(\mathscr{N}_\infty^{2,\theta}(x) - \mathbb{1}_{(1/2,\infty)}(x))\,\mathrm{d}x \\ &= \frac{\alpha^\theta}{12}(\mathfrak{q}_1^\theta)^3 - \frac{\beta^\theta}{12}(1 - \mathfrak{q}_2^\theta)^3. \end{aligned} \tag{2.389}$$

This establishes that $\alpha^\theta \beta^\theta > 0$. Combining this and (2.387) proves that $\min\{\alpha^\theta, \beta^\theta\} > 0$. This and (2.386) imply that for all $x \in [0, 1]$ it holds that $\mathscr{N}_\infty^{2,\theta}(x) \le (\alpha^\theta + \beta^\theta)x$. Therefore, we obtain that

$$1 < \mathscr{N}_\infty^{2,\theta}(1) \le \alpha^\theta + \beta^\theta. \tag{2.390}$$

This shows that $\max\{\alpha^\theta, \beta^\theta\} \ge 1/2$. In the following we distinguish between the case $\alpha^\theta \ge 1/2$ and the case $\beta^\theta \ge 1/2$. We first prove (2.385) in the case

$$\alpha^\theta \ge \frac{1}{2}. \tag{2.391}$$

Observe that Lemma 2.6, (2.386), (2.390), and (2.391) demonstrate that

$$\begin{aligned} \mathcal{L}_\infty^2(\theta) &\ge \int_0^{\mathfrak{q}_1^\theta} (\mathscr{N}_\infty^{2,\theta}(x))^2\,\mathrm{d}x + \int_{\mathfrak{q}_1^\theta}^{\frac{1}{2}} (\mathscr{N}_\infty^{2,\theta}(x))^2\,\mathrm{d}x \\ &\ge \frac{1}{12}(\alpha^\theta)^2(\mathfrak{q}_1^\theta)^3 + \frac{1}{12}(\alpha^\theta + \beta^\theta)^2\left(\frac{1}{2} - \mathfrak{q}_1^\theta\right)^3 \ge \frac{1}{48}(\mathfrak{q}_1^\theta)^3 + \frac{1}{12}\left(\frac{1}{2} - \mathfrak{q}_1^\theta\right)^3 \\ &\ge \frac{1}{48}\left(\frac{1}{3}\right)^3 + \frac{1}{12}\left(\frac{1}{2} - \frac{1}{3}\right)^3 = \frac{1}{864}. \end{aligned} \tag{2.392}$$

This proves (2.385) in the case $\alpha^\theta \ge 1/2$. Next we establish (2.385) in the case

$$\beta^\theta \ge \frac{1}{2}. \tag{2.393}$$

Note that Lemma 2.6, (2.386), (2.390), and (2.393) assure that

$$\begin{aligned} \mathcal{L}_\infty^2(\theta) &\ge \int_{\frac{1}{2}}^{\mathfrak{q}_2^\theta} (\mathscr{N}_\infty^{2,\theta}(x) - 1)^2\,\mathrm{d}x + \int_{\mathfrak{q}_2^\theta}^1 (\mathscr{N}_\infty^{2,\theta}(x) - 1)^2\,\mathrm{d}x \\ &\ge \frac{1}{12}(\alpha^\theta + \beta^\theta)^2\left(\mathfrak{q}_2^\theta - \frac{1}{2}\right)^3 + \frac{1}{12}(\beta^\theta)^2(1 - \mathfrak{q}_2^\theta)^3 \\ &\ge \frac{1}{12}\left(\mathfrak{q}_2^\theta - \frac{1}{2}\right)^3 + \frac{1}{48}(1 - \mathfrak{q}_2^\theta)^3 \ge \frac{1}{12}\left(\frac{2}{3} - \frac{1}{2}\right)^3 + \frac{1}{48}\left(1 - \frac{2}{3}\right)^3 = \frac{1}{864}. \end{aligned} \tag{2.394}$$

This demonstrates (2.385) in the case $\beta^\theta \ge 1/2$. The proof of Lemma 2.51 is thus complete. □

Lemma 2.52. *Assume Setting 2.1. Then it holds for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ that*

$$\mathcal{L}^2_\infty(\theta) \geq \frac{1}{864}. \tag{2.395}$$

Proof of Lemma 2.52. Observe that for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ with $\{k \in \{1,2\}\colon I^\theta_k = \varnothing\} \neq \varnothing$ there exists $\vartheta \in (\mathscr{G}^1)^{-1}(\{0\})$ such that for all $x \in [0,1]$ it holds that $\mathscr{N}^{2,\theta}_\infty(x) = \mathscr{N}^{1,\vartheta}_\infty(x)$. Combining this and Proposition 2.42 shows that for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ with $\{k \in \{1,2\}\colon I^\theta_k = \varnothing\} \neq \varnothing$ there exists $\vartheta \in (\mathscr{G}^1)^{-1}(\{0\})$ such that

$$\mathcal{L}^2_\infty(\theta) = \mathcal{L}^1_\infty(\vartheta) \geq \frac{1}{36}. \tag{2.396}$$

Furthermore, note that for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ with $\mathfrak{v}^\theta_1\mathfrak{v}^\theta_2 = 0$ there exists $\vartheta \in (\mathscr{G}^1)^{-1}(\{0\})$ such that for all $x \in [0,1]$ it holds that $\mathscr{N}^{2,\theta}_\infty(x) = \mathscr{N}^{1,\vartheta}_\infty(x)$. Combining this and Proposition 2.42 establishes that for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ with $\mathfrak{v}^\theta_1\mathfrak{v}^\theta_2 = 0$ there exists $\vartheta \in (\mathscr{G}^1)^{-1}(\{0\})$ such that

$$\mathcal{L}^2_\infty(\theta) = \mathcal{L}^1_\infty(\vartheta) \geq \frac{1}{36}. \tag{2.397}$$

Furthermore, observe that Corollary 2.34 ensures that for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ with $\{k \in \{1,2\}\colon (0,1) \subseteq I^{2,\theta}_k\} \neq \varnothing$ it holds that

$$\mathcal{L}^2_\infty(\theta) \geq \frac{1}{36}. \tag{2.398}$$

Next, note that Corollary 2.35 and Proposition 2.36 assure that for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ with $\mathfrak{v}^\theta_1\mathfrak{v}^\theta_2 \neq 0$, $\mathfrak{q}^\theta_1 = \mathfrak{q}^\theta_2 \in (0,1)$ it holds that

$$\mathcal{L}^2_\infty(\theta) \geq \frac{1}{36}. \tag{2.399}$$

In addition, observe that Lemma 2.37 and Lemma 2.38 demonstrate that for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ with $\mathfrak{v}^\theta_1\mathfrak{v}^\theta_2 \neq 0 < \mathfrak{q}^\theta_1 < \mathfrak{q}^\theta_2 < 1$ and $1/2 \notin (\mathfrak{q}^\theta_1, \mathfrak{q}^\theta_2)$ it holds that

$$\mathcal{L}^2_\infty(\theta) \geq \frac{1}{36}. \tag{2.400}$$

Moreover, note that Lemma 2.49, Lemma 2.50, and Lemma 2.51 prove that for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ with $\mathfrak{v}^\theta_1\mathfrak{v}^\theta_2 \neq 0 < \mathfrak{q}^\theta_1 < \mathfrak{q}^\theta_2 < 1$ and $1/2 \in (\mathfrak{q}^\theta_1, \mathfrak{q}^\theta_2)$ it holds that

$$\mathcal{L}^2_\infty(\theta) \geq \frac{1}{864}. \tag{2.401}$$

Combining this, (2.396), (2.397), (2.398), (2.399), and (2.400) implies that for all $\theta \in (\mathscr{G}^2)^{-1}(\{0\})$ it holds that

$$\mathcal{L}^2_\infty(\theta) \geq \frac{1}{864}. \tag{2.402}$$

The proof of Lemma 2.52 is thus complete. ☐

Lemma 2.53. *Let $\mathbb{A}_r \in C(\mathbb{R},\mathbb{R})$, $r \in \mathbb{N}\cup\{\infty\}$, satisfy for all $x \in \mathbb{R}$ that $(\bigcup_{r\in\mathbb{N}}\{\mathbb{A}_r\}) \subseteq C^1(\mathbb{R},\mathbb{R})$, $\mathbb{A}_\infty(x) = \max\{x,0\}$, $\sup_{r\in\mathbb{N}}\sup_{y\in[-|x|,|x|]}|(\mathbb{A}_r)'(y)| < \infty$, and*

$$\limsup_{r\to\infty}\big(|\mathbb{A}_r(x) - \mathbb{A}_\infty(x)| + |(\mathbb{A}_r)'(x) - \mathbb{1}_{(0,\infty)}(x)|\big) = 0, \tag{2.403}$$

let $\mathcal{L}_r\colon \mathbb{R}^7 \to \mathbb{R}$, $r \in \mathbb{N}\cup\{\infty\}$, satisfy for all $r \in \mathbb{N}\cup\{\infty\}$, $\theta = (\theta_1,\dots,\theta_7) \in \mathbb{R}^7$ that

$$\mathcal{L}_r(\theta) = \int_0^1 \big(\mathbb{1}_{(1/2,\infty)}(x) - \theta_7 - \textstyle\sum_{i=1}^2 \theta_{4+i}\big[\mathbb{A}_r(\theta_{2+i} + \theta_i x)\big]\big)^2\,\mathrm{d}x, \tag{2.404}$$

and let $\mathcal{G}\colon \mathbb{R}^7 \to \mathbb{R}^7$ satisfy for all $\theta \in \{\vartheta \in \mathbb{R}^7\colon ((\nabla\mathcal{L}_r)(\vartheta))_{r\in\mathbb{N}}$ is convergent$\}$ that $\mathcal{G}(\theta) = \lim_{r\to\infty}(\nabla\mathcal{L}_r)(\theta)$. Then it holds for all $\Theta \in C([0,\infty),\mathbb{R}^7)$ with $\forall\, t \in [0,\infty)\colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}(\Theta_s)\,\mathrm{d}s$ and $\mathcal{L}_\infty(\Theta_0) < 1/864$ that $\liminf_{t\to\infty}\|\Theta_t\| = \infty$.

Proof of Lemma 2.53. Observe that Lemma 2.52 demonstrates that for all $\theta \in \mathcal{G}^{-1}(\{0\})$ it holds that

$$\mathcal{L}_\infty(\theta) \geq \frac{1}{864}. \tag{2.405}$$

Furthermore, note that, e.g., [23, Lemma 3.1] implies that $[0,\infty) \ni t \mapsto \mathcal{L}_\infty(\Theta_t) \in \mathbb{R}$ is non-increasing. Combining this with (2.405) and Proposition 2.45 assures that for all $\Theta \in C([0,\infty),\mathbb{R}^7)$ with $\forall\, t \in [0,\infty)\colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}(\Theta_s)\,\mathrm{d}s$ and $\mathcal{L}_\infty(\Theta_0) < 1/864$ it holds that $\liminf_{t\to\infty}\|\Theta_t\| = \infty$. The proof of Lemma 2.53 is thus complete. □

2.8 Upper bounds for GFs

Proposition 2.54. *Let $\mathfrak{d} \in \mathbb{N}$, $\Theta \in C([0,\infty),\mathbb{R}^{\mathfrak{d}})$, let $\mathcal{L}\colon \mathbb{R}^{\mathfrak{d}} \to [0,\infty)$ and $\mathcal{G}\colon \mathbb{R}^{\mathfrak{d}} \to \mathbb{R}^{\mathfrak{d}}$ be measurable, and assume for all $t \in [0,\infty)$ that $\mathcal{L}(\Theta_t) = \mathcal{L}(\Theta_0) - \int_0^t \|\mathcal{G}(\Theta_s)\|^2\,\mathrm{d}s$ and $\Theta_t = \Theta_0 - \int_0^t \mathcal{G}(\Theta_s)\,\mathrm{d}s$. Then it holds for all $t \in [0,\infty)$ that*

$$\|\Theta_t\| \leq \|\Theta_0\| + t^{1/2}|\mathcal{L}(\Theta_0) - \mathcal{L}(\Theta_t)|^{1/2} \leq \|\Theta_0\| + [t\mathcal{L}(\Theta_0)]^{1/2}. \tag{2.406}$$

Proof of Proposition 2.54. Observe that the assumption that for all $t \in [0,\infty)$ it holds that $\Theta_t = \Theta_0 - \int_0^t \mathcal{G}(\Theta_s)\,\mathrm{d}s$, the triangle inequality, and the Cauchy-Schwarz inequality ensure for all $t \in [0,\infty)$ that

$$\|\Theta_t\| \leq \|\Theta_0\| + \int_0^t \|\mathcal{G}(\Theta_s)\|\,\mathrm{d}s \leq \|\Theta_0\| + t^{1/2}\left[\int_0^t \|\mathcal{G}(\Theta_s)\|^2\,\mathrm{d}s\right]^{1/2}. \tag{2.407}$$

Hence, we obtain that for all $t \in [0,\infty)$ it holds that

$$\|\Theta_t\| \leq \|\Theta_0\| + t^{1/2}|\mathcal{L}(\Theta_0) - \mathcal{L}(\Theta_t)|^{1/2} \leq \|\Theta_0\| + [t\mathcal{L}(\Theta_0)]^{1/2}. \tag{2.408}$$

The proof of Proposition 2.54 is thus complete. □

3 Non-existence of global minima of the risk and divergence of GFs and gradient descent (GD) for widely used activation functions

In this section we establish, in the case of at least two neurons on the hidden layer, the non-existence of global minima employing various activation and target functions. Next we show the blow up of GFs under a specific asymptotic optimality assumption regarding the risk values. The key idea is to prove that there exists a sequence of ANN parameters such that the risk converges to zero and to consequently demonstrate that the set of global minima is empty. After proving this, Corollary 3.29, Corollary 3.30, and Corollary 3.31 in Subsection 3.8 assure the divergence of GFs under the assumption that the risk of GFs converges to the infimum of the risk while Corollary 3.33, Corollary 3.35, and Corollary 3.37 in Subsection 3.8 prove the corresponding result in the discrete-time case. Related results can be found in [30, Proposition 3.6]. Corollary 3.29, Corollary 3.30, and Corollary 3.31 are based on Lemma 3.28, demonstrated using compactness and continuity properties, and the well-known deterministic Itô-type formula for continuously differentiable functions, see, e.g., [9, Lemma 3.1]. Corollary 3.33, Corollary 3.35, and Corollary 3.37 follow from Lemma 3.32.

Choosing the square function as target function we establish the non-existence of global minima in the case of softplus activation in Lemma 3.11 in Subsection 3.3, in the case of standard logistic, hyperbolic tangent, arctangent, and inverse square root unit activation in Lemma 3.21 in Subsection 3.4, in the case of exponential linear unit activation in Lemma 3.25 in Subsection 3.6, and in the case of softsign activation in Lemma 3.27 in Subsection 3.7. The proofs of Lemma 3.11 and Lemma 3.21 employ properties of real analytic functions and are inspired by [30, Theorem 3.3]. The proofs of Lemma 3.25 and Lemma 3.27 instead use a comparison between target and realization function derivatives.

Employing an indicator function as target function we demonstrate the non-existence of global minima in the case of ReLU and leaky ReLU activation function in Lemma 3.6 in Subsection 3.2. Its proof uses Proposition 3.2, Proposition 3.3, Lipschitz continuity results in Lemma 3.4, and elementary properties of realization functions in Lemma 3.5.

In the case of rectified power unit activation we instead establish in Lemma 3.23 in Subsection 3.5 the non-existence of global minima using as target function the rectified power unit itself with a smaller exponent. Ingredients employed in the proof are Proposition 3.22 and a continuity study inspired by [30, Theorem 3.3].

Also notably, we show the non-existence of global minima for every number of hidden neurons in Lemma 3.9 in Subsection 3.3 employing the ReLU function as target function and the softplus activation function and in Lemma 3.20 in Subsection 3.4 using the identity function as target function and standard logistic, arctangent, and inverse square root unit activation. The proofs of Lemma 3.9 and Lemma 3.20 are inspired by [30, Theorem 3.3].

