Trends and Applications in Constructive Approximation
(Eds.) M.G. de **Bruin**, D.H. **Mache** & J. **Szabados**
International Series of Numerical Mathematics Vol. 1??

A Markov–Bernstein inequality for Gaussian networks

H. N. Mhaskar

Abstract

Let $s \geq 1$ be an integer. A Gaussian network is a function on $\mathbb{R}^s$ of the form $g(\mathbf{x}) = \sum_{k=1}^{N} a_k \exp(-\|\mathbf{x} - \mathbf{x}_k\|^2)$. The minimal separation among the centers, defined by $\min_{1\leq j\neq k\leq N} \|\mathbf{x}_j - \mathbf{x}_k\|$, is an important characteristic of the network that determines the stability of interpolation by Gaussian networks, the degree of approximation by such networks, etc. We prove that if $g(\mathbf{x}) = \sum_{k=1}^{N} a_k \exp(-\|\mathbf{x} - \mathbf{x}_k\|^2)$, the minimal separation of g exceeds $1/m$, and $\log N = \mathcal{O}(m^2)$ then for any integer $r \geq 1$, any partial derivative $\mathcal{D}g$ of order r of g satisfies $\|\mathcal{D}g\|_{p,\mathbb{R}^s} \leq cm^r\|g\|_{p,\mathbb{R}^s}$.

1 Introduction

Let $s, N \geq 1$ be integers. A *Gaussian network* with N *neurons* is a function on the Euclidean space $\mathbb{R}^s$ of the form $\mathbf{x} \mapsto \sum_{k=1}^{N} a_k \exp(-\|\mathbf{x}-\mathbf{x}_k\|^2)$, where $\|\circ\|$ denotes the Euclidean norm on $\mathbb{R}^s$, the *centers* $\mathbf{x}_k$ are in $\mathbb{R}^s$, and $a_k \in \mathbb{R}$, $k = 1, \cdots, N$. These functions can be evaluated in hardware using parallel computation of the exponential terms, and are used extensively in many applications in pattern recognition, computer graphics, antenna array theory, probability density estimation, etc. A typical problem in all these applications is to approximate an unknown function (the *target function*) by such networks.

An important characteristic of Gaussian networks is the minimal separation among the centers, defined by $\min_{1\leq j\neq k\leq N} \|\mathbf{x}_j - \mathbf{x}_k\|$. Many results in the theory of stability of interpolation by Gaussian networks, the degree of approximation by such networks, etc. depend upon the minimal separation. For example, Narcowich and Ward [9] have estimated the condition numbers of the interpolation matrices in the context of a general scattered data interpolation. Their estimates are in terms of the minimal separation between the interpolation points, independent of the number of points (and hence, of neurons) involved. In [7], we have argued that treating the minimal separation among the centers as the "cost of approximation" (rather than the more apparent cost in terms of the number of neurons) leads to matching direct and converse theorems in the theory of approximation by Gaussian networks. In particular, under certain conditions, if a function can be approximated by Gaussian networks at a polynomial rate, measured in terms of the minimal separation of the networks, then it can also be approximated at the same rate by the linear processes of weighted polynomial approximation.

The purpose of this paper is to prove a Markov-Bernstein inequality for Gaussian networks in terms of the minimal separation. We note that such an inequality was obtained by Erdélyi [3] in terms of the number of neurons. Also, if $r \ge 1$ is an integer, and $\mathcal{D}$ is a partial derivative operator of order r, our results in [7] immediately yield an inequality of the form $\|\mathcal{D}g\|_{p,\mathbb{R}^s} \le c\exp(Am^2)\|g\|_{p,\mathbb{R}^s}$ for networks g where the minimal separation exceeds $1/m$. In this paper, we prove a substantially better inequality of the form $\|\mathcal{D}g\|_{p,\mathbb{R}^s} \le cm^r\|g\|_{p,\mathbb{R}^s}$ for such networks, provided that the number of neurons is not too large (cf. Theorem 2.1 below). Our proof involves a good deal of book-keeping in estimating the degree of weighted polynomial approximation in a more careful way than what is available in the literature that we are aware of so far.

In Section 2, we formulate our main result (Theorem 2.1) regarding Gaussian networks. In Section 3, we discuss the background and prove the necessary new results on weighted polynomial approximation. In Section 4, we review some results regarding Gaussian networks, and prove Theorem 2.1.