3.1 Mathematical description of ANNs

Setting 3.1. *Let $\mathscr{a} \in \mathbb{R}$, $\mathscr{b} \in (\mathscr{a}, \infty)$, $\xi \in (0, \infty)$, $h, \mathfrak{d} \in \mathbb{N}$ satisfy $\mathfrak{d} = 3h + 1$, let $\mathfrak{w} = ((\mathfrak{w}^\theta_1, \dots, \mathfrak{w}^\theta_h))_{\theta \in \mathbb{R}^\mathfrak{d}} \colon \mathbb{R}^\mathfrak{d} \to \mathbb{R}^h$, $\mathfrak{b} = ((\mathfrak{b}^\theta_1, \dots, \mathfrak{b}^\theta_h))_{\theta \in \mathbb{R}^\mathfrak{d}} \colon \mathbb{R}^\mathfrak{d} \to \mathbb{R}^h$, $\mathfrak{v} = ((\mathfrak{v}^\theta_1, \dots, \mathfrak{v}^\theta_h))_{\theta \in \mathbb{R}^\mathfrak{d}} \colon \mathbb{R}^\mathfrak{d} \to \mathbb{R}^h$, and $\mathfrak{c} = (\mathfrak{c}^\theta)_{\theta \in \mathbb{R}^\mathfrak{d}} \colon \mathbb{R}^\mathfrak{d} \to \mathbb{R}$ satisfy for all $\theta = (\theta_1, \dots, \theta_\mathfrak{d}) \in \mathbb{R}^\mathfrak{d}$, $j \in \{1, 2, \dots, h\}$ that $\mathfrak{w}^\theta_j = \theta_j$, $\mathfrak{b}^\theta_j = \theta_{h+j}$, $\mathfrak{v}^\theta_j = \theta_{2h+j}$, and $\mathfrak{c}^\theta = \theta_\mathfrak{d}$, for every $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ let $A_{k,\gamma} \colon \mathbb{R} \to \mathbb{R}$ satisfy for all $x \in \mathbb{R}$ that*

$$A_{k,\gamma}(x) = \begin{cases} x(1 + |x|)^{-1} & : k < -5 \\ \arctan(x) & : k = -5 \\ x(1 + \xi x^2)^{-1/2} & : k = -4 \\ x \mathbb{1}_{(0,\infty)}(x) + (\exp(x) - 1) \mathbb{1}_{(-\infty,0]}(x) & : k = -3 \\ (\exp(x) - \exp(-x))(\exp(x) + \exp(-x))^{-1} & : k = -2 \\ (1 + \exp(-x))^{-1} & : k = -1 \\ \ln(1 + \exp(x)) & : k = 0 \\ (\max\{x, 0\})^k + \min\{\gamma x, 0\} & : k > 0 \end{cases} \tag{3.1}$$

and let $\mathbb{A}^r_{k,\gamma} \in C(\mathbb{R}, \mathbb{R})$, $r \in \mathbb{N} \cup \{\infty\}$, satisfy for all $x \in \mathbb{R}$ that $(\bigcup_{r \in \mathbb{N}} \{\mathbb{A}^r_{k,\gamma}\}) \subseteq C^1(\mathbb{R}, \mathbb{R})$, $\mathbb{A}^\infty_{k,\gamma}(x) = A_{k,\gamma}(x)$, $\sup_{r \in \mathbb{N}} \sup_{y \in [-|x|, |x|]} |(\mathbb{A}^r_{k,\gamma})'(y)| < \infty$, and

$$\limsup_{r \to \infty} \left(|\mathbb{A}^r_{k,\gamma}(x) - \mathbb{A}^\infty_{k,\gamma}(x)| + \left|(\mathbb{A}^r_{k,\gamma})'(x) - \lim_{h \nearrow 0} \tfrac{\mathbb{A}^\infty_{k,\gamma}(x+h) - \mathbb{A}^\infty_{k,\gamma}(x)}{h}\right| \right) = 0, \tag{3.2}$$

for every $\theta \in \mathbb{R}^\mathfrak{d}$, $r \in \mathbb{N} \cup \{\infty\}$, $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ let $\mathscr{N}^{\theta,r}_{k,\gamma} \colon \mathbb{R} \to \mathbb{R}$ satisfy for all $x \in \mathbb{R}$ that

$$\mathscr{N}^{\theta,r}_{k,\gamma}(x) = \mathfrak{c}^\theta + \sum_{i=1}^h \mathfrak{v}^\theta_i \big[\mathbb{A}^r_{k,\gamma}(\mathfrak{w}^\theta_i x + \mathfrak{b}^\theta_i)\big], \tag{3.3}$$

let $f \colon \mathbb{R} \to \mathbb{R}$ be measurable, for every $r \in \mathbb{N} \cup \{\infty\}$, $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ let $\mathcal{L}^r_{k,\gamma} \colon \mathbb{R}^\mathfrak{d} \to \mathbb{R}$ satisfy for all $\theta \in \mathbb{R}^\mathfrak{d}$ that

$$\mathcal{L}^r_{k,\gamma}(\theta) = \int_\mathscr{a}^\mathscr{b} \big(f(x) - \mathscr{N}^{\theta,r}_{k,\gamma}(x)\big)^2 \, \mathrm{d}x, \tag{3.4}$$

and for every $k \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ let $\mathcal{G}_{k,\gamma} \colon \mathbb{R}^\mathfrak{d} \to \mathbb{R}^\mathfrak{d}$ satisfy for all $\theta \in \{\vartheta \in \mathbb{R}^\mathfrak{d} \colon ((\nabla \mathcal{L}^r_{k,\gamma})(\vartheta))_{r \in \mathbb{N}}$ is convergent$\}$ that $\mathcal{G}_{k,\gamma}(\theta) = \lim_{r \to \infty} (\nabla \mathcal{L}^r_{k,\gamma})(\theta)$.

3.2 ANNs with ReLU and leaky ReLU activation

Proposition 3.2. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in \mathbb{R}$ that $f(x) = \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)$, let $\gamma \in (-\infty, 0]$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = \mathfrak{w}_2^{\theta_n} = (n+3)(2\mathscr{b}-2\mathscr{a})^{-1}$, $\mathfrak{b}_1^{\theta_n} = -(1+n)4^{-1}-\mathscr{a}(n+3)(2\mathscr{b}-2\mathscr{a})^{-1}$, $\mathfrak{b}_2^{\theta_n} = -(5+n)4^{-1}-\mathscr{a}(n+3)(2\mathscr{b}-2\mathscr{a})^{-1}$, $\mathfrak{v}_1^{\theta_n} = -\mathfrak{v}_2^{\theta_n} = (1-\gamma)^{-1}$, and $|\mathfrak{c}^{\theta_n}| + \sum_{j=3}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup\nolimits_{n\to\infty} \mathcal{L}_{1,\gamma}^{\infty}(\theta_n) = 0. \tag{3.5}$$

Proof of Proposition 3.2. Note that (3.3) ensures that for all $n \in \mathbb{N}$, $x \in [\mathscr{a}, \mathscr{b}]$ it holds that

$$\begin{aligned}\mathscr{N}_{1,\gamma}^{\theta_n,\infty}(x) = \max&\left\{-\frac{1+n}{4} - \frac{\mathscr{a}(n+3)}{2(\mathscr{b}-\mathscr{a})} + \frac{n+3}{2(\mathscr{b}-\mathscr{a})}x, 0\right\}\\ &- \max\left\{-\frac{5+n}{4} - \frac{\mathscr{a}(n+3)}{2(\mathscr{b}-\mathscr{a})} + \frac{n+3}{2(\mathscr{b}-\mathscr{a})}x, 0\right\}.\end{aligned} \tag{3.6}$$

This implies that for all $n \in \mathbb{N}$ it holds that

$$\begin{aligned}\mathcal{L}_{1,\gamma}^{\infty}(\theta_n) &= \int_{\mathscr{a}}^{\mathscr{b}} (\mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x) - \mathscr{N}_{1,\gamma}^{\theta_n,\infty}(x))^2 \,\mathrm{d}x\\ &= \int_{\mathscr{a}+\frac{(\mathscr{b}-\mathscr{a})(1+n)}{2(n+3)}}^{\frac{\mathscr{a}+\mathscr{b}}{2}} \left(-\frac{1+n}{4} - \frac{\mathscr{a}(n+3)}{2(\mathscr{b}-\mathscr{a})} + \frac{n+3}{2(\mathscr{b}-\mathscr{a})}x\right)^2 \mathrm{d}x\\ &\quad+ \int_{\frac{\mathscr{a}+\mathscr{b}}{2}}^{\mathscr{a}+\frac{(\mathscr{b}-\mathscr{a})(5+n)}{2(n+3)}} \left(-\frac{5+n}{4} - \frac{\mathscr{a}(n+3)}{2(\mathscr{b}-\mathscr{a})} + \frac{n+3}{2(\mathscr{b}-\mathscr{a})}x\right)^2 \mathrm{d}x\\ &\le \int_{\mathscr{a}+\frac{(\mathscr{b}-\mathscr{a})(1+n)}{2(n+3)}}^{\frac{\mathscr{a}+\mathscr{b}}{2}} \left(\frac{1}{2}\right)^2 \mathrm{d}x + \int_{\frac{\mathscr{a}+\mathscr{b}}{2}}^{\mathscr{a}+\frac{(\mathscr{b}-\mathscr{a})(5+n)}{2(n+3)}} \left(\frac{1}{2}\right)^2 \mathrm{d}x = \frac{\mathscr{b}-\mathscr{a}}{2(n+3)}.\end{aligned} \tag{3.7}$$

Therefore, we obtain that

$$\limsup\nolimits_{n\to\infty} \mathcal{L}_{1,\gamma}^{\infty}(\theta_n) = 0. \tag{3.8}$$

The proof of Proposition 3.2 is thus complete. □

Proposition 3.3. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in \mathbb{R}$ that $f(x) = \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)$, let $\gamma \in (0, \infty)\backslash\{1\}$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = \mathfrak{w}_2^{\theta_n} = (n+3)(2\mathscr{b}-2\mathscr{a})^{-1}$, $\mathfrak{b}_1^{\theta_n} = -(1+n)4^{-1} - \mathscr{a}(n+3)(2\mathscr{b}-2\mathscr{a})^{-1}$, $\mathfrak{b}_2^{\theta_n} = -(5+n)4^{-1} - \mathscr{a}(n+3)(2\mathscr{b}-2\mathscr{a})^{-1}$, $\mathfrak{v}_1^{\theta_n} = -\mathfrak{v}_2^{\theta_n} = (1-\gamma)^{-1}$, $\mathfrak{c}^{\theta_n} = -\gamma(1-\gamma)^{-1}$, and $\sum_{j=3}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup\nolimits_{n\to\infty} \mathcal{L}_{1,\gamma}^{\infty}(\theta_n) = 0. \tag{3.9}$$

Proof of Proposition 3.3. Observe that (3.3) ensures that for all $n \in \mathbb{N}$, $x \in \mathbb{R}$ it holds that

$$\mathscr{N}_{1,\gamma}^{\theta_n,\infty}(x) = \begin{cases} 0 & : x \in [\mathscr{a}, \mathscr{a} + \frac{(\mathscr{b}-\mathscr{a})(1+n)}{2(n+3)}] \\ -\frac{1+n}{4} - \frac{\mathscr{a}(n+3)}{2(\mathscr{b}-\mathscr{a})} + \frac{n+3}{2(\mathscr{b}-\mathscr{a})}x. & : x \in (\mathscr{a} + \frac{(\mathscr{b}-\mathscr{a})(1+n)}{2(n+3)}, \mathscr{a} + \frac{(\mathscr{b}-\mathscr{a})(5+n)}{2(n+3)}] \\ 1 & : x \in (\mathscr{a} + \frac{(\mathscr{b}-\mathscr{a})(5+n)}{2(n+3)}, \mathscr{b}]. \end{cases} \tag{3.10}$$

This implies that for all $x \in \mathbb{R}$ it holds that

$$\limsup\nolimits_{n\to\infty} |\mathscr{N}_{1,\gamma}^{\theta_n,\infty}(x) - \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)| = 0. \tag{3.11}$$

Furthermore, note that (3.10) assures that for all $n \in \mathbb{N}$, $x \in [\mathscr{a}, \mathscr{b}]$ it holds that $|\mathscr{N}_{1,\gamma}^{\theta_n}(x) - f(x)| \le 1$. Combining this with (3.11) and Lebesgue's dominated convergence theorem demonstrates that

$$\limsup\nolimits_{n\to\infty} \mathcal{L}_{1,\gamma}^{\infty}(\theta_n) = 0. \tag{3.12}$$

The proof of Proposition 3.3 is thus complete. □

Lemma 3.4. *Let $\mathscr{a} \in \mathbb{R}$, $\mathscr{b} \in (\mathscr{a}, \infty)$, let $f\colon [\mathscr{a}, \mathscr{b}] \to \mathbb{R}$ satisfy for all $x \in [\mathscr{a}, \mathscr{b}]$ that $f(x) = \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2, \infty)}(x)$, and let $L \in \mathbb{R}$, $g \in C([\mathscr{a}, \mathscr{b}], \mathbb{R})$ satisfy for all $x, y \in [\mathscr{a}, \mathscr{b}]$ that $|g(x) - g(y)| \leq L|x - y|$. Then*

$$\int_{\mathscr{a}}^{\mathscr{b}} (f(x) - g(x))^2 \, \mathrm{d}x \geq \frac{1}{32 \max\{L, 1/(\mathscr{b}-\mathscr{a})\}}. \tag{3.13}$$

Proof of Lemma 3.4. Throughout this proof let $c \in \mathbb{R}$ satisfy

$$c = \max\left\{L, 1/(\mathscr{b}-\mathscr{a})\right\}. \tag{3.14}$$

Observe that (3.14) assures that for all $x, y \in [\mathscr{a}, \mathscr{b}]$ it holds that

$$|g(x) - g(y)| \leq c|x - y|. \tag{3.15}$$

In the following we distinguish between the case $g((\mathscr{a}+\mathscr{b})/2) > 1/2$ and the case $g((\mathscr{a}+\mathscr{b})/2) \leq 1/2$. We first prove (3.13) in the case

$$g\left(\frac{\mathscr{a}+\mathscr{b}}{2}\right) > \frac{1}{2}. \tag{3.16}$$

Note that (3.15) and (3.16) imply that for all $x \in [(\mathscr{a}+\mathscr{b})/2 - 1/2c, (\mathscr{a}+\mathscr{b})/2]$ it holds that

$$0 \leq \frac{1}{2} - c\left(\frac{\mathscr{a}+\mathscr{b}}{2} - x\right) < g\left(\frac{\mathscr{a}+\mathscr{b}}{2}\right) - c\left(\frac{\mathscr{a}+\mathscr{b}}{2} - x\right) \leq g(x). \tag{3.17}$$

This proves that

$$\int_{\frac{\mathscr{a}+\mathscr{b}}{2} - \frac{1}{2c}}^{\frac{\mathscr{a}+\mathscr{b}}{2}} |g(x)| \, \mathrm{d}x \geq \int_{\frac{\mathscr{a}+\mathscr{b}}{2} - \frac{1}{2c}}^{\frac{\mathscr{a}+\mathscr{b}}{2}} \frac{1}{2}(1 - c(\mathscr{a}+\mathscr{b})) + cx \, \mathrm{d}x = \frac{1}{8c}. \tag{3.18}$$

Combining this and the Cauchy-Schwarz inequality demonstrates that

$$\begin{aligned} \int_{\mathscr{a}}^{\mathscr{b}} |f(x) - g(x)|^2 \, \mathrm{d}x &\geq \int_{\frac{\mathscr{a}+\mathscr{b}}{2} - \frac{1}{2c}}^{\frac{\mathscr{a}+\mathscr{b}}{2}} |f(x) - g(x)|^2 \, \mathrm{d}x \\ &\geq 2c\left(\int_{\frac{\mathscr{a}+\mathscr{b}}{2} - \frac{1}{2c}}^{\frac{\mathscr{a}+\mathscr{b}}{2}} |g(x)| \, \mathrm{d}x\right)^2 = \frac{1}{32c}. \end{aligned} \tag{3.19}$$

This establishes (3.13) in the case $g((\mathscr{a}+\mathscr{b})/2) > 1/2$. In the next step we prove (3.13) in the case

$$g\left(\frac{\mathscr{a}+\mathscr{b}}{2}\right) \leq \frac{1}{2}. \tag{3.20}$$

Observe that (3.15) and (3.20) imply that for all $x \in [(\mathscr{a}+\mathscr{b})/2, (\mathscr{a}+\mathscr{b})/2 + 1/2c]$ it holds that

$$g(x) \leq c\left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right) + g\left(\frac{\mathscr{a}+\mathscr{b}}{2}\right) \leq c\left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right) + \frac{1}{2} \leq 1. \tag{3.21}$$

This proves that

$$\int_{\frac{\mathscr{a}+\mathscr{b}}{2}}^{\frac{\mathscr{a}+\mathscr{b}}{2} + \frac{1}{2c}} |f(x) - g(x)| \, \mathrm{d}x \geq \int_{\frac{\mathscr{a}+\mathscr{b}}{2}}^{\frac{\mathscr{a}+\mathscr{b}}{2} + \frac{1}{2c}} 1 - \frac{1}{2}(1 - c(\mathscr{a}+\mathscr{b})) - cx \, \mathrm{d}x = \frac{1}{8c}. \tag{3.22}$$

Combining this and the Cauchy-Schwarz inequality demonstrates that

$$\begin{aligned} \int_{\mathscr{a}}^{\mathscr{b}} |f(x) - g(x)|^2 \, \mathrm{d}x &\geq \int_{\frac{\mathscr{a}+\mathscr{b}}{2}}^{\frac{\mathscr{a}+\mathscr{b}}{2} + \frac{1}{2c}} |f(x) - g(x)|^2 \, \mathrm{d}x \\ &\geq 2c\left(\int_{\frac{\mathscr{a}+\mathscr{b}}{2}}^{\frac{\mathscr{a}+\mathscr{b}}{2} + \frac{1}{2c}} |1 - g(x)| \, \mathrm{d}x\right)^2 = \frac{1}{32c}. \end{aligned} \tag{3.23}$$

This establishes (3.13) in the case $g\left((\mathscr{a}+\mathscr{b})/2\right) \leq 1/2$ The proof of Lemma 3.4 is thus complete. ☐

Lemma 3.5. *Assume Setting 3.1 and let $\theta \in \mathbb{R}^{\mathfrak{d}}$, $\gamma \in \mathbb{R}\backslash\{1\}$. Then*

$$\sup_{x,y\in[\mathscr{a},\mathscr{b}],\, x\neq y}(|\mathscr{N}^{\theta,\infty}_{1,\gamma}(x) - \mathscr{N}^{\theta,\infty}_{1,\gamma}(y)||x-y|^{-1}) \leq (1+|\gamma|)\|\theta\|^2. \tag{3.24}$$

Proof of Lemma 3.5. Note that the fact that $\mathscr{N}^{\theta,\infty}_{1,\gamma}$ is continuous and the fact that $\mathscr{N}^{\theta,\infty}_{1,\gamma}$ is piecewise affine linear imply for all $x, y \in [\mathscr{a}, \mathscr{b}]$ that

$$|\mathscr{N}^{\theta,\infty}_{1,\gamma}(x) - \mathscr{N}^{\theta,\infty}_{1,\gamma}(y)| \leq \max\{1,\gamma\}\left(\textstyle\sum_{i=1}^{H} |\mathfrak{w}^{\theta}_i \mathfrak{v}^{\theta}_i|\right)|x-y| \leq (1+|\gamma|)\|\theta\|^2|x-y|. \tag{3.25}$$

This demonstrates that

$$\sup_{x,y\in[\mathscr{a},\mathscr{b}],\, x\neq y}(|\mathscr{N}^{\theta,\infty}_{1,\gamma}(x) - \mathscr{N}^{\theta,\infty}_{1,\gamma}(y)||x-y|^{-1}) \leq (1+|\gamma|)\|\theta\|^2. \tag{3.26}$$

The proof of Lemma 3.5 is thus complete. ☐

Lemma 3.6. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in \mathbb{R}$ that $f(x) = \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)$, and let $\gamma \in \mathbb{R}\backslash\{1\}$. Then*

$$\left\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}^{\infty}_{1,\gamma}(\vartheta) = \inf\nolimits_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{1,\gamma}(\theta)\right\} = \varnothing. \tag{3.27}$$

Proof of Lemma 3.6. Observe that Proposition 3.2 and Proposition 3.3 imply that $\inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{1,\gamma}(\theta) = 0$. Furthermore, note that Lemma 3.4 and Lemma 3.5 ensure for all $\theta \in \mathbb{R}^{\mathfrak{d}}$ that

$$\mathcal{L}^{\infty}_{1,\gamma}(\theta) \geq \frac{1}{32\max\{(1+|\gamma|)\|\theta\|^2, 1/(\mathscr{b}-\mathscr{a})\}}. \tag{3.28}$$

This implies that $\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}^{\infty}_{1,\gamma}(\vartheta) = 0\} = \varnothing$. The proof of Lemma 3.6 is thus complete. ☐

3.3 ANNs with softplus activation

Lemma 3.7. *Assume Setting 3.1 and let $r \in (0,\infty)$, $\theta \in \mathbb{R}^4$ satisfy $\mathfrak{c}^{\theta} = 0$, $\mathfrak{b}^{\theta}_1 = -r(\mathscr{a}+\mathscr{b})2^{-1}$, $\mathfrak{w}^{\theta}_1 = r$, and $\mathfrak{v}^{\theta}_1 = 1/r$. Then it holds for all $x \in \mathbb{R}$ that $0 \leq \mathscr{N}^{\theta,\infty}_{0,0}(x) - \max\{x - (\mathscr{a}+\mathscr{b})/2, 0\} \leq 1/r$.*

Proof of Lemma 3.7. Observe that (3.3) ensures that for all $x \in \mathbb{R}$ it holds that

$$\mathscr{N}^{\theta,\infty}_{0,0}(x) = \frac{1}{r}\ln\left(1+\exp\left(rx - r\frac{\mathscr{a}+\mathscr{b}}{2}\right)\right). \tag{3.29}$$

Furthermore, note that for all $x \in [0,\infty)$ it holds that

$$x \leq \ln(1+\exp(x)) \leq x+1. \tag{3.30}$$

Combining this and (3.29) establishes for all $x \in [(\mathscr{a}+\mathscr{b})/2, \infty)$ that

$$0 \leq \mathscr{N}^{\theta,\infty}_{0,0}(x) - x + \frac{\mathscr{a}+\mathscr{b}}{2} \leq \frac{1}{r}. \tag{3.31}$$

Moreover, observe that for all $x \in (-\infty, 0]$ it holds that

$$0 \leq \ln(1+\exp(rx)) \leq \exp(rx) \leq 1. \tag{3.32}$$

Combining this and (3.29) implies for all $x \in (-\infty, (\mathscr{a}+\mathscr{b})/2]$ that $0 \leq \mathscr{N}^{\theta,\infty}_{0,0}(x) \leq 1/r$. This and (3.31) show that for all $x \in \mathbb{R}$ it holds that $0 \leq \mathscr{N}^{\theta,\infty}_{0,0}(x) - \max\{x - (\mathscr{a}+\mathscr{b})/2, 0\} \leq 1/r$. The proof of Lemma 3.7 is thus complete. ☐

Proposition 3.8. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in \mathbb{R}$ that $f(x) = x^2$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = -\mathfrak{w}_2^{\theta_n} = 1/n$, $\mathfrak{b}_1^{\theta_n} = \mathfrak{b}_2^{\theta_n} = 0$, $\mathfrak{v}_1^{\theta_n} = \mathfrak{v}_2^{\theta_n} = 4n^2$, $\mathfrak{c}^{\theta_n} = -8n^2\ln(2)$, and $\sum_{j=3}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup\nolimits_{n\to\infty} \mathcal{L}_{0,0}^{\infty}(\theta_n) = 0. \tag{3.33}$$

Proof of Proposition 3.8. Note that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}_{0,0}^{\theta_n,\infty}(x) = 4n^2 \ln\left(1 + \exp\left(\frac{x}{n}\right)\right) + 4n^2 \ln\left(1 + \exp\left(-\frac{x}{n}\right)\right) - 8n^2\ln(2). \tag{3.34}$$

This and the fact that there exists $c \in (0,\infty)$ such that for all $x \in [-1,1]$ it holds that $|\ln(1+\exp(x)) - \ln(2) - x/2 - x^2/8| \le c|x^3|$ assure that there exist $c, M \in (0,\infty)$ such that for all $x \in [\mathscr{a}, \mathscr{b}]$, $n \ge M$ it holds that

$$|\mathscr{N}_{0,0}^{\theta_n,\infty}(x) - x^2| \le c\left|\frac{1}{n^3}\right|. \tag{3.35}$$

This and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup\nolimits_{n\to\infty} \mathcal{L}_{0,0}^{\infty}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup\nolimits_{n\to\infty} \big(x^2 - \mathscr{N}_{0,0}^{\theta_n,\infty}(x)\big)^2 \,\mathrm{d}x = 0. \tag{3.36}$$

The proof of Proposition 3.8 is thus complete. □

Lemma 3.9. *Assume Setting 3.1 and assume for all $x \in \mathbb{R}$ that $f(x) = \max\{x - (\mathscr{a}+\mathscr{b})/2, 0\}$. Then*

$$\big\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}_{0,0}^{\infty}(\vartheta) = \inf\nolimits_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{0,0}^{\infty}(\theta)\big\} = \varnothing. \tag{3.37}$$

Proof of Lemma 3.9. Observe that Lemma 3.7 proves that for every $r \in (0,\infty)$ there exists $\vartheta_r \in \mathbb{R}^{\mathfrak{d}}$ which satisfies for all $x \in \mathbb{R}$ that $|\mathscr{N}_{0,0}^{\vartheta_r,\infty}(x) - \max\{x - (\mathscr{a}+\mathscr{b})/2, 0\}| \le 1/r$. This implies that for every $r \in (0,\infty)$ there exists $\vartheta_r \in \mathbb{R}^{\mathfrak{d}}$ such that $\mathcal{L}_{0,0}^{\infty}(\vartheta_r) \le (\mathscr{b}-\mathscr{a})/r^2$. Hence, we obtain that $\inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{0,0}^{\infty}(\theta) = 0$. Furthermore, note that for all $\theta \in \mathbb{R}^{\mathfrak{d}}$ it holds that $\mathscr{N}_{0,0}^{\theta,\infty} \in C^{\infty}(\mathbb{R},\mathbb{R})$. We prove (3.37) by contradiction. Assume that there exists $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ which satisfies that

$$\mathcal{L}_{0,0}^{\infty}(\vartheta) = 0. \tag{3.38}$$

Observe that (3.38) ensures that for all $x \in [\mathscr{a}, \mathscr{b}]$ it holds that $\mathscr{N}_{0,0}^{\vartheta,\infty}(x) = \max\{x - (\mathscr{a}+\mathscr{b})/2, 0\}$. This demonstrates that $\mathscr{N}_{0,0}^{\vartheta,\infty} \in C([\mathscr{a},\mathscr{b}],\mathbb{R}) \backslash C^1([\mathscr{a},\mathscr{b}],\mathbb{R})$ which is a contradiction. The proof of Lemma 3.9 is thus complete. □

Lemma 3.10. *Assume Setting 3.1, assume $h > 1$, and assume for all $x \in \mathbb{R}$ that $f(x) = x$. Then*

$$\big\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}_{0,0}^{\infty}(\vartheta) = \inf\nolimits_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{0,0}^{\infty}(\theta)\big\} \neq \varnothing. \tag{3.39}$$

Proof of Lemma 3.10. Let $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ satisfy for all $i \in \{1,2,\ldots,h\}\backslash\{1,2\}$ that $\mathfrak{w}_1^{\vartheta} = -\mathfrak{w}_2^{\vartheta} = \mathfrak{v}_1^{\vartheta} = -\mathfrak{v}_2^{\vartheta} = 1$ and $\mathfrak{b}_1^{\vartheta} = \mathfrak{b}_2^{\vartheta} = \mathfrak{w}_i^{\vartheta} = \mathfrak{v}_i^{\vartheta} = \mathfrak{b}_i^{\vartheta} = \mathfrak{c}^{\vartheta} = 0$. This proves that for all $x \in \mathbb{R}$ it holds that

$$\mathscr{N}_{0,0}^{\vartheta,\infty}(x) = \ln(1+\exp(x)) - \ln(1+\exp(-x)) = \ln\left(\frac{1+\exp(x)}{1+\exp(-x)}\right) = \ln(\exp(x)) = x. \tag{3.40}$$

Therefore, we obtain that $\mathcal{L}_{0,0}^{\infty}(\vartheta) = 0$. This and the fact that for all $\theta \in \mathbb{R}^{\mathfrak{d}}$ it holds that $\mathcal{L}_{0,0}^{\infty}(\theta) \ge 0$ demonstrates that $\vartheta \in \{v \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}_{0,0}^{\infty}(v) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{0,0}^{\infty}(\theta)\}$. The proof of Lemma 3.10 is thus complete. □

Lemma 3.11. *Assume Setting 3.1, assume $h > 1$, and assume for all $x \in \mathbb{R}$ that $f(x) = x^2$. Then*

$$\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}^{\infty}_{0,0}(\vartheta) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{0,0}(\theta)\} = \varnothing. \tag{3.41}$$

Proof of Lemma 3.11. Note that f is real analytic. Furthermore, observe that for all $\theta \in \mathbb{R}^{\mathfrak{d}}$ it holds that $\mathscr{N}^{\theta,\infty}_{0,0}$ is real analytic. Moreover, note that Proposition 3.8 ensures that $\inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{0,0}(\theta) = 0$. We prove (3.41) by contradiction. Assume that there exists $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ such that

$$\mathcal{L}^{\infty}_{0,0}(\vartheta) = 0. \tag{3.42}$$

Observe that (3.42) establishes that for all $x \in [\mathscr{a}, \mathscr{b}]$ it holds that $f(x) = \mathscr{N}^{\vartheta,\infty}_{0,0}(x)$. Combining this with the fact that f and $\mathscr{N}^{\vartheta,\infty}_{0,0}$ are real analytic implies for all $x \in \mathbb{R}$ that

$$f(x) = \mathscr{N}^{\vartheta,\infty}_{0,0}(x). \tag{3.43}$$

Note that for all $x \in \mathbb{R}$ it holds that

$$|(\mathscr{N}^{\vartheta,\infty}_{0,0})'(x)| = \left| \sum_{k=1}^{h} \frac{\mathfrak{v}^{\vartheta}_k \mathfrak{w}^{\vartheta}_k \exp(\mathfrak{w}^{\vartheta}_k x + \mathfrak{b}^{\vartheta}_k)}{1 + \exp(\mathfrak{w}^{\vartheta}_k x + \mathfrak{b}^{\vartheta}_k)} \right| \leq \sum_{k=1}^{h} |\mathfrak{v}^{\vartheta}_k \mathfrak{w}^{\vartheta}_k|. \tag{3.44}$$

This, the fact that f' is unbounded, and (3.43) show the contradiction. The proof of Lemma 3.11 is thus complete. ☐

3.4 ANNs with standard logistic, hyperbolic tangent, arctangent, and inverse square root unit activation

Proposition 3.12. *Assume Setting 3.1, assume for all $x \in \mathbb{R}$ that $f(x) = x$, and let $(\theta_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}^{\theta_n}_1 = 1/n$, $\mathfrak{b}^{\theta_n}_1 = 0$, $\mathfrak{v}^{\theta_n}_1 = 4n$, $\mathfrak{c}^{\theta_n} = -n$, and $\sum_{j=2}^{h} |\mathfrak{w}^{\theta_n}_j| + |\mathfrak{b}^{\theta_n}_j| + |\mathfrak{v}^{\theta_n}_j| = 0$. Then*

$$\limsup_{n \to \infty} \mathcal{L}^{\infty}_{-1,0}(\theta_n) = 0. \tag{3.45}$$

Proof of Proposition 3.12. Observe that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}^{\theta_n,\infty}_{-1,0}(x) = a_{-1,0}\left(\frac{x}{n}\right) 4n - 2n. \tag{3.46}$$

This and the fact that there exists $c \in (0, \infty)$ such that for all $x \in [-1, 1]$ it holds that $|a_{-1,0}(x) - 1/2 - x/4 + x^3/48| \leq c|x^4|$ demonstrate that for all $x \in \mathbb{R}$ it holds that

$$\limsup_{n \to \infty} |\mathscr{N}^{\theta_n,\infty}_{-1,0}(x) - x| = 0. \tag{3.47}$$

Furthermore, note that the mean-value theorem demonstrates that for all $x \in \mathbb{R}$ there exists $\tilde{x} \in [\min\{0, x\}, \max\{0, x\}]$ which satisfies that $a_{-1,0}(x) - 1/2 = x\, a'_{-1,0}(\tilde{x})$. This and the fact that $a'_{-1,0}$ is continuous imply that there exists $L \in \mathbb{R}$ such that for all $x \in [\mathscr{a}, \mathscr{b}]$, $n \in \mathbb{N}$ it holds that

$$\begin{aligned} |\mathscr{N}^{\theta_n,\infty}_{-1,0}(x) - x| &= |4n(a_{-1,0}(x/n) - 1/2) - x| = |x||4a'_{-1,0}(\tilde{x}) - 1| \\ &\leq \max\{|\mathscr{a}|, |\mathscr{b}|\}|4a'_{-1,0}(\tilde{x}) - 1| \leq L. \end{aligned} \tag{3.48}$$

Combining this, (3.47), and Lebesgue's dominated convergence theorem proves that

$$\limsup_{n \to \infty} \mathcal{L}^{\infty}_{-1,0}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup_{n \to \infty} \big(\mathscr{N}^{\theta_n,\infty}_{-1,0}(x) - x\big)^2 \, \mathrm{d}x = 0. \tag{3.49}$$

The proof of Proposition 3.12 is thus complete. ☐

Proposition 3.13. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in \mathbb{R}$ that $f(x) = x^2$, and let $(\theta_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = -\mathfrak{w}_2^{\theta_n} = -1/n$, $\mathfrak{b}_1^{\theta_n} = \mathfrak{b}_2^{\theta_n} = -1$, $\mathfrak{v}_1^{\theta_n} = \mathfrak{v}_2^{\theta_n} = n^2(1+e)^3(e(e-1))^{-1}$, $\mathfrak{c}^{\theta_n} = -2n^2(1+e)^2(e(e-1))^{-1}$, and $\sum_{j=3}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup_{n\to\infty} \mathcal{L}_{-1,0}^{\infty}(\theta_n) = 0. \tag{3.50}$$

Proof of Proposition 3.13. Observe that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}_{-1,0}^{\theta_n,\infty}(x) = n^2 \frac{(1+e)^3}{e(e-1)} \left(\frac{1}{1+e^{-\frac{x}{n}+1}} + \frac{1}{1+e^{\frac{x}{n}+1}} \right) - 2n^2 \frac{(1+e)^2}{e(e-1)}. \tag{3.51}$$

This and the fact that there exists $c \in (0,\infty)$ such that for all $x \in [-1,1]$ it holds that

$$\left| \frac{1}{1+e^{x+1}} - \frac{1}{1+e} + \frac{e\,x}{(1+e)^2} - \frac{e(e-1)x^2}{2(1+e)^3} \right| \le c|x^3| \tag{3.52}$$

assure that there exist $c, M \in (0,\infty)$ such that for all $x \in [\mathscr{a}, \mathscr{b}]$, $n \ge M$ it holds that

$$|\mathscr{N}_{-1,0}^{\theta_n,\infty}(x) - x^2| \le c \left| \frac{1}{n} \right|. \tag{3.53}$$

This and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup_{n\to\infty} \mathcal{L}_{-1,0}^{\infty}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup_{n\to\infty} \big(x^2 - \mathscr{N}_{-1,0}^{\theta_n,\infty}(x)\big)^2 \, dx = 0. \tag{3.54}$$

The proof of Proposition 3.13 is thus complete. □

Proposition 3.14. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in \mathbb{R}$ that $f(x) = x^2$, and let $(\theta_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = -\mathfrak{w}_2^{\theta_n} = 1/n$, $\mathfrak{b}_1^{\theta_n} = \mathfrak{b}_2^{\theta_n} = -1$, $\mathfrak{v}_1^{\theta_n} = \mathfrak{v}_2^{\theta_n} = n^2(1+e^2)^3(8e^2(e^2-1))^{-1}$, $\mathfrak{c}^{\theta_n} = n^2(1+e^2)^2(4e^2)^{-1}$, and $\sum_{j=3}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup_{n\to\infty} \mathcal{L}_{-2,0}^{\infty}(\theta_n) = 0. \tag{3.55}$$