2 Main result

Let $s \ge 1$ be an integer. The notation for the class of Gaussian networks will involve different bounds on the centers as well as the number of neurons involved. Thus, for $m, M, N > 0$, the symbol $\mathbb{G}_{N,M,m,s}$ denotes the class of functions of the form

$$\mathbf{x} \mapsto \sum_{1\le k\le N,\ k\in\mathbb{Z}} a_k \exp(-\|\mathbf{x}-\mathbf{x}_k\|^2), \qquad \mathbf{x}, \mathbf{x}_k \in \mathbb{R}^s,\ a_k \in \mathbb{R},\ 1 \le k \le N, \tag{1}$$

where $\max_{1\le k\le N}\|\mathbf{x}_k\| \le M$ and the *minimal separation,* $\min\limits_{1\le k,j\le N, k\ne j}\|\mathbf{x}_j - \mathbf{x}_k\| \ge m^{-1}$. Also, the union of the class of networks over a certain parameter will be denoted by writing the symbol ∞ in place of that parameter; for example, $\mathbb{G}_{N,\infty,m,s} := \cup_{M>0}\mathbb{G}_{N,M,m,s}$, etc. For $A, C, m > 0$, we write

$$\mathbb{B}(A, C; m, s) := \{g \in \mathbb{G}_{N,\infty,m,s} \ :\ N \le C\exp(Am^2)\}. \tag{2}$$

We remark that if $A_1 \ge A$, $C_1 \ge C$, $m_1 \ge m$, then $\mathbb{B}(A, C; m, s) \subseteq \mathbb{B}(A_1, C_1; m_1, s)$.

If $1 \le p \le \infty$, $f : \mathbb{R}^s \to \mathbb{R}$ is a Lebesgue measurable function, and $S \subseteq \mathbb{R}^s$ is a Lebsegue measurable set having positive measure, we write

$$\|f\|_{p,S} := \begin{cases} \{\int_S |f(\mathbf{x})|^p d\mathbf{x}\}^{1/p}, & \text{if } 1 \le p < \infty, \\ \text{ess sup}_{\mathbf{x}\in S}|f(\mathbf{x})|, & \text{if } p = \infty. \end{cases} \tag{3}$$

The set of all functions for which $\|f\|_{p,S} < \infty$ is denoted by $L^p(S)$, where, as usual, two functions that are equal almost everywhere on S are considered equal as elements of $L^p(S)$. Let D_j denote the operation of partial differentiation with respect the j-th variable, and $\mathbf{D}^{\mathbf{k}} := \prod_{j=1}^s D_j^{k_j}$. For a suitably smooth function f, we write

$$\|f\|_{p,r,S} := \sum_{|\mathbf{k}|\le r} \|\mathbf{D}^{\mathbf{k}} f\|_{p,S}. \tag{4}$$

In the sequel, we adopt the following convention regarding constants. The symbols $c, c_1, \cdots$ will denote positive constants depending only on A, C, s, p, r, and other similarly fixed parameters, but their values may be different at different occurrences, even within a single formula. Constants denoted by capital letters retain their values, subject to the choice of the parameters on which they depend.

Our main theorem in this paper is the following.

Theorem 2.1 *Let $1 \le p \le \infty$, $m \ge 1$, $A, C > 0$, $s, r \ge 1$ be integers. Then there exists a positive constant c depending only on A, C, p, r, and s, such that*

$$\|g\|_{p,r,\mathbb{R}^s} \le cm^r \|g\|_{p,\mathbb{R}^s}, \qquad g \in \mathbb{B}(A, C; m, s). \tag{5}$$

The idea behind the proof of Theorem 2.1 is the following. In [7], we have established a connection between the ℓ^1 norm of the coefficients of a Gaussian network, and the norm of this network. Using the partial sums of the series in (43), we will approximate the basic Gaussian $\exp(-\|\circ - \mathbf{x}_k\|^2)$ by weighted polynomials, and hence, approximate the network by weighted polynomials. Unfortunately, this can be done adequately only if $\|\mathbf{x}_k\| \le cm$. Therefore, we use a partition of unity so that the norms of different networks and their derivatives are essentially confined to cubes with side proportional to m. However, this involves estimating a norm of the form $\|\phi g\|_{p,r,\mathbb{R}^s}$ for a compactly supported ϕ. Here, we use Theorem 3.2 below to approximate ϕ by weighted polynomials, the partial sums of the series in (43) to approximate the part of g with centers in a cube of side proportional to m by weighted polynomials as well, and estimate the remaining part of g using the results in [7]. A repeated application of the Markov-Bernstein inequality (8) enables us to estimate the derivatives of the approximating weighted polynomials thus obtained in terms of their norms. A reverse process then takes us to (5).

3 Weighted polynomial approximation

In this section, we write $w(x) := \exp(-x^2)$. Our results here are in the univariate case, $s = 1$, but can be extended easily to the multivariate case by a simple tensor product argument. For $x \geq 0$, let Π_x denote the class of all univariate algebraic polynomials of degree at most x. First, we recall a few properties of polynomials. The following proposition will be used often, sometimes without an explicit reference.