Proof of Proposition 3.14. Note that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}_{-2,0}^{\theta_n,\infty}(x) = \frac{n^2(1+e^2)^3}{8e^2(e^2-1)} \left(\frac{e^{\frac{x}{n}-1} - e^{-\frac{x}{n}+1}}{e^{\frac{x}{n}-1} + e^{-\frac{x}{n}+1}} + \frac{e^{-\frac{x}{n}-1} - e^{\frac{x}{n}+1}}{e^{-\frac{x}{n}-1} + e^{\frac{x}{n}+1}} \right) + \frac{n^2(1+e^2)^2}{4e^2}. \tag{3.56}$$

This and the fact that there exists $c \in (0,\infty)$ such that for all $x \in [-1,1]$ it holds that

$$\left| \frac{e^{x-1} - e^{-x+1}}{e^{x-1} + e^{-x+1}} - \frac{1-e^2}{1+e^2} - \frac{4e^2 x}{(1+e^2)^2} - \frac{4e^2(e^2-1)x^2}{(1+e^2)^3} \right| \le c|x^3| \tag{3.57}$$

assure that there exist $c, M \in (0,\infty)$ such that for all $x \in [\mathscr{a}, \mathscr{b}]$, $n \ge M$ it holds that

$$|\mathscr{N}_{-2,0}^{\theta_n,\infty}(x) - x^2| \le c \left| \frac{1}{n} \right|. \tag{3.58}$$

This and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup_{n\to\infty} \mathcal{L}_{-2,0}^{\infty}(\theta_n) = \int_{-1}^{1} \limsup_{n\to\infty} \big(x^2 - \mathscr{N}_{-2,0}^{\theta_n,\infty}(x)\big)^2 \, dx = 0. \tag{3.59}$$

The proof of Proposition 3.14 is thus complete. □

Proposition 3.15. *Assume Setting 3.1, assume for all $x \in \mathbb{R}$ that $f(x) = x$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = 1/n$, $\mathfrak{v}_1^{\theta_n} = n$, and $|\mathfrak{b}_1^{\theta_n}| + |\mathfrak{c}^{\theta_n}| + \sum_{j=2}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup_{n\to\infty} \mathcal{L}_{-4,0}^{\infty}(\theta_n) = 0. \tag{3.60}$$

Proof of Proposition 3.15. Observe that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}_{-4,0}^{\theta_n,\infty}(x) = \frac{x}{\sqrt{1+\xi(\frac{x}{n})^2}}. \tag{3.61}$$

This shows that for all $x \in \mathbb{R}$ it holds that

$$\limsup_{n\to\infty} |\mathscr{N}_{-4,0}^{\theta_n,\infty}(x) - x| = 0. \tag{3.62}$$

Furthermore, note that (3.61) implies that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$|\mathscr{N}_{-4,0}^{\theta_n,\infty}(x)| \le x. \tag{3.63}$$

This, (3.62), and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup_{n\to\infty} \mathcal{L}_{-4,0}^{\infty}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup_{n\to\infty} \bigl(x - \mathscr{N}_{-4,0}^{\theta_n,\infty}(x)\bigr)^2 \,\mathrm{d}x = 0. \tag{3.64}$$

The proof of Proposition 3.15 is thus complete. □

Proposition 3.16. *Assume Setting 3.1, assume $h > 1$ and $\xi < 3$, assume for all $x \in \mathbb{R}$ that $f(x) = x^2$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = -\mathfrak{w}_2^{\theta_n} = 1/n$, $\mathfrak{v}_1^{\theta_n} = \mathfrak{v}_2^{\theta_n} = -(\xi+4)^{\frac{5}{2}} n^2 (48\xi)^{-1}$, $\mathfrak{b}_1^{\theta_n} = \mathfrak{b}_2^{\theta_n} = 1/2$, $\mathfrak{c}^{\theta_n} = (\xi+4)^{\frac{5}{2}} n^2 (48\xi\sqrt{1+\xi/4})^{-1}$, and $\sum_{j=3}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup_{n\to\infty} \mathcal{L}_{-4,0}^{\infty}(\theta_n) = 0. \tag{3.65}$$

Proof of Proposition 3.16. Observe that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}_{-4,0}^{\theta_n,\infty}(x) = -\frac{(\xi+4)^{\frac{5}{2}} n^2}{48\xi} \left(\frac{\frac{x}{n} + \frac{1}{2}}{\sqrt{1+\xi(\frac{x}{n}+\frac{1}{2})^2}} + \frac{-\frac{x}{n} + \frac{1}{2}}{\sqrt{1+\xi(-\frac{x}{n}+\frac{1}{2})^2}} - \frac{1}{\sqrt{1+\frac{\xi}{4}}} \right). \tag{3.66}$$

This and the fact that there exists $c \in (0,\infty)$ such that for all $x \in [-1,1]$ it holds that

$$\left| \frac{x+\frac{1}{2}}{\sqrt{(1+\xi(x+\frac{1}{2})^2)}} - \frac{1}{(\xi+4)^{\frac{1}{2}}} - \frac{8x}{(\xi+4)^{\frac{3}{2}}} + \frac{24\xi x^2}{(\xi+4)^{\frac{5}{2}}} \right| \le c|x^3| \tag{3.67}$$

assure that there exist $c, M \in (0,\infty)$ such that for all $x \in [\mathscr{a}, \mathscr{b}]$, $n \ge M$ it holds that

$$|\mathscr{N}_{-4,0}^{\theta_n,\infty}(x) - x^2| \le c\left|\frac{1}{n}\right|. \tag{3.68}$$

This and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup_{n\to\infty} \mathcal{L}_{-4,0}^{\infty}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup_{n\to\infty} \bigl(x^2 - \mathscr{N}_{-4,0}^{\theta_n,\infty}(x)\bigr)^2 \,\mathrm{d}x = 0. \tag{3.69}$$

The proof of Proposition 3.16 is thus complete. □

Proposition 3.17. *Assume Setting 3.1, assume for all $x \in \mathbb{R}$ that $f(x) = x$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = 1/n$, $\mathfrak{v}_1^{\theta_n} = n$, and $|\mathfrak{b}_1^{\theta_n}| + |\mathfrak{c}^{\theta_n}| + \sum_{j=2}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup_{n\to\infty} \mathcal{L}_{-5,0}^{\infty}(\theta_n) = 0. \tag{3.70}$$

Proof of Proposition 3.17. Note that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}^{\theta_n,\infty}_{-5,0}(x) = n \arctan\left(\frac{x}{n}\right). \tag{3.71}$$

This shows that for all $x \in \mathbb{R}$ it holds that

$$\limsup_{n\to\infty} |\mathscr{N}^{\theta_n,\infty}_{-5,0}(x) - x| = 0. \tag{3.72}$$

Furthermore, observe that (3.71) implies that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$|\mathscr{N}^{\theta_n,\infty}_{-5,0}(x)| \le |x|. \tag{3.73}$$

This, (3.72), and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup_{n\to\infty} \mathcal{L}^{\infty}_{-5,0}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup_{n\to\infty} \bigl(x - \mathscr{N}^{\theta_n,\infty}_{-5,0}(x)\bigr)^2 \, \mathrm{d}x = 0. \tag{3.74}$$

The proof of Proposition 3.17 is thus complete. ☐

Proposition 3.18. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in \mathbb{R}$ that $f(x) = x^2$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}^{\theta_n}_1 = -\mathfrak{w}^{\theta_n}_2 = 1/n$, $\mathfrak{v}^{\theta_n}_1 = \mathfrak{v}^{\theta_n}_2 = -2n^2$, $\mathfrak{b}^{\theta_n}_1 = \mathfrak{b}^{\theta_n}_2 = 1$, $\mathfrak{c}^{\theta_n} = n^2\pi$, and $\sum_{j=3}^{h} |\mathfrak{w}^{\theta_n}_j| + |\mathfrak{b}^{\theta_n}_j| + |\mathfrak{v}^{\theta_n}_j| = 0$. Then*

$$\limsup_{n\to\infty} \mathcal{L}^{\infty}_{-5,0}(\theta_n) = 0. \tag{3.75}$$

Proof of Proposition 3.18. Note that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}^{\theta_n,\infty}_{-5,0}(x) = -2n^2 \arctan\left(\frac{x}{n} + 1\right) - 2n^2 \arctan\left(-\frac{x}{n} + 1\right) + \pi n^2. \tag{3.76}$$

This and the fact that there exists $c \in (0,\infty)$ such that for all $x \in [-1,1]$ it holds that

$$\left|\arctan(x+1) - \frac{\pi}{4} - \frac{x}{2} + \frac{x^2}{4}\right| \le c|x^3| \tag{3.77}$$

assure that there exist $c, M \in (0,\infty)$ such that for all $x \in [\mathscr{a}, \mathscr{b}]$, $n \ge M$ it holds that

$$|\mathscr{N}^{\theta_n,\infty}_{-5,0}(x) - x^2| \le c\left|\frac{1}{n}\right|. \tag{3.78}$$

This and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup_{n\to\infty} \mathcal{L}^{\infty}_{-5,0}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup_{n\to\infty} \bigl(x^2 - \mathscr{N}^{\theta_n,\infty}_{-5,0}(x)\bigr)^2 \, \mathrm{d}x = 0. \tag{3.79}$$

The proof of Proposition 3.18 is thus complete. ☐

Lemma 3.19. *Assume Setting 3.1 and let $\theta \in \mathbb{R}^{\mathfrak{d}}$. Then there exists $c \in \mathbb{R}$ such that for all $k \in \{-1,-2,-4,-5\}$, $x \in \mathbb{R}$ it holds that*

$$|\mathscr{N}^{\theta,\infty}_{k,0}(x)| \le c. \tag{3.80}$$

Proof of Lemma 3.19. Observe that (3.1) ensures that for all $j \in \{-1,-2,-4,-5\}$, $x \in \mathbb{R}$ it holds that

$$|a_{j,0}(x)| \le \begin{cases} \frac{\pi}{2} & : k = -5 \\ \frac{|x|}{\sqrt{\xi x^2}} \le \frac{1}{\sqrt{\xi}} & : k = -4 \\ \frac{\exp(x)+\exp(-x)}{\exp(x)+\exp(-x)} \le 1 & : k = -2 \\ 1 & : k = -1. \end{cases} \tag{3.81}$$

Hence, we obtain that for all $j \in \{-1,-2,-4,-5\}$, $x \in \mathbb{R}$ it holds that

$$|\mathscr{N}_{j,0}^{\theta,\infty}(x)| \leq \begin{cases} |\mathfrak{c}^\theta| + \sum_{i=1}^h \frac{\pi}{2}|\mathfrak{v}_i^\theta| \leq \frac{\pi}{2}\|\theta\|^2 & : k=-5 \\ |\mathfrak{c}^\theta| + \sum_{i=1}^h \frac{1}{\sqrt{\xi}}|\mathfrak{v}_i^\theta| \leq \frac{1}{\sqrt{\xi}}\|\theta\|^2 & : k=-4 \\ |\mathfrak{c}^\theta| + \sum_{i=1}^h |\mathfrak{v}_i^\theta| \leq \|\theta\|^2 & : k=-2 \\ |\mathfrak{c}^\theta| + \sum_{i=1}^h |\mathfrak{v}_i^\theta| \leq \|\theta\|^2 & : k=-1. \end{cases} \tag{3.82}$$

This implies that for all $j \in \{-1,-2,-4,-5\}$, $x \in \mathbb{R}$ it holds that

$$|\mathscr{N}_{j,0}^{\theta,\infty}(x)| \leq \max\{\frac{\pi}{2}\|\theta\|^2, \frac{1}{\sqrt{\xi}}\|\theta\|^2\}. \tag{3.83}$$

The proof of Lemma 3.19 is thus complete. □

Lemma 3.20. *Assume Setting 3.1, let $k \in \{-1,-4,-5\}$, and assume for all $x \in \mathbb{R}$ that $f(x) = x$. Then*

$$\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}_{k,0}^\infty(\vartheta) = \inf\nolimits_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{k,0}^\infty(\theta)\} = \varnothing. \tag{3.84}$$

Proof of Lemma 3.20. Note that f is real analytic. Furthermore, observe that for all $\theta \in \mathbb{R}^{\mathfrak{d}}$ it holds that $\mathscr{N}_{k,0}^{\theta,\infty}$ is real analytic. Moreover, note that Proposition 3.12, Proposition 3.16, and Proposition 3.17 ensure that for all $j \in \{-1,-4,-5\}$ it holds that $\inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{j,0}^\infty(\theta) = 0$. We prove (3.84) by contradiction. Assume that there exists $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ which satisfies that

$$\mathcal{L}_{k,0}^\infty(\vartheta) = 0. \tag{3.85}$$

Observe that (3.85) establishes that for all $x \in [\mathscr{a}, \mathscr{b}]$ it holds that $f(x) = \mathscr{N}_{k,0}^{\vartheta,\infty}(x)$. Combining this with the fact that f and $\mathscr{N}_{k,0}^{\vartheta,\infty}$ are real analytic implies for all $x \in \mathbb{R}$ that

$$f(x) = \mathscr{N}_{k,0}^{\vartheta,\infty}(x). \tag{3.86}$$

Note that Lemma 3.19 assures that there exists $c \in \mathbb{R}$ such that for all $j \in \{-1,-4,-5\}$, $x \in \mathbb{R}$ it holds that

$$|\mathscr{N}_{j,0}^{\vartheta,\infty}(x)| \leq c. \tag{3.87}$$

Combining this, the fact that f is unbounded, and (3.86) shows the contradiction. The proof of Lemma 3.20 is thus complete. □

Lemma 3.21. *Assume Setting 3.1, let $k \in \{-1,-2,-4,-5\}$, assume $h > 1$, assume $\xi < 3$, and assume for all $x \in \mathbb{R}$ that $f(x) = x^2$. Then*

$$\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}_{k,0}^\infty(\vartheta) = \inf\nolimits_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{k,0}^\infty(\theta)\} = \varnothing. \tag{3.88}$$

Proof of Lemma 3.21. Observe that f is real analytic. Furthermore, note that for all $\theta \in \mathbb{R}^{\mathfrak{d}}$ it holds that $\mathscr{N}_{k,0}^{\theta,\infty}$ is real analytic. Moreover, observe that Proposition 3.13, Proposition 3.14, Proposition 3.18, and Proposition 3.16 ensure that for all $j \in \{-1,-2,-4,-5\}$ it holds that $\inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{j,0}^\infty(\theta) = 0$. We prove (3.88) by contradiction. Assume that there exists $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ such that

$$\mathcal{L}_{k,0}^\infty(\vartheta) = 0. \tag{3.89}$$

Note that (3.89) establishes that for all $x \in [\mathscr{a}, \mathscr{b}]$ it holds that $f(x) = \mathscr{N}_{k,0}^{\vartheta,\infty}(x)$. Combining this with the fact that f and $\mathscr{N}_{k,0}^{\vartheta,\infty}$ are real analytic implies for all $x \in \mathbb{R}$ that

$$f(x) = \mathscr{N}_{k,0}^{\vartheta,\infty}(x). \tag{3.90}$$

In addition, observe that Lemma 3.19 assures that there exists $c \in \mathbb{R}$ such that for all $j \in \{-1,-2,-4,-5\}$, $x \in \mathbb{R}$ it holds that

$$|\mathscr{N}_{j,0}^{\vartheta,\infty}(x)| \leq c. \tag{3.91}$$

This shows that for all $j \in \{-1,-2,-4,-5\}$, $x \in \mathbb{R}$ it holds that $\sup_{x\in\mathbb{R}} |\mathscr{N}_{j,0}^{\vartheta,\infty}(x)| < \infty$. Combining this with (3.90) assures that

$$\infty = \sup_{x\in\mathbb{R}} |x^2| = \sup_{x\in\mathbb{R}} |f(x)| = \sup_{x\in\mathbb{R}} |\mathscr{N}_{k,0}^{\vartheta,\infty}(x)| < \infty. \tag{3.92}$$

This contradiction establishes (3.88). The proof of Lemma 3.21 is thus complete. □

3.5 ANNs with rectified power unit activation

Proposition 3.22. *Assume Setting 3.1, assume $h > 1$, let $k \in \mathbb{N}\backslash\{1\}$ satisfy for all $x \in \mathbb{R}$ that $f(x) = (\max\{x - (\mathscr{a}+\mathscr{b})/2, 0\})^{k-1}$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = \mathfrak{w}_2^{\theta_n} = 1$, $\mathfrak{b}_1^{\theta_n} = 1/n - (\mathscr{a}+\mathscr{b})/2$, $\mathfrak{b}_2^{\theta_n} = -(\mathscr{a}+\mathscr{b})/2$, $\mathfrak{v}_1^{\theta_n} = -\mathfrak{v}_2^{\theta_n} = n/k$, and $|\mathfrak{c}^{\theta_n}| + \sum_{j=3}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup_{n\to\infty} \mathcal{L}_{k,0}^{\infty}(\theta_n) = 0. \tag{3.93}$$

Proof of Proposition 3.22. Note that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\begin{aligned}\mathscr{N}_{k,0}^{\theta_n,\infty}(x) &= \frac{n}{k}(\max\{x + 1/n - (\mathscr{a}+\mathscr{b})/2, 0\})^k - \frac{n}{k}(\max\{x - (\mathscr{a}+\mathscr{b})/2, 0\})^k \\ &= \begin{cases} 0 & : x \in \left(-\infty, -\frac{1}{n} + \frac{\mathscr{a}+\mathscr{b}}{2}\right) \\ \frac{1}{k}\sum_{i=0}^{k}\binom{k}{i}\left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right)^{k-i}\left(\frac{1}{n}\right)^{i-1} & : x \in \left[-\frac{1}{n} + \frac{\mathscr{a}+\mathscr{b}}{2}, \frac{\mathscr{a}+\mathscr{b}}{2}\right) \\ \left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right)^{k-1} + \frac{n}{k}\sum_{i=2}^{k}\binom{k}{i}\left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right)^{k-i}\left(\frac{1}{n}\right)^{i} & : x \in \left[\frac{\mathscr{a}+\mathscr{b}}{2}, \infty\right). \end{cases}\end{aligned} \tag{3.94}$$

This implies that for all $x \in \mathbb{R}$ it holds that

$$\limsup_{n\to\infty} |\mathscr{N}_{k,0}^{\theta_n,\infty}(x) - (\max\{x - (\mathscr{a}+\mathscr{b})/2, 0\})^{k-1}| = 0. \tag{3.95}$$

Furthermore, observe that (3.94) proves that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$|\mathscr{N}_{k,0}^{\theta_n,\infty}(x)| \leq \begin{cases} 0 & : x \in \left(-\infty, -\frac{1}{n} + \frac{\mathscr{a}+\mathscr{b}}{2}\right) \\ \frac{1}{k} & : x \in \left[-\frac{1}{n} + \frac{\mathscr{a}+\mathscr{b}}{2}, \frac{\mathscr{a}+\mathscr{b}}{2}\right) \\ \left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right)^{k-1} + \frac{1}{k}\sum_{i=2}^{k}\binom{k}{i}\left(x - \frac{\mathscr{a}+\mathscr{b}}{2}\right)^{k-i} & : x \in \left[\frac{\mathscr{a}+\mathscr{b}}{2}, \infty\right). \end{cases} \tag{3.96}$$

This, (3.95), and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup_{n\to\infty} \mathcal{L}_{k,0}^{\infty}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup_{n\to\infty} \left((\max\{x - (\mathscr{a}+\mathscr{b})/2, 0\})^{k-1} - \mathscr{N}_{k,0}^{\theta_n,\infty}(x)\right)^2 \mathrm{d}x = 0. \tag{3.97}$$

The proof of Proposition 3.22 is thus complete. □

Lemma 3.23. *Assume Setting 3.1, assume $h > 1$, and let $k \in \mathbb{N}\backslash\{1\}$ satisfy for all $x \in \mathbb{R}$ that $f(x) = (\max\{x - (\mathscr{a}+\mathscr{b})/2, 0\})^{k-1}$. Then*

$$\left\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}_{k,0}^{\infty}(\vartheta) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{k,0}^{\infty}(\theta)\right\} = \varnothing. \tag{3.98}$$

Proof of Lemma 3.23. Note that the fact that $k \in \mathbb{N}\backslash\{1\}$ ensures that for all $\theta \in \mathbb{R}^{\mathfrak{d}}$ it holds that $\mathscr{N}_{k,0}^{\theta,\infty} \in C^{k-1}([\mathscr{a},\mathscr{b}],\mathbb{R})$. Furthermore, observe that Proposition 3.22 establishes that $\inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{k,0}^{\infty}(\theta) = 0$. We prove (3.98) by contradiction. Assume that there exists $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ such that

$$\mathcal{L}_{k,0}^{\infty}(\vartheta) = 0. \tag{3.99}$$

Note that (3.99) implies that for all $x \in [\mathscr{a},\mathscr{b}]$ it holds that

$$\mathscr{N}_{k,0}^{\vartheta,\infty}(x) = (\max\{x - (\mathscr{a}+\mathscr{b})/2, 0\})^{k-1}. \tag{3.100}$$

This establishes that $\mathscr{N}_{k,0}^{\vartheta,\infty} \in C^{k-2}([\mathscr{a},\mathscr{b}],\mathbb{R})\backslash C^{k-1}([\mathscr{a},\mathscr{b}],\mathbb{R})$ which is a contradiction. The proof of Lemma 3.23 is thus complete. □

3.6 ANNs with exponential linear unit activation

Proposition 3.24. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in \mathbb{R}$ that $f(x) = x^2$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = -\mathfrak{w}_2^{\theta_n} = -1/n$, $\mathfrak{b}_1^{\theta_n} = \mathfrak{b}_2^{\theta_n} = -2|\mathscr{b}|$, $\mathfrak{v}_1^{\theta_n} = \mathfrak{v}_2^{\theta_n} = n^2 e^{2|\mathscr{b}|}$, $\mathfrak{c}^{\theta_n} = -2n^2(1 - e^{2|\mathscr{b}|})$, and $\sum_{j=3}^{h} |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup_{n\to\infty} \mathcal{L}_{-3,0}^{\infty}(\theta_n) = 0. \tag{3.101}$$

Proof of Proposition 3.24. Observe that (3.3) ensures that for all $x \in [\mathscr{a}, \mathscr{b}]$, $n \in \mathbb{N}$ it holds that

$$\begin{aligned} \mathscr{N}_{-3,0}^{\theta_n,\infty}(x) &= e^{2|\mathscr{b}|} n^2 (e^{-2|\mathscr{b}| - \frac{x}{n}} - 1 + e^{-2|\mathscr{b}| + \frac{x}{n}} - 1) - 2n^2(1 - e^{2|\mathscr{b}|}) \\ &= n^2(e^{-\frac{x}{n}} + e^{\frac{x}{n}} - 2). \end{aligned} \tag{3.102}$$

This and the fact that there exists $c \in (0, \infty)$ such that for all $x \in [-1, 1]$ it holds that

$$\left| e^x - 1 - x - \frac{x^2}{2} \right| \leq c|x^3| \tag{3.103}$$

assure that there exist $c, M \in (0, \infty)$ such that for all $x \in [\mathscr{a}, \mathscr{b}]$, $n \geq M$ it holds that

$$|\mathscr{N}_{-3,0}^{\theta_n,\infty}(x) - x^2| \leq c \left| \frac{1}{n} \right|. \tag{3.104}$$

This and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup_{n\to\infty} \mathcal{L}_{-3,0}^{\infty}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup_{n\to\infty} \big(x^2 - \mathscr{N}_{-3,0}^{\theta_n,\infty}(x)\big)^2 \, \mathrm{d}x = 0. \tag{3.105}$$

The proof of Proposition 3.24 is thus complete. □

Lemma 3.25. *Assume Setting 3.1, assume $h > 1$, and assume for all $x \in \mathbb{R}$ that $f(x) = x^2$. Then*

$$\big\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}_{-3,0}^{\infty}(\vartheta) = \inf\nolimits_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{-3,0}^{\infty}(\theta)\big\} = \varnothing. \tag{3.106}$$

Proof of Lemma 3.25. Note that Proposition 3.24 implies that $\inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{-3,0}^{\infty}(\theta) = 0$. We prove (3.106) by contradiction. We thus assume that there exists $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ which satisfies that

$$\mathcal{L}_{-3,0}^{\infty}(\vartheta) = 0. \tag{3.107}$$

Observe that (3.107) implies that for all $x \in [\mathscr{a}, \mathscr{b}]$ it holds that

$$\mathscr{N}_{-3,0}^{\vartheta,\infty}(x) = x^2. \tag{3.108}$$

Therefore, we obtain $\mathscr{N}_{-3,0}^{\vartheta,\infty} \in C^{\infty}([\mathscr{a}, \mathscr{b}], \mathbb{R})$. This demonstrates that for all $i \in \{1, 2, \dots, h\}$ it holds that $\{x \in (\mathscr{a}, \mathscr{b}) \colon \mathfrak{w}_i^{\vartheta} x + \mathfrak{b}_i^{\vartheta} \leq 0\} \in \{(\mathscr{a}, \mathscr{b}), \varnothing\}$. For every $i \in \{1, 2, \dots, h\}$ let $Q^i \subseteq \mathbb{R}$ satisfy $Q^i = \{x \in (\mathscr{a}, \mathscr{b}) \colon \mathfrak{w}_i^{\vartheta} x + \mathfrak{b}_i^{\vartheta} \leq 0\}$ and let $S \subseteq \mathbb{N}$ satisfy $S = \{i \in \{1, 2, \dots, h\} \colon Q^i = (\mathscr{a}, \mathscr{b})\}$. In the following we distinguish between the case $|S| = 0$ and the case $|S| > 0$. We first establish the contradiction in the case

$$|S| = 0. \tag{3.109}$$

Note that (3.109) ensures that there exist $\alpha, \beta \in \mathbb{R}$ such that for all $x \in [\mathscr{a}, \mathscr{b}]$ it holds that $\mathscr{N}_{-3,0}^{\vartheta,\infty}(x) = \alpha x + \beta$ which is a contradiction. In the next step we establish the contradiction in the case

$$|S| > 0. \tag{3.110}$$

For every $i \in S$ let $y_i \in \mathbb{R}$ satisfy $y_i = \mathfrak{v}_i^\vartheta \exp\left(\mathfrak{w}_i^\vartheta(\mathscr{a}+\mathscr{b}/2) + \mathfrak{b}_i^\vartheta\right)$, let $k \in \mathbb{N}$ satisfy $|S| = k$, and assume without loss of generality that $S = \{1, 2, \ldots, k\}$. Observe that (3.108) proves that for all $n \in \mathbb{N} \cap (2, \infty)$, $x \in [\mathscr{a}, \mathscr{b}]$ it holds that

$$0 = (\mathscr{N}_{-3,0}^{\vartheta,\infty})^{(n)}(x) = \textstyle\sum_{i \in S} \mathfrak{v}_i^\vartheta (\mathfrak{w}_i^\vartheta)^n \exp(\mathfrak{w}_i^\vartheta x + \mathfrak{b}_i^\vartheta) \qquad \text{and} \qquad 2 = (\mathscr{N}_{-3,0}^{\vartheta,\infty})^{(2)}(x). \tag{3.111}$$

This implies that

$$\begin{aligned} (\mathfrak{w}_1^\vartheta)^2 y_1 + \ldots + (\mathfrak{w}_k^\vartheta)^2 y_k &= 2, \\ (\mathfrak{w}_1^\vartheta)^3 y_1 + \ldots + (\mathfrak{w}_k^\vartheta)^3 y_k &= 0, \\ &\ldots \\ \text{and} \qquad (\mathfrak{w}_1^\vartheta)^{k+2} y_1 + \ldots + (\mathfrak{w}_k^\vartheta)^{k+2} y_k &= 0. \end{aligned} \tag{3.112}$$

Hence, we obtain there exists $\eta = (\eta_1, \ldots, \eta_k) \in \mathbb{R}^k$ such that for all $j \in S$ it holds that

$$\textstyle\sum_{i=1}^k \eta_i (\mathfrak{w}_j^\vartheta)^{2+i} = (\mathfrak{w}_j^\vartheta)^2. \tag{3.113}$$

This and (3.112) show that $0 = (\mathfrak{w}_1^\vartheta)^2 y_1 + \ldots + (\mathfrak{w}_k^\vartheta)^2 y_k = 2$ which is a contradiction. The proof of Lemma 3.25 is thus complete. □

3.7 ANNs with softsign activation

Proposition 3.26. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in \mathbb{R}$ that $f(x) = x^2$, let $\mathscr{c} \in \mathbb{R}$ satisfy $\mathscr{c} = \max\{|\mathscr{a}|, |\mathscr{b}|\}$, and let $(\theta_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^{\mathfrak{d}}$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}_1^{\theta_n} = -\mathfrak{w}_2^{\theta_n} = 1/n$, $\mathfrak{v}_1^{\theta_n} = \mathfrak{v}_2^{\theta_n} = -(\mathscr{c}+1)^3 n^2/2$, $\mathfrak{b}_1^{\theta_n} = \mathfrak{b}_2^{\theta_n} = \mathscr{c}$, $\mathfrak{c}^{\theta_n} = \mathscr{c}(\mathscr{c}+1)^2 n^2$, and $\sum_{j=3}^h |\mathfrak{w}_j^{\theta_n}| + |\mathfrak{b}_j^{\theta_n}| + |\mathfrak{v}_j^{\theta_n}| = 0$. Then*

$$\limsup\nolimits_{n \to \infty} \mathcal{L}_{-6,0}^{\infty}(\theta_n) = 0. \tag{3.114}$$

Proof of Proposition 3.26. Note that (3.3) ensures that for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}_{-6,0}^{\theta_n,\infty}(x) = -\frac{(\mathscr{c}+1)^3 n^2}{2} \left(\frac{\frac{x}{n} + \mathscr{c}}{1 + |\frac{x}{n} + \mathscr{c}|} + \frac{-\frac{x}{n} + \mathscr{c}}{1 + |-\frac{x}{n} + \mathscr{c}|} \right) + \mathscr{c}(\mathscr{c}+1)^2 n^2. \tag{3.115}$$

This and the fact that there exists $c \in (0, \infty)$ such that for all $x \in [-1, 1]$ it holds that

$$\left| \frac{x + \mathscr{c}}{1 + x + \mathscr{c}} - \frac{\mathscr{c}}{\mathscr{c}+1} - \frac{x}{(\mathscr{c}+1)^2} + \frac{x^2}{(\mathscr{c}+1)^3} \right| \le c|x^3| \tag{3.116}$$

assure that there exist $c, M \in (0, \infty)$ such that for all $x \in [\mathscr{a}, \mathscr{b}]$, $n \ge M$ it holds that

$$|\mathscr{N}_{-6,0}^{\theta_n,\infty}(x) - x^2| \le c \left| \frac{1}{n} \right|. \tag{3.117}$$

This and Lebesgue's dominated convergence theorem demonstrate that

$$\limsup\nolimits_{n \to \infty} \mathcal{L}_{-6,0}^{\infty}(\theta_n) = \int_{\mathscr{a}}^{\mathscr{b}} \limsup\nolimits_{n \to \infty} \left(x^2 - \mathscr{N}_{-6,0}^{\theta_n,\infty}(x)\right)^2 \mathrm{d}x = 0. \tag{3.118}$$

The proof of Proposition 3.26 is thus complete. □

Lemma 3.27. *Assume Setting 3.1, assume $h > 1$, and assume for all $x \in \mathbb{R}$ that $f(x) = x^2$. Then*

$$\left\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}_{-6,0}^{\infty}(\vartheta) = \inf\nolimits_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}_{-6,0}^{\infty}(\theta)\right\} = \varnothing. \tag{3.119}$$

Proof of Lemma 3.27. Observe that Proposition 3.26 implies that $\inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{-6,0}(\theta) = 0$. We prove (3.119) by contradiction. Assume that there exists $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ which satisfies

$$\mathcal{L}^{\infty}_{-6,0}(\vartheta) = 0. \tag{3.120}$$

Note that (3.120) implies that for all $x \in [\mathscr{a}, \mathscr{b}]$ it holds that

$$\mathscr{N}^{\vartheta,\infty}_{-6,0}(x) = x^2. \tag{3.121}$$

Therefore we obtain that $\mathscr{N}^{\vartheta,\infty}_{-6,0} \in C^{\infty}([\mathscr{a}, \mathscr{b}], \mathbb{R})$. This demonstrates that for all $i \in \{1, 2, \dots, h\}$ it holds that $\{x \in (\mathscr{a}, \mathscr{b}) \colon \mathfrak{w}^{\vartheta}_i x + \mathfrak{b}^{\vartheta}_i = 0\} = \varnothing$. This implies that for all $i \in \{1, 2, \dots, h\}$ there exists $k_i \in \{-1, 1\}$ which satisfies for all $x \in [\mathscr{a}, \mathscr{b}]$ that

$$\mathscr{N}^{\vartheta,\infty}_{-6,0}(x) = \textstyle\sum_{i=1}^{h} \mathfrak{v}^{\vartheta}_i \frac{\mathfrak{w}^{\vartheta}_i x + \mathfrak{b}^{\vartheta}_i}{1 + k_i(\mathfrak{w}^{\vartheta}_i x + \mathfrak{b}^{\vartheta}_i)}. \tag{3.122}$$

Combining this with (3.121) proves that for all $n \in \mathbb{N} \cap (2, \infty)$, $x \in [\mathscr{a}, \mathscr{b}]$ it holds that

$$0 = (\mathscr{N}^{\vartheta,\infty}_{-6,0})^{(n)}(x) = \textstyle\sum_{i=1}^{h} \mathfrak{v}^{\vartheta}_i \frac{n!\, \mathfrak{w}^{\vartheta}_i (-k_i \mathfrak{w}^{\vartheta}_i)^{n-1}}{(1 + k_i(\mathfrak{w}^{\vartheta}_i x + \mathfrak{b}^{\vartheta}_i))^{n+1}} \qquad \text{and} \qquad 2 = (\mathscr{N}^{\vartheta,\infty}_{-6,0})^{(2)}(x). \tag{3.123}$$

Let $S \subseteq \mathbb{N}$ satisfy $S = \{i \in \{1, 2, \dots, h\} \colon \mathfrak{w}^{\vartheta}_i \neq 0\}$ and for every $i \in S$ let $\xi_i \in \mathbb{R}$ satisfy

$$\xi_i = \frac{-k_i \mathfrak{w}^{\vartheta}_i}{1 + k_i(\mathfrak{w}^{\vartheta}_i \frac{\mathscr{a} + \mathscr{b}}{2} + \mathfrak{b}^{\vartheta}_i)}. \tag{3.124}$$