Proposition 3.1 *Let $m \geq 0$, $1 \leq p \leq \infty$, $\lambda > 0$, and $P \in \Pi_{m^2}$.*
(a) (Infinite-finite range inequality) *For any $\gamma > 0$, there exists $a > 1$, depending only on γ and λ such that*

$$\|w^\lambda P\|_{p,\mathbb{R}\setminus[-am,am]} \leq c(\lambda,\gamma)\exp(-\gamma^2 m^2)\|w^\lambda P\|_{p,[-am,am]}. \tag{6}$$

In particular,

$$\|xw^\lambda P\|_{p,\mathbb{R}} \leq c(\lambda) m \|w^\lambda P\|_{p,\mathbb{R}}. \tag{7}$$

(b) (Markov-Bernstein inequality)

$$\|(w^\lambda P)'\|_{p,\mathbb{R}} \leq c(\lambda) m \|w^\lambda P\|_{p,\mathbb{R}}. \tag{8}$$

Proof. The inequality (6) follows from [6, Theorem 6.2.4, Lemma 7.2.2]. The inequality (7) is then clear. Part (b) follows from [6, Theorem 6.2.9, Theorem 3.4.2] and (7). ■

In this section, we adopt the following notation. The space of all 2π-periodic continuous functions on $\mathbb{R}$, equipped with the norm $\|F\|^* := \|F\|_{\infty,[-\pi,\pi]}$, will be denoted by C^*, the class of all trigonometric polynomials of order at most n will be denoted by $\mathbf{H}_n$, and

$$E_n^*(F) := \inf_{T\in\mathbf{H}_n} \|F - T\|^*.$$

Our main theorem in this section is the following.

Theorem 3.2 *Let $0 < a < 1$, $f : \mathbb{R} \to \mathbb{R}$ be a continuous function, with $f(x) = 0$ if $|x| \geq a$. Let $F(\theta) := f(3\cos(2\theta))$, $\theta \in \mathbb{R}$. There exists a sequence of polynomials $P_n \in \Pi_n$ such that*

$$\max_{x\in\mathbb{R}} |f(x) - P_n(x)\exp(-nx^2)| \leq c_1 E^*_{c\sqrt{n}}(F) + \exp(-c_2\sqrt{n})\|f\|_{\infty,\mathbb{R}}, \tag{9}$$

where c, c_1, c_2 are positive constants depending only on a.

We note that the fact that the left hand side of (9) tends to zero has been known for a very long time, and has been generalized a great deal (cf. [6, 1] and references therein). The novelty here is the rate of convergence. Our proof consists of a book keeping in the proof of Theorem 10.1.1 in [6], obtaining an intermediate polynomial involved in that proof using the following Theorem 3.3 of Gaier [4] instead of Jackson's theorem. The same proof can also be adapted at least to the more general situation discussed in [6, Section 10.1]. However, we do not wish to introduce here the additional notation that would be necessary to formulate this general version. We will give a simple proof of Theorem 3.3 which does not involve complex variable techniques, and yields estimates using the norm of the functions on the interval rather than unspecified constants depending upon the functions.

Theorem 3.3 *There exists a sequence of linear operators $\mathcal{G}_n$ on $C[-3,3]$, such that for each $f \in C[-3,3]$, and integer $n \ge 1$, $\mathcal{G}_n(f) \in \Pi_n$, and satisfies the following conditions: Writing $F(t) := f(3\cos(2t))$ ($F \in C^*$),*

$$\|f - \mathcal{G}_n(f)\|_{\infty,[-3,3]} \le c_1 E^*_{n/3}(F) + c_2 \exp(-c_3 n)\|f\|_{\infty,[-3,3]}, \tag{10}$$

where c_1, c_2, c_3 are absolute positive constants. Further, if $a \in (0,3)$, $f(x) = 0$ for $|x| \ge a$, and $a < b < 3$, then

$$\|\mathcal{G}_n(f)\|_{\infty,[-3,-b]\cup[b,3]} \le c_4 \exp(-cn)\|f\|_{\infty,[-3,3]}, \tag{11}$$

where c_4, c are positive constants depending only on a and b.

Proof. We recall (cf. [2, Chapter 9, Theorem 3.1]) that the expression

$$v^*_\ell(t) := \frac{1}{\ell}\sum_{m=\ell}^{2\ell-1}\sum_{|k|\le m} e^{ikt} = \frac{\cos \ell t - \cos(2\ell t)}{2\ell \sin^2(t/2)}$$

is an even, trigonometric polynomial of order at most $2\ell - 1$, the operator

$$V^*_\ell(F,x) := \frac{1}{2\pi}\int_{-\pi}^{\pi} F(t) v^*_\ell(x-t)dt, \qquad F \in C^*, \tag{12}$$

satisfies $\|V^*_\ell(F)\|^* \le c\|F\|^*$, and

$$E^*_{2\ell}(F) \le \|F - V^*_\ell(F)\|^* \le cE^*_\ell(F), \qquad F \in C^*. \tag{13}$$