Observe that (3.121) assures that $S \neq \varnothing$. Let $k \in \mathbb{N}$ satisfy $|S| = k$ and assume without loss of generality that $S = \{1, 2, \dots, k\}$. Combining this with (3.123) shows that

$$\begin{aligned} (\eta_1)^3 \frac{\mathfrak{v}^{\vartheta}_1}{\mathfrak{w}^{\vartheta}_1} + \ldots + (\eta_k)^3 \frac{\mathfrak{v}^{\vartheta}_k}{\mathfrak{w}^{\vartheta}_k} &= 2, \\ (\eta_1)^4 \frac{\mathfrak{v}^{\vartheta}_1}{\mathfrak{w}^{\vartheta}_1} + \ldots + (\eta_k)^4 \frac{\mathfrak{v}^{\vartheta}_k}{\mathfrak{w}^{\vartheta}_k} &= 0, \\ &\ \,\cdots \\ \text{and} \qquad (\eta_1)^{k+3} \frac{\mathfrak{v}^{\vartheta}_1}{\mathfrak{w}^{\vartheta}_1} + \ldots + (\eta_k)^{k+3} \frac{\mathfrak{v}^{\vartheta}_k}{\mathfrak{w}^{\vartheta}_k} &= 0. \end{aligned} \tag{3.125}$$

Therefore, we obtain that there exists $\mathscr{c} = (\mathscr{c}_1, \dots, \mathscr{c}_k) \in \mathbb{R}^k$ such that for all $j \in S$ it holds that

$$\textstyle\sum_{i=1}^{k} \mathscr{c}_i (\eta_j)^{3+i} = (\eta_j)^3. \tag{3.126}$$

This and (3.125) show that $0 = (\eta_1)^3 \frac{\mathfrak{v}^{\vartheta}_1}{\mathfrak{w}^{\vartheta}_1} + \ldots + (\eta_k)^3 \frac{\mathfrak{v}^{\vartheta}_k}{\mathfrak{w}^{\vartheta}_k} = 2$ which is a contradiction. The proof of Lemma 3.27 is thus complete. ☐

3.8 Divergence of GFs

Lemma 3.28. *Let $\mathfrak{d} \in \mathbb{N}$, $\Theta \in C([0, \infty), \mathbb{R}^{\mathfrak{d}})$, $\mathcal{L} \in C(\mathbb{R}^{\mathfrak{d}}, \mathbb{R})$ satisfy $\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}(\vartheta) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}(\theta)\} = \varnothing$ and $\liminf_{t \to \infty} \mathcal{L}(\Theta_t) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}(\theta)$, let $\mathcal{G} \colon \mathbb{R}^{\mathfrak{d}} \to \mathbb{R}^{\mathfrak{d}}$ be measurable, and assume for all $t \in [0, \infty)$ that $\mathcal{L}(\Theta_t) = \mathcal{L}(\Theta_0) - \int_0^t \|\mathcal{G}(\Theta_s)\|^2 \, \mathrm{d}s$. Then*

$$\liminf_{t \to \infty} \|\Theta_t\| = \infty. \tag{3.127}$$

Proof of Lemma 3.28. Note that the assumption that for all $t \in [0,\infty)$ it holds that $\mathcal{L}(\Theta_t) = \mathcal{L}(\Theta_0) - \int_0^t \|\mathcal{G}(\Theta_s)\|^2 \, \mathrm{d}s$ assures that $[0,\infty) \ni t \mapsto \mathcal{L}(\Theta_t) \in \mathbb{R}$ is non-increasing. This demonstrates that

$$\limsup_{t\to\infty} \mathcal{L}(\Theta_t) = \liminf_{t\to\infty} \mathcal{L}(\Theta_t) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}(\theta). \tag{3.128}$$

We prove (3.127) by contradiction. We thus assume that $\liminf_{t\to\infty} \|\Theta_t\| < \infty$. Therefore, by compactness, there exist $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ and $\tau_n \in [0,\infty)$, $n \in \mathbb{N}$, which satisfy $\liminf_{n\to\infty} \tau_n = \infty$ and

$$\limsup_{n\to\infty} \|\Theta_{\tau_n} - \vartheta\| = 0. \tag{3.129}$$

Hence, continuity of $\mathcal{L}$ shows that $\limsup_{n\to\infty} |\mathcal{L}(\Theta_{\tau_n}) - \mathcal{L}(\vartheta)| = 0$. Combining this with (3.128) proves that

$$\mathcal{L}(\vartheta) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}(\theta). \tag{3.130}$$

This implies that $\vartheta \in \{\theta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}(\theta) = \inf_{v\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}(v)\}$ which is a contradiction. The proof of Lemma 3.28 is thus complete. □

Corollary 3.29. *Assume Setting 3.1, assume $h > 1$, assume $\xi < 3$, let $k \in \mathbb{Z}\backslash\mathbb{N}$, assume for all $x \in \mathbb{R}$ that $f(x) = x^2$, and let $\Theta \in C([0,\infty), \mathbb{R}^{\mathfrak{d}})$ satisfy $\liminf_{t\to\infty} \mathcal{L}^{\infty}_{k,0}(\Theta_t) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)$ and $\forall\, t \in [0,\infty) \colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}_{k,0}(\Theta_s) \, \mathrm{d}s$. Then $\liminf_{t\to\infty} \|\Theta_t\| = \infty$.*

Proof of Corollary 3.29. Observe that, e.g., [9, Lemma 3.1] implies that for all $t \in [0,\infty)$ it holds that

$$\mathcal{L}^{\infty}_{k,0}(\Theta_t) = \mathcal{L}^{\infty}_{k,0}(\Theta_0) - \int_0^t \|\mathcal{G}_{k,0}(\Theta_s)\|^2 \, \mathrm{d}s. \tag{3.131}$$

Furthermore, note that Lemma 3.11, Lemma 3.21, Lemma 3.25, and Lemma 3.27 assure that for all $j \in \mathbb{Z}\backslash\mathbb{N}$ it holds that $\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}^{\infty}_{j,0}(\vartheta) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{j,0}(\theta)\} = \varnothing$. Combining this and (3.131) with Lemma 3.28 proves that $\liminf_{t\to\infty} \|\Theta_t\| = \infty$. The proof of Corollary 3.29 is thus complete. □

Corollary 3.30. *Assume Setting 3.1, assume $h > 1$, let $k \in \mathbb{N}\backslash\{1\}$ satisfy for all $x \in \mathbb{R}$ that $f(x) = (\max\{x - (\mathscr{a}+\mathscr{b})/2, 0\})^{k-1}$, and let $\Theta \in C([0,\infty), \mathbb{R}^{\mathfrak{d}})$ satisfy $\liminf_{t\to\infty} \mathcal{L}^{\infty}_{k,0}(\Theta_t) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)$ and $\forall\, t \in [0,\infty) \colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}_{k,0}(\Theta_s) \, \mathrm{d}s$. Then $\liminf_{t\to\infty} \|\Theta_t\| = \infty$.*

Proof of Corollary 3.30. Observe that, e.g., [9, Lemma 3.1] assures that for all $t \in [0,\infty)$ it holds that

$$\mathcal{L}^{\infty}_{k,0}(\Theta_t) = \mathcal{L}^{\infty}_{k,0}(\Theta_0) - \int_0^t \|\mathcal{G}_{k,0}(\Theta_s)\|^2 \, \mathrm{d}s. \tag{3.132}$$

Furthermore, note that Lemma 3.23 shows that $\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}^{\infty}_{k,0}(\vartheta) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)\} = \varnothing$. Combining this and (3.132) with Lemma 3.28 demonstrates that $\liminf_{t\to\infty} \|\Theta_t\| = \infty$. The proof of Corollary 3.30 is thus complete. □

Corollary 3.31. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in [\mathscr{a}, \mathscr{b}]$ that $f(x) = \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)$, let $\gamma \in \mathbb{R}\backslash\{1\}$, and let $\Theta \in C([0,\infty), \mathbb{R}^{\mathfrak{d}})$ satisfy $\liminf_{t\to\infty} \mathcal{L}^{\infty}_{1,\gamma}(\Theta_t) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{1,\gamma}(\theta)$ and $\forall\, t \in [0,\infty) \colon \Theta_t = \Theta_0 - \int_0^t \mathcal{G}_{1,\gamma}(\Theta_s) \, \mathrm{d}s$. Then $\liminf_{t\to\infty} \|\Theta_t\| = \infty$.*

Proof of Corollary 3.31. Observe that the assumption that for all $t \in [0,\infty)$ it holds that $\Theta_t = \Theta_0 - \int_0^t \mathcal{G}_{1,\gamma}(\Theta_s) \, \mathrm{d}s$ assures that for all $t \in [0,\infty)$ it holds that

$$\mathcal{L}^{\infty}_{1,\gamma}(\Theta_t) = \mathcal{L}^{\infty}_{1,\gamma}(\Theta_0) - \int_0^t \|\mathcal{G}_{1,\gamma}(\Theta_s)\|^2 \, \mathrm{d}s \tag{3.133}$$

(cf., e.g., Cheridito et al. [9, Lemma 3.5]). Furthermore, note that Lemma 3.6 shows that $\big\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}^{\infty}_{1,\gamma}(\vartheta) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{1,\gamma}(\theta)\big\} = \varnothing$. Combining this and (3.133) with Lemma 3.28 demonstrates that $\liminf_{t\to\infty} \|\Theta_t\| = \infty$. The proof of Corollary 3.31 is thus complete. □

3.9 Divergence of GD

Lemma 3.32. *Let* $\mathfrak{d} \in \mathbb{N}$*,* $\mathcal{L} \in C(\mathbb{R}^{\mathfrak{d}}, \mathbb{R})$ *satisfy* $\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}(\vartheta) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}(\theta)\} = \varnothing$ *and let* $\Theta = (\Theta_n)_{n \in \mathbb{N}_0} \colon \mathbb{N}_0 \to \mathbb{R}^{\mathfrak{d}}$ *satisfy* $\limsup_{n \to \infty} \mathcal{L}(\Theta_n) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}(\theta)$*. Then*

$$\liminf_{n \to \infty} \|\Theta_n\| = \infty. \tag{3.134}$$

Proof of Lemma 3.32. We prove (3.134) by contradiction. We assume that $\liminf_{n \to \infty} \|\Theta_n\| < \infty$. Therefore, by compactness, there exist $\vartheta \in \mathbb{R}^{\mathfrak{d}}$ and a strictly increasing $n \colon \mathbb{N} \to \mathbb{N}$ which satisfies that

$$\limsup_{k \to \infty} \|\Theta_{n(k)} - \vartheta\| = 0. \tag{3.135}$$

Hence, continuity of $\mathcal{L}$ shows that $\limsup_{k \to \infty} |\mathcal{L}(\Theta_{n(k)}) - \mathcal{L}(\vartheta)| = 0$. Combining this with the assumption that $\limsup_{n \to \infty} \mathcal{L}(\Theta_n) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}(\theta)$ proves that

$$\mathcal{L}(\vartheta) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}(\theta). \tag{3.136}$$

This implies that $\vartheta \in \{\theta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}(\theta) = \inf_{v \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}(v)\}$ which is a contradiction. The proof of Lemma 3.32 is thus complete. □

Corollary 3.33. *Assume Setting 3.1, assume* $h > 1$*, assume* $\xi < 3$*, let* $k \in \mathbb{Z} \backslash \mathbb{N}$*, assume for all* $x \in \mathbb{R}$ *that* $f(x) = x^2$*, and let* $\Theta \colon \mathbb{N}_0 \to \mathbb{R}^{\mathfrak{d}}$ *satisfy* $\limsup_{n \to \infty} \mathcal{L}^{\infty}_{k,0}(\Theta_n) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)$*. Then* $\liminf_{n \to \infty} \|\Theta_n\| = \infty$*.*

Proof of Corollary 3.33. Observe that Lemma 3.11, Lemma 3.21, Lemma 3.25, and Lemma 3.27 assure that for all $j \in \mathbb{Z} \backslash \mathbb{N}$ it holds that $\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}^{\infty}_{j,0}(\vartheta) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{j,0}(\theta)\} = \varnothing$. Combining this with Lemma 3.32 proves that $\liminf_{n \to \infty} \|\Theta_n\| = \infty$. The proof of Corollary 3.33 is thus complete. □

Corollary 3.34. *Assume Setting 3.1, assume* $h > 1$*, assume* $\xi < 3$*, let* $k \in \mathbb{Z} \backslash \mathbb{N}$*, assume for all* $x \in \mathbb{R}$ *that* $f(x) = x^2$*, and let* $\Theta \colon \mathbb{N}_0 \to \mathbb{R}^{\mathfrak{d}}$ *satisfy* $\liminf_{n \to \infty} \mathcal{L}^{\infty}_{k,0}(\Theta_n) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)$*. Then* $\limsup_{n \to \infty} \|\Theta_n\| = \infty$*.*

Proof of Corollary 3.34. Note that the assumption that $\liminf_{n \to \infty} \mathcal{L}^{\infty}_{k,0}(\Theta_n) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)$ assures that there exists $n \colon \mathbb{N} \to \mathbb{N}$ which satisfies that

$$\lim_{j \to \infty} \mathcal{L}^{\infty}_{k,0}(\Theta_{n(j)}) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta). \tag{3.137}$$

This and Corollary 3.33 imply that $\liminf_{j \to \infty} \|\Theta_{n(j)}\| = \infty$. Hence, we obtain that $\limsup_{j \to \infty} \|\Theta_j\| = \infty$. The proof of Corollary 3.34 is thus complete. □

Corollary 3.35. *Assume Setting 3.1, assume* $h > 1$*, let* $k \in \mathbb{N} \backslash \{1\}$ *satisfy for all* $x \in \mathbb{R}$ *that* $f(x) = (\max\{x - (\mathscr{a}+\mathscr{b})/2, 0\})^{k-1}$*, and let* $\Theta \colon \mathbb{N}_0 \to \mathbb{R}^{\mathfrak{d}}$ *satisfy* $\limsup_{n \to \infty} \mathcal{L}^{\infty}_{k,0}(\Theta_n) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)$*. Then* $\liminf_{n \to \infty} \|\Theta_n\| = \infty$*.*

Proof of Corollary 3.35. Observe that Lemma 3.23 demonstrates that $\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}^{\infty}_{k,0}(\vartheta) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)\} = \varnothing$. Combining this with Lemma 3.32 shows that $\liminf_{n \to \infty} \|\Theta_n\| = \infty$. The proof of Corollary 3.35 is thus complete. □

Corollary 3.36. *Assume Setting 3.1, assume* $h > 1$*, let* $k \in \mathbb{N} \backslash \{1\}$ *satisfy for all* $x \in \mathbb{R}$ *that* $f(x) = (\max\{x - (\mathscr{a}+\mathscr{b})/2, 0\})^{k-1}$*, and let* $\Theta \colon \mathbb{N}_0 \to \mathbb{R}^{\mathfrak{d}}$ *satisfy* $\liminf_{n \to \infty} \mathcal{L}^{\infty}_{k,0}(\Theta_n) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)$*. Then* $\limsup_{n \to \infty} \|\Theta_n\| = \infty$*.*

Proof of Corollary 3.36. Note that the assumption that $\liminf_{n \to \infty} \mathcal{L}^{\infty}_{k,0}(\Theta_n) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta)$ assures that there exists $n \colon \mathbb{N} \to \mathbb{N}$ which satisfies that

$$\lim_{j \to \infty} \mathcal{L}^{\infty}_{k,0}(\Theta_{n(j)}) = \inf_{\theta \in \mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{k,0}(\theta). \tag{3.138}$$

This and Corollary 3.35 imply that $\liminf_{j \to \infty} \|\Theta_{n(j)}\| = \infty$. Therefore, we obtain that $\limsup_{j \to \infty} \|\Theta_j\| = \infty$. The proof of Corollary 3.36 is thus complete. □

Corollary 3.37. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in [\mathscr{a}, \mathscr{b}]$ that $f(x) = \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)$, let $\gamma \in \mathbb{R}\backslash\{1\}$, and let $\Theta \colon \mathbb{N}_0 \to \mathbb{R}^{\mathfrak{d}}$ satisfy $\limsup_{n\to\infty} \mathcal{L}^{\infty}_{1,\gamma}(\Theta_n) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{1,\gamma}(\theta)$. Then $\liminf_{n\to\infty} \|\Theta_n\| = \infty$.*

Proof of Corollary 3.37. Observe that Lemma 3.6 assures that $\{\vartheta \in \mathbb{R}^{\mathfrak{d}} \colon \mathcal{L}^{\infty}_{1,\gamma}(\vartheta) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{1,\gamma}(\theta)\} = \varnothing$. This and Lemma 3.32 proves that $\liminf_{n\to\infty} \|\Theta_n\| = \infty$. The proof of Corollary 3.37 is thus complete. □

Corollary 3.38. *Assume Setting 3.1, assume $h > 1$, assume for all $x \in [\mathscr{a}, \mathscr{b}]$ that $f(x) = \mathbb{1}_{((\mathscr{a}+\mathscr{b})/2,\infty)}(x)$, let $\gamma \in \mathbb{R}\backslash\{1\}$, and let $\Theta \colon \mathbb{N}_0 \to \mathbb{R}^{\mathfrak{d}}$ satisfy $\liminf_{n\to\infty} \mathcal{L}^{\infty}_{1,\gamma}(\Theta_n) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{1,\gamma}(\theta)$. Then $\limsup_{n\to\infty} \|\Theta_n\| = \infty$.*

Proof of Corollary 3.38. Note that the assumption that $\liminf_{n\to\infty} \mathcal{L}^{\infty}_{1,\gamma}(\Theta_n) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{1,\gamma}(\theta)$ assures that there exists $n \colon \mathbb{N} \to \mathbb{N}$ which satisfies that

$$\lim_{j\to\infty} \mathcal{L}^{\infty}_{1,\gamma}(\Theta_{n(j)}) = \inf_{\theta\in\mathbb{R}^{\mathfrak{d}}} \mathcal{L}^{\infty}_{1,\gamma}(\theta). \tag{3.139}$$

This and Corollary 3.37 imply that $\liminf_{j\to\infty} \|\Theta_{n(j)}\| = \infty$. Hence, we obtain that $\limsup_{j\to\infty} \|\Theta_j\| = \infty$. The proof of Corollary 3.38 is thus complete. □

4 Blow up phenomena for data driven supervised learning problems

In this section we analyze the existence of global minima in the case where the risk is defined using a discrete measure, the activation function is the standard logistic function, and the hidden layer is made up of one neuron. In Lemma 4.7 in Subsection 4.4 and Lemma 4.8 in Subsection 4.4, assuming to have three non-strictly increasing or decreasing and non-constant data points $\mathscr{y}_1, \mathscr{y}_2, \mathscr{y}_3 \in \mathbb{R}$, we prove the non-existence of global minima of the risk function. The proofs of Lemma 4.7 and Lemma 4.8 are based on Proposition 4.5 and on Proposition 4.6. In Proposition 4.5 we find an upper bound for the infimum of the risk assuming that the data points do not coincide, $\max\{|\mathscr{y}_1 - \mathscr{y}_2|, |\mathscr{y}_3 - \mathscr{y}_2|\} > 0$, and are not non-strictly increasing or decreasing, $0 \le (\mathscr{y}_1 - \mathscr{y}_2)(\mathscr{y}_3 - \mathscr{y}_2)$. In Proposition 4.6 we provide a lower bound for the risk in the case where the realization function is constant. The proof of Proposition 4.6 employs the elementary result for the first derivative of the realization function in Proposition 4.4.

In Lemma 4.2 in Subsection 4.2 and Lemma 4.3 in Subsection 4.3 we establish the existence of global minima of the risk function in the case of two data points and in the case of three data points.

4.1 Mathematical description of ANNs

Setting 4.1. *Let $\mathfrak{w}, \mathfrak{b}, \mathfrak{v}, \mathfrak{c} \in C(\mathbb{R}^4, \mathbb{R})$ satisfy for all $\theta = (\theta_1, \dots, \theta_4) \in \mathbb{R}^4$ that $\mathfrak{w}^\theta = \theta_1$, $\mathfrak{b}^\theta = \theta_2$, $\mathfrak{v}^\theta = \theta_3$, and $\mathfrak{c}^\theta = \theta_4$, let $A \colon \mathbb{R} \to \mathbb{R}$ satisfy for all $x \in \mathbb{R}$ that*

$$A(x) = \frac{1}{1 + \exp(-x)}, \tag{4.1}$$

for every $\theta \in \mathbb{R}^4$ let $\mathscr{N}^\theta \colon \mathbb{R} \to \mathbb{R}$ satisfy for all $x \in \mathbb{R}$ that

$$\mathscr{N}^\theta(x) = \mathfrak{c}^\theta + \mathfrak{v}^\theta\big[A(\mathfrak{w}^\theta x + \mathfrak{b}^\theta)\big], \tag{4.2}$$

let $M \in \mathbb{N}$, $\mathscr{x} = (\mathscr{x}_1, \dots, \mathscr{x}_M) \in \mathbb{R}^M$, $\mathscr{y} = (\mathscr{y}_1, \dots, \mathscr{y}_M) \in \mathbb{R}^M$, $\mathcal{L} \in C(\mathbb{R}^4, \mathbb{R})$ satisfy for all $\theta \in \mathbb{R}^4$ that

$$\mathcal{L}(\theta) = \tfrac{1}{M} \textstyle\sum_{i=1}^M \big(\mathscr{N}^\theta(\mathscr{x}_i) - \mathscr{y}_i\big)^2, \tag{4.3}$$

and let $\operatorname{sgn} \colon \mathbb{R} \to \mathbb{R}$ satisfy for all $x \in \mathbb{R}$ that

$$\operatorname{sgn}(x) = \begin{cases} 1 & : x \ge 0 \\ -1 & : x < 0. \end{cases} \tag{4.4}$$

4.2 Existence of global minima for two data points

Lemma 4.2. *Assume Setting 4.1, assume $M = 2$, and assume $\mathscr{x}_1 < \mathscr{x}_2$. Then there exists $\theta \in \mathbb{R}^4$ such that $\mathcal{L}(\theta) = 0$.*

Proof of Lemma 4.2. Throughout this proof let $\vartheta \in \mathbb{R}^4$ satisfy

$$\begin{aligned} \mathfrak{w}^\vartheta &= 1, \qquad \mathfrak{v}^\vartheta = (\mathscr{y}_2 - \mathscr{y}_1)\Big(\frac{1}{1+\exp(-\mathscr{x}_2)} - \frac{1}{1+\exp(-\mathscr{x}_1)}\Big)^{-1}, \\ \mathfrak{b}^\vartheta &= 0, \qquad \text{and} \qquad \mathfrak{c}^\vartheta = \mathscr{y}_1 - \frac{\mathfrak{v}^\vartheta}{1+\exp(-\mathscr{x}_1)}. \end{aligned} \tag{4.5}$$

This implies that

$$\begin{aligned} \mathscr{N}^\vartheta(\mathscr{x}_1) &= \mathscr{y}_1 - \frac{\mathfrak{v}^\vartheta}{1+\exp(-\mathscr{x}_1)} + \frac{\mathfrak{v}^\vartheta}{1+\exp(-\mathscr{x}_1)} = \mathscr{y}_1 \\ \text{and} \qquad \mathscr{N}^\vartheta(\mathscr{x}_2) &= \mathscr{y}_1 - \frac{\mathfrak{v}^\vartheta}{1+\exp(-\mathscr{x}_1)} + \frac{\mathfrak{v}^\vartheta}{1+\exp(-\mathscr{x}_2)} = \mathscr{y}_2. \end{aligned} \tag{4.6}$$

Therefore, we obtain that $\mathcal{L}(\vartheta) = 0$. The proof of Lemma 4.2 is thus complete. □

4.3 Existence of global minima for three data points

Lemma 4.3. *Assume Setting 4.1 and assume $M = 3$, $\mathscr{x}_1 < \mathscr{x}_2 < \mathscr{x}_3$, and $\min\{\mathscr{y}_1, \mathscr{y}_3\} < \mathscr{y}_2 < \max\{\mathscr{y}_1, \mathscr{y}_3\}$. Then there exists $\theta \in \mathbb{R}^4$ such that $\mathcal{L}(\theta) = 0$.*

Proof of Lemma 4.3. Throughout this proof let $f\colon \mathbb{R}^2 \to \mathbb{R}$ satisfy for all $(x_1, x_2) \in \mathbb{R}^2$ that

$$f(x_1, x_2) = \frac{\big(\exp(-\mathscr{x}_1 x_1) - \exp(-\mathscr{x}_3 x_1)\big)\big(1 + \exp(-\mathscr{x}_2 x_1 - x_2)\big)}{\big(\exp(-\mathscr{x}_1 x_1) - \exp(-\mathscr{x}_2 x_1)\big)\big(1 + \exp(-\mathscr{x}_3 x_1 - x_2)\big)}. \tag{4.7}$$

Observe that (4.7) assures that

$$\liminf\nolimits_{x_1 \to \infty} f(x_1, -\mathscr{x}_3 x_1) = \infty \qquad \text{and} \qquad \limsup\nolimits_{x_1 \to -\infty} |f(x_1, 0) - 1| = 0. \tag{4.8}$$

Combining this with intermediate value theorem implies that for all $y \in (1, \infty)$ there exist $x_1, x_2 \in \mathbb{R}$ such that $f(x_1, x_2) = y$. Note that $(\mathscr{y}_3 - \mathscr{y}_1)(\mathscr{y}_2 - \mathscr{y}_1)^{-1} > 1$. Throughout this proof let $\vartheta \in \mathbb{R}^4$ satisfy

$$\begin{aligned} (\mathfrak{w}^\vartheta, \mathfrak{b}^\vartheta) &= f^{-1}\Big(\frac{\mathscr{y}_3 - \mathscr{y}_1}{\mathscr{y}_2 - \mathscr{y}_1}\Big), \qquad \mathfrak{c}^\vartheta = \mathscr{y}_1 - \frac{\mathfrak{v}^\vartheta}{1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta)}, \\ \text{and} \qquad \mathfrak{v}^\vartheta &= (\mathscr{y}_2 - \mathscr{y}_1)\Big(\frac{(1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta))(1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta)}{\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta) - \exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta)}\Big). \end{aligned} \tag{4.9}$$

This shows that

$$\begin{aligned} \mathscr{N}^\vartheta(\mathscr{x}_1) &= \mathscr{y}_1 - \frac{\mathfrak{v}^\vartheta}{1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta)} + \frac{\mathfrak{v}^\vartheta}{1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta)} = \mathscr{y}_1, \\ \mathscr{N}^\vartheta(\mathscr{x}_2) &= \mathscr{y}_1 - (\mathscr{y}_2 - \mathscr{y}_1)\frac{1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta)}{\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta) - \exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta)} \\ &\quad + (\mathscr{y}_2 - \mathscr{y}_1)\frac{1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta)}{\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta) - \exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta)} = \mathscr{y}_2, \qquad \text{and} \\ \mathscr{N}^\vartheta(\mathscr{x}_3) &= \mathscr{y}_1 - (\mathscr{y}_2 - \mathscr{y}_1)\frac{1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta)}{\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta) - \exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta)} + (\mathscr{y}_2 - \mathscr{y}_1) \\ &\quad \frac{(1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta))(1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta)}{(\exp(-\mathfrak{w}^\vartheta \mathscr{x}_1 - \mathfrak{b}^\vartheta) - \exp(-\mathfrak{w}^\vartheta \mathscr{x}_2 - \mathfrak{b}^\vartheta))(1+\exp(-\mathfrak{w}^\vartheta \mathscr{x}_3 - \mathfrak{b}^\vartheta))} \\ &= \mathscr{y}_1 + (\mathscr{y}_2 - \mathscr{y}_1)\frac{\mathscr{y}_3 - \mathscr{y}_1}{\mathscr{y}_2 - \mathscr{y}_1} = \mathscr{y}_3. \end{aligned} \tag{4.10}$$

Hence, we obtain that $\mathcal{L}(\vartheta) = 0$. The proof of Lemma 4.3 is thus complete. □

4.4 Non-existence of global minima for three data points

Proposition 4.4. *Assume Setting 4.1. Then it holds for all $x \in \mathbb{R}$, $\theta \in \mathbb{R}^4$ with $\mathfrak{w}^\theta \mathfrak{v}^\theta \neq 0$ that*

$$\operatorname{sgn}(\mathfrak{w}^\theta \mathfrak{v}^\theta)(\mathscr{N}^\theta)'(x) > 0. \tag{4.11}$$

Proof of Proposition 4.4. Observe that (4.2) ensures that for all $x \in \mathbb{R}$, $\theta \in \mathbb{R}^4$ it holds that

$$(\mathscr{N}^\theta)'(x) = (\mathfrak{w}^\theta \mathfrak{v}^\theta)\frac{\exp(-\mathfrak{w}^\theta x - \mathfrak{b}^\theta)}{(1+\exp(-\mathfrak{w}^\theta x - \mathfrak{b}^\theta))^2}. \tag{4.12}$$

This implies that for all $x \in \mathbb{R}$, $\theta \in \mathbb{R}^4$ with $\mathfrak{w}^\theta \mathfrak{v}^\theta \neq 0$ it holds that

$$\operatorname{sgn}(\mathfrak{w}^\theta \mathfrak{v}^\theta)(\mathscr{N}^\theta)'(x) > 0. \tag{4.13}$$

The proof of Proposition 4.4 is thus complete. □

Proposition 4.5. *Assume Setting 4.1, assume $M = 3$, $\mathscr{x}_1 < \mathscr{x}_2 < \mathscr{x}_3$, and $-\max\{|\mathscr{y}_1 - \mathscr{y}_2|, |\mathscr{y}_3 - \mathscr{y}_2|\} < 0 \leq (\mathscr{y}_1-\mathscr{y}_2)(\mathscr{y}_3-\mathscr{y}_2)$, let $I = \{i \in \{1,3\} \colon |\mathscr{y}_i - \mathscr{y}_2| = \max\{|\mathscr{y}_1-\mathscr{y}_2|, |\mathscr{y}_3-\mathscr{y}_2|\}\}$, let $j,k \in \mathbb{N}$ satisfy $j = \min I$, $k \in \{1,3\}\backslash\{j\}$, and let $(\theta_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^4$ satisfy for all $n \in \mathbb{N}$ that $\mathfrak{w}^{\theta_n} = (2-j)n$, $\mathfrak{b}^{\theta_n} = (j-2)n\mathscr{x}_j$, $\mathfrak{c}^{\theta_n} = 2\mathscr{y}_j - ((\mathscr{y}_2+\mathscr{y}_k)/2)$, and $\mathfrak{v}^{\theta_n} = \mathscr{y}_2 + \mathscr{y}_k - 2\mathscr{y}_j$. Then*

$$\limsup_{n\to\infty}\left|3\mathcal{L}(\theta_n) - 2\Big(\frac{\mathscr{y}_2 - \mathscr{y}_k}{2}\Big)^2\right| = 0. \tag{4.14}$$

Proof of Proposition 4.5. Note that (4.2) ensures thst for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ it holds that

$$\mathscr{N}^{\theta_n}(x) = \frac{\mathscr{y}_2 + \mathscr{y}_k - 2\mathscr{y}_j}{1+\exp(n(j-2)(x-\mathscr{x}_j))} + 2\mathscr{y}_j - \frac{\mathscr{y}_2+\mathscr{y}_k}{2}. \tag{4.15}$$

This implies that for all $n \in \mathbb{N}$ it holds that

$$\begin{aligned}3\mathcal{L}(\theta_n) &= (\mathscr{N}^{\theta_n}(\mathscr{x}_j) - \mathscr{y}_j)^2 + (\mathscr{N}^{\theta_n}(\mathscr{x}_2) - \mathscr{y}_2)^2 + (\mathscr{N}^{\theta_n}(\mathscr{x}_k) - \mathscr{y}_k)^2 \\ &= 0 + \Big(\frac{\mathscr{y}_2 + \mathscr{y}_k - 2\mathscr{y}_j}{1+\exp(n(j-2)(\mathscr{x}_2-\mathscr{x}_j))} + 2\mathscr{y}_j - \frac{3\mathscr{y}_2+\mathscr{y}_k}{2}\Big)^2 \\ &\quad + \Big(\frac{\mathscr{y}_2 + \mathscr{y}_k - 2\mathscr{y}_j}{1+\exp(n(j-2)(\mathscr{x}_k-\mathscr{x}_j))} + 2\mathscr{y}_j - \frac{\mathscr{y}_2+3\mathscr{y}_k}{2}\Big)^2.\end{aligned} \tag{4.16}$$

Therefore, we obtain that

$$\begin{aligned}&\limsup_{n\to\infty}\left|3\mathcal{L}(\theta_n) - 2\Big(\frac{\mathscr{y}_2 - \mathscr{y}_k}{2}\Big)^2\right| \\ &= \left|\Big(\mathscr{y}_2 + \mathscr{y}_k - \frac{3\mathscr{y}_2+\mathscr{y}_k}{2}\Big)^2 + \Big(\mathscr{y}_2 + \mathscr{y}_k - \frac{\mathscr{y}_2+3\mathscr{y}_k}{2}\Big)^2 - 2\Big(\frac{\mathscr{y}_2 - \mathscr{y}_k}{2}\Big)^2\right| \\ &= \left|\Big(\frac{\mathscr{y}_k - \mathscr{y}_2}{2}\Big)^2 + \Big(\frac{\mathscr{y}_2 - \mathscr{y}_k}{2}\Big)^2 - 2\Big(\frac{\mathscr{y}_2 - \mathscr{y}_k}{2}\Big)^2\right| = 0.\end{aligned} \tag{4.17}$$

The proof of Proposition 4.5 is thus complete. □

Proposition 4.6. *Assume Setting 4.1, assume $M = 3$, $\mathscr{x}_1 < \mathscr{x}_2 < \mathscr{x}_3$, and $\max\{|\mathscr{y}_1-\mathscr{y}_2|, |\mathscr{y}_3-\mathscr{y}_2|\} > 0$, let $k \in \mathbb{N}$ satisfy $|\mathscr{y}_k - \mathscr{y}_2| = \min\{|\mathscr{y}_1-\mathscr{y}_2|, |\mathscr{y}_3-\mathscr{y}_2|\}$, and let $\vartheta \in \mathbb{R}^4$ satisfy $\mathfrak{w}^\vartheta \mathfrak{v}^\vartheta = 0$. Then*

$$3\mathcal{L}(\vartheta) > 2\Big(\frac{\mathscr{y}_2 - \mathscr{y}_k}{2}\Big)^2. \tag{4.18}$$

Proof of Proposition 4.6. Observe that the assumption that $\mathfrak{w}^{\vartheta}\mathfrak{v}^{\vartheta} = 0$ assures that there exists $r \in \mathbb{R}$ which satisfies for all $x \in \mathbb{R}$ that $\mathscr{N}^{\vartheta}(x) = r$. This implies that

$$3\mathcal{L}(\vartheta) = (r - \mathscr{y}_1)^2 + (r - \mathscr{y}_2)^2 + (r - \mathscr{y}_3)^2. \tag{4.19}$$

Assume without loss of generality that $k = 1$. This and the assumption that $\max\{|\mathscr{y}_1 - \mathscr{y}_2|, |\mathscr{y}_3 - \mathscr{y}_2|\} > 0$ assure that $|\mathscr{y}_3 - \mathscr{y}_2| > 0$. Hence, we obtain that in the case $r = \mathscr{y}_3$ it holds that

$$3\mathcal{L}(\vartheta) = (\mathscr{y}_3 - \mathscr{y}_1)^2 + (\mathscr{y}_3 - \mathscr{y}_2)^2 \geq (\mathscr{y}_3 - \mathscr{y}_1)^2 + (\mathscr{y}_1 - \mathscr{y}_2)^2 > \frac{1}{2}(\mathscr{y}_2 - \mathscr{y}_1)^2 \tag{4.20}$$

and in the case $r \neq \mathscr{y}_3$ it holds that

$$3\mathcal{L}(\vartheta) = (r - \mathscr{y}_1)^2 + (r - \mathscr{y}_2)^2 + (r - \mathscr{y}_3)^2 > (r - \mathscr{y}_1)^2 + (r - \mathscr{y}_2)^2 \geq \frac{1}{2}(\mathscr{y}_2 - \mathscr{y}_1)^2. \tag{4.21}$$

Combining this and (4.20) shows that

$$3\mathcal{L}(\vartheta) > 2\Big(\frac{\mathscr{y}_2 - \mathscr{y}_1}{2}\Big)^2. \tag{4.22}$$

The proof of Proposition 4.6 is thus complete. ☐

Lemma 4.7. *Assume Setting 4.1 and assume* $M = 3$, $\mathscr{x}_1 < \mathscr{x}_2 < \mathscr{x}_3$, $\max\{|\mathscr{y}_1 - \mathscr{y}_2|, |\mathscr{y}_3 - \mathscr{y}_2|\} > 0$, *and* $\min\{\mathscr{y}_1, \mathscr{y}_3\} \geq \mathscr{y}_2$. *Then*

$$\big\{\vartheta \in \mathbb{R}^4 \colon \mathcal{L}(\vartheta) = \inf\nolimits_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta)\big\} = \varnothing. \tag{4.23}$$

Proof of Lemma 4.7. We prove (4.23) by contradiction. We thus assume that there exists $\vartheta \in \mathbb{R}^4$ such that $\mathcal{L}(\vartheta) = \inf_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta)$ and let $a_1, a_2, a_3 \in \mathbb{R}$ satisfy for all $n \in \{1, 2, 3\}$ that $\mathscr{N}^{\vartheta}(\mathscr{x}_n) = a_n$. Note that Proposition 4.5 implies that

$$\mathcal{L}(\vartheta) = \inf\nolimits_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta) \leq \frac{2}{3}\Big(\frac{\mathscr{y}_2 - \min\{\mathscr{y}_1, \mathscr{y}_3\}}{2}\Big)^2. \tag{4.24}$$

This and Proposition 4.6 show that $\mathfrak{w}^{\vartheta}\mathfrak{v}^{\vartheta} \neq 0$. Combining this with Proposition 4.4 demonstrates that for all $x \in \mathbb{R}$ it holds that $(\mathscr{N}^{\vartheta})'(x) \neq 0$. In the following we distinguish between the case $\min_{x \in [\mathscr{x}_1, \mathscr{x}_3]}(\mathscr{N}^{\vartheta})'(x) > 0$ and the case $\max_{x \in [\mathscr{x}_1, \mathscr{x}_3]}(\mathscr{N}^{\vartheta})'(x) < 0$. We first establish the contradiction in the case

$$\min\nolimits_{x \in [\mathscr{x}_1, \mathscr{x}_3]}(\mathscr{N}^{\vartheta})'(x) > 0. \tag{4.25}$$

Observe that (4.25) assures that $a_1 < a_2 < a_3$. Combining this and the assumption that $\min\{\mathscr{y}_1, \mathscr{y}_3\} \geq \mathscr{y}_2$ proves that

$$(a_1 - \mathscr{y}_1)^2 + (a_2 - \mathscr{y}_2)^2 > \begin{cases} (a_1 - \mathscr{y}_1)^2 + (a_1 - \mathscr{y}_2)^2 & : a_1 \geq \mathscr{y}_2 \\ (\mathscr{y}_1 - \mathscr{y}_2)^2 + (a_2 - \mathscr{y}_2)^2 & : a_1 < \mathscr{y}_2. \end{cases} \tag{4.26}$$

This implies that

$$3\mathcal{L}(\vartheta) \geq (a_1 - \mathscr{y}_1)^2 + (a_2 - \mathscr{y}_2)^2 > 2\Big(\frac{\mathscr{y}_2 - \mathscr{y}_1}{2}\Big)^2 \geq 2\Big(\frac{\mathscr{y}_2 - \min\{\mathscr{y}_1, \mathscr{y}_3\}}{2}\Big)^2. \tag{4.27}$$

Combining this with (4.24) shows that $\mathcal{L}(\vartheta) > \inf_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta)$ which is a contradiction. In the next step we establish the contradiction in the case

$$\max\nolimits_{x \in [\mathscr{x}_1, \mathscr{x}_3]}(\mathscr{N}^{\vartheta})'(x) < 0. \tag{4.28}$$

Note that (4.28) assures that $a_1 > a_2 > a_3$. Combining this and the assumption that $\min\{y_1, y_3\} \geq y_2$ proves that

$$(a_2 - y_2)^2 + (a_3 - y_3)^2 > \begin{cases} (a_3 - y_2)^2 + (a_3 - y_3)^2 & : a_3 \geq y_2 \\ (a_2 - y_2)^2 + (y_2 - y_3)^2 & : a_3 < y_2. \end{cases} \tag{4.29}$$

This implies that

$$3\mathcal{L}(\vartheta) \geq (a_2 - y_2)^2 + (a_3 - y_3)^2 > 2\Big(\frac{y_2 - y_3}{2}\Big)^2 \geq 2\Big(\frac{y_2 - \min\{y_1, y_3\}}{2}\Big)^2. \tag{4.30}$$

Combining this with (4.24) shows that $\mathcal{L}(\vartheta) > \inf_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta)$ which is a contradiction. The proof of Lemma 4.7 is thus complete. ☐

Lemma 4.8. *Assume Setting 4.1 and assume $M = 3$, $x_1 < x_2 < x_3$, $\max\{|y_1 - y_2|, |y_3 - y_2|\} > 0$, $\max\{y_1, y_3\} \leq y_2$. Then*

$$\big\{\vartheta \in \mathbb{R}^4 \colon \mathcal{L}(\vartheta) = \inf_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta)\big\} = \varnothing. \tag{4.31}$$

Proof of Lemma 4.8. We prove (4.31) by contradiction. Assume that there exists $\vartheta \in \mathbb{R}^4$ such that $\mathcal{L}(\vartheta) = \inf_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta)$ and let $a_1, a_2, a_3 \in \mathbb{R}$ satisfy for all $n \in \{1, 2, 3\}$ that $\mathscr{N}^{\vartheta}(x_n) = a_n$. Observe that Proposition 4.5 implies that

$$\mathcal{L}(\vartheta) = \inf_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta) \leq \frac{2}{3}\Big(\frac{y_2 - \max\{y_1, y_3\}}{2}\Big)^2. \tag{4.32}$$

This and Proposition 4.6 show that $\mathfrak{w}^{\vartheta}\mathfrak{v}^{\vartheta} \neq 0$. Combining this with Proposition 4.4 demonstrates that for all $x \in \mathbb{R}$ it holds that $(\mathscr{N}^{\vartheta})'(x) \neq 0$. In the following we distinguish between the case $\min_{x \in [x_1, x_3]} (\mathscr{N}^{\vartheta})'(x) > 0$ and the case $\max_{x \in [x_1, x_3]} (\mathscr{N}^{\vartheta})'(x) < 0$. We first establish the contradiction in the case

$$\min_{x \in [x_1, x_3]} (\mathscr{N}^{\vartheta})'(x) > 0. \tag{4.33}$$

Note that (4.33) assures that $a_1 < a_2 < a_3$. Combining this and the assumption that $\max\{y_1, y_3\} \leq y_2$ proves that

$$(a_2 - y_2)^2 + (a_3 - y_3)^2 > \begin{cases} (a_2 - y_2)^2 + (y_3 - y_2)^2 & : a_3 \geq y_2 \\ (a_3 - y_2)^2 + (a_3 - y_3)^2 & : a_3 < y_2. \end{cases} \tag{4.34}$$

This implies that

$$3\mathcal{L}(\vartheta) \geq (a_2 - y_2)^2 + (a_3 - y_3)^2 > 2\Big(\frac{y_2 - y_3}{2}\Big)^2 \geq 2\Big(\frac{y_2 - \max\{y_1, y_3\}}{2}\Big)^2. \tag{4.35}$$

Combining this with (4.32) shows that $\mathcal{L}(\vartheta) > \inf_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta)$ which is a contradiction. In the next step we establish the contradiction in the case

$$\max_{x \in [x_1, x_3]} (\mathscr{N}^{\vartheta})'(x) < 0. \tag{4.36}$$

Observe that (4.36) assures that $a_1 > a_2 > a_3$. Combining this and the assumption that $\max\{y_1, y_3\} \leq y_2$ proves that

$$(a_1 - y_1)^2 + (a_2 - y_2)^2 > \begin{cases} (a_1 - y_1)^2 + (a_1 - y_2)^2 & : a_1 \leq y_2 \\ (y_2 - y_1)^2 + (a_2 - y_2)^2 & : a_1 > y_2. \end{cases} \tag{4.37}$$

This implies that

$$3\mathcal{L}(\vartheta) \geq (a_1 - y_1)^2 + (a_2 - y_2)^2 > 2\Big(\frac{y_2 - y_1}{2}\Big)^2 \geq 2\Big(\frac{y_2 - \max\{y_1, y_3\}}{2}\Big)^2. \tag{4.38}$$

Combining this with (4.32) shows that $\mathcal{L}(\vartheta) > \inf_{\theta \in \mathbb{R}^4} \mathcal{L}(\theta)$ which is a contradiction. The proof of Lemma 4.8 is thus complete. ☐

Acknowledgements

Adrian Riekert is gratefully acknowledged for several useful comments. The second author acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044-390685587, Mathematics Muenster: Dynamics-Geometry-Structure.

References

[1] Pierre-Antoine Absil, Robert Mahony, and Ben Andrews. Convergence of the iterates of descent methods for analytic cost functions. *SIAM Journal on Optimization*, 16(2):531–547, 2005. doi:10.1137/040605266.

[2] Ömer Deniz Akyildiz and Sotirios Sabanis. Nonasymptotic analysis of Stochastic Gradient Hamiltonian Monte Carlo under local conditions for nonconvex optimization, 2021. arXiv:2002.05465.

[3] Zeyuan Allen-Zhu, Yuanzhi Li, and Yingyu Liang. Learning and generalization in overparameterized neural networks, going beyond two layers. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32, pages 6158–6169. Curran Associates, Inc., 2019. URL: https://proceedings.neurips.cc/paper/2019/file/62dad6e273d32235ae02b7d321578ee8-Paper.pdf.

[4] Zeyuan Allen-Zhu, Yuanzhi Li, and Zhao Song. A convergence theory for deep learning via over-parameterization. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 242–252. PMLR, 09–15 Jun 2019. URL: http://proceedings.mlr.press/v97/allen-zhu19a.html.

[5] Hedy Attouch and Jérôme Bolte. On the convergence of the proximal algorithm for nonsmooth functions involving analytic features. *Math. Program.*, 116(1-2, Ser. B):5–16, 2009. doi:10.1007/s10107-007-0133-5.

[6] Francis Bach and Eric Moulines. Non-strongly-convex smooth stochastic approximation with convergence rate $O(1/n)$. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26, pages 773–781. Curran Associates, Inc., 2013. URL: http://papers.nips.cc/paper/4900-non-strongly-convex-smooth-stochastic-approximation-with-convergence-rate-o1n.pdf.

[7] Dimitri P. Bertsekas and John N. Tsitsiklis. Gradient convergence in gradient methods with errors. *SIAM Journal on Optimization*, 10(3):627–642, 2000. doi:10.1137/S1052623497331063.

[8] Patrick Cheridito, Arnulf Jentzen, and Florian Rossmannek. Landscape analysis for shallow ReLU neural networks: complete classification of critical points for affine target functions. *J Nonlinear Sci* , 32, 64, 2022. doi:10.1007/s00332-022-09823-8.

[9] Patrick Cheridito, Arnulf Jentzen, Adrian Riekert, and Florian Rossmannek. A proof of convergence for gradient descent in the training of artificial neural networks for constant target functions. *Journal of Complexity*, 101646, 2022.

[10] Patrick Cheridito, Arnulf Jentzen, and Florian Rossmannek. Non-convergence of stochastic gradient descent in the training of deep neural networks. *Journal of Complexity*, 101540, 2020.

[11] Steffen Dereich and Thomas Müller-Gronbach. General multilevel adaptations for stochastic approximation algorithms of Robbins-Monro and Polyak-Ruppert type. *Numer. Math.*, 142(2):279–328, 2019. doi:10.1007/s00211-019-01024-y.

[12] Steffen Dereich and Sebastian Kassing. Convergence of stochastic gradient descent schemes for Lojasiewicz-landscapes, 2021. arXiv:2102.09385.

[13] Simon Du, Jason Lee, Haochuan Li, Liwei Wang, and Xiyu Zhai. Gradient descent finds global minima of deep neural networks. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 1675–1685, Long Beach, California, USA, 6 2019. PMLR. URL: http://proceedings.mlr.press/v97/du19c.html.

[14] Simon Du, Xiyu Zhai, Barnabas Poczos, and Aarti Singh. Gradient descent provably optimizes over-parameterized neural networks. In *International Conference on Learning Representations*, 2019. URL: https://openreview.net/forum?id=S1eK3i09YQ.

[15] Simon Eberle, Arnulf Jentzen, Adrian Riekert, Georg S. Weiss. Existence, uniqueness, and convergence rates for gradient flows in the training of artificial neural networks with ReLU activation, 2021. arXiv:2108.08106.

[16] Benjamin Fehrman, Benjamin Gess, and Arnulf Jentzen. Convergence rates for the stochastic gradient descent method for non-convex objective functions. *J. Mach. Learn. Res.*, 21:Paper No. 136, 48, 2020.

[17] Arnulf Jentzen and Philippe von Wurstemberger. Lower error bounds for the stochastic gradient descent optimization algorithm: Sharp convergence rates for slowly and fast decaying learning rates. *Journal of Complexity 57*, 101438, 2021.

[18] Arnulf Jentzen, Benno Kuckuck, Ariel Neufeld, and Philippe von Wurstemberger. Strong error analysis for stochastic gradient descent optimization algorithms. *IMA J. Numer. Anal.*, 41(1):455–492, 2021. doi:10.1093/imanum/drz055.

[19] Arnulf Jentzen and Timo Kröger. Convergence rates for gradient descent in the training of overparameterized artificial neural networks with biases, 2021. arXiv:2102.11840.

[20] Arnulf Jentzen and Adrian Riekert. On the existence of global minima and convergence analyses for gradient descent methods in the training of deep neural networks. *Journal of Machine Learning*, 1(2):141–246, 2022.

[21] Arnulf Jentzen and Adrian Riekert. A proof of convergence for the gradient descent optimization method with random initializations in the training of neural networks with ReLU activation for piecewise linear target functions. Accepted in *J. Mach. Learn. Res.*, 2021. arXiv:2108.04620.

[22] Arnulf Jentzen and Adrian Riekert. A proof of convergence for stochastic gradient descent in the training of artificial neural networks with ReLU activation for constant target functions. *Z. Angew. Math. Phys.*, 73, 188, 2022. doi:10.1007/s00033-022-01716-w.

[23] Arnulf Jentzen and Adrian Riekert. Convergence analysis for gradient flows in the training of artificial neural networks with ReLU activation. *J. Math. Anal. Appl.*, 517(2): 126601, 2022. doi:10.1016/j.jmaa.2022.126601.

[24] Clemens Karner, Vladimir Kazeev, Philipp Christian Petersen. Limitations of neural network training due to numerical instability of backpropagation, 2022. arXiv:2210.00805.

[25] Yunwen Lei, Ting Hu, Guiying Li, and Ke Tang. Stochastic gradient descent for nonconvex learning without bounded gradient assumptions. *IEEE Transactions on Neural Networks and Learning Systems*, 31(10):4394–4400, 2020. doi:10.1109/TNNLS.2019.2952219.

[26] Yuanzhi Li and Yingyu Liang. Learning overparameterized neural networks via stochastic gradient descent on structured data. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31, pages 8157–8166. Curran Associates, Inc., 2018. URL: https://proceedings.neurips.cc/paper/2018/file/54fe976ba170c19ebae453679b362263-Paper.pdf.

[27] Attila Lovas, Iosif Lytras, Miklós Rásonyi, and Sotirios Sabanis. Taming neural networks with TUSLA: Non-convex learning via adaptive stochastic gradient Langevin algorithms, 2020. arXiv:2006.14514.

[28] Lu Lu, Yeonjong Shin, Yanhui Su, and George Em Karniadakis. Dying ReLU and initialization: Theory and numerical examples. *Communications in Computational Physics*, 28(5):1671–1706, 2020. doi:10.4208/cicp.OA-2020-0165.

[29] Eric Moulines and Francis Bach. Non-asymptotic analysis of stochastic approximation algorithms for machine learning. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 24, pages 451–459. Curran Associates, Inc., 2011. URL: https://proceedings.neurips.cc/paper/2011/file/40008b9a5380fcacce3976bf7c08af5b-Paper.pdf.

[30] Philipp Petersen, Mones Raslan, Felix Voigtlaender. Topological properties of the set of functions generated by neural networks of fixed size. *Foundations of Computational Mathematics*, 2020.

[31] Weinan E, Chao Ma, Stephan Wojtowytsch, and Lei Wu. Towards a mathematical understanding of neural network-based machine learning: what we know and what we don't, 2020. arXiv:2009.10713.