We write $V^*_0(F) := \frac{1}{2\pi}\int_{-\pi}^{\pi} F(t)dt$, $v^*_0(t) := 1$, and

$$\mathcal{G}^*_n(F) = 2^{-n}\sum_{\ell=0}^{n}\binom{n}{\ell} V^*_\ell(F). \tag{14}$$

Since [5, 8]

$$2^{-n}\sum_{\ell=0}^{\lfloor n/3\rfloor-1}\binom{n}{\ell} \le c_1 \exp(-cn),$$

we conclude that

$$\|F - \mathcal{G}^*_n(F)\|^* \le c_1 \exp(-cn)\|F\|^* + c_2 E^*_{n/3}(F). \tag{15}$$

Next, we observe that for $\ell \geq 1$,

$$v_\ell^*(t) = \frac{1}{2\sin^2(t/2)} \int_t^{2t} \sin(\ell u)du.$$

Therefore,

$$\begin{aligned} 2^{-n} \sum_{\ell=1}^{n} \binom{n}{\ell} v_\ell^*(t) &= \frac{1}{2^{n+1}\sin^2(t/2)} \int_t^{2t} \left\{ \sum_{\ell=0}^{n} \binom{n}{\ell} \sin(\ell u) \right\} du \\ &= \frac{1}{2\sin^2(t/2)} \int_t^{2t} \sin(nu/2)\cos^n(u/2)du \\ &= \frac{1}{\sin^2(t/2)} \int_{t/2}^{t} \sin(nu)\cos^n(u)du. \end{aligned}$$

Thus,

$$\left| 2^{-n} \sum_{\ell=1}^{n} \binom{n}{\ell} v_\ell^*(t) \right| \leq c_1(\delta)\exp(-c(\delta)n), \qquad t \in [\delta, \pi-\delta].$$

Now, let $F(t) = 0$ for $t \notin [\delta, \pi-\delta]$ (and hence, also for $t \notin [-\pi+\delta, -\delta]$). Then for $0 < \delta_1 < \delta$, and $|x| \leq \delta_1$,

$$\begin{aligned} |\mathcal{G}_n^*(F,x)| &\leq c \int_{\delta-\delta_1 \leq |t| \leq \pi-\delta+\delta_1} |F(x-t)| \left| \left\{ 2^{-n} \sum_{\ell=0}^{n} \binom{n}{\ell} v_\ell^*(t) \right\} \right| dt \\ &\leq c_1(\delta,\delta_1)\exp(-c(\delta,\delta_1)n)\|F\|^*. \end{aligned} \tag{16}$$

Now, we observe that $F(t) = F(-t) = F(\pi - t)$ for $t \in \mathbb{R}$. Therefore, since each v_ℓ^* is an even function, it is not difficult to see that for $x \in \mathbb{R}$

$$\begin{aligned} 2\pi V_\ell^*(F, \pi - x) &= \int_0^\pi (F(\pi-x-t) + F(\pi-x+t))v_\ell^*(t)dt \\ &= \int_0^\pi (F(x+t) + F(x-t))v_\ell^*(t)dt = 2\pi V_\ell^*(F,x). \end{aligned}$$

Hence, $\mathcal{G}_n^*(F,x)$ is a linear combination of $\cos 2kx$, $k = 0, \cdots, n$. Thus, the operator defined by $\mathcal{G}_n(f, 3\cos 2\theta) := \mathcal{G}_n^*(F,\theta)$ satisfies $\mathcal{G}_n(f) \in \Pi_n$. The estimates (15) and (16) lead to (10) and (11) respectively. ∎

We resume our proof of Theorem 3.2 as in [6, Section 10.1], sketching only enough details to make the paper self-sufficient, and to point out the necessary differences.

Let T_n be the extremal polynomial satisfying

$$\|w\mathsf{T}_n\|_{\infty,\mathbb{R}} = \inf_{P\in\Pi_{n-1}} \|((\circ)^n - P)w\|_{\infty,\mathbb{R}}$$

Let

$$\xi_n = \max\{\xi \in \mathbb{R} \ : \ w(\xi)\mathsf{T}_n(\xi) = \|w\mathsf{T}_n\|_{\infty,\mathbb{R}}\},$$

and $Z_n = \xi_n/\sqrt{n}$. It is known [6, Chapter 6] that $Z_n \in (-1,1)$, $\lim_{n\to\infty} Z_n = 1$, and that T_n has n zeros in $[-\sqrt{n}, \sqrt{n}]$. In this section only, we will write $T_n(x) := n^{-n/2}\mathsf{T}_n(\sqrt{n}x)$, $\mathcal{E}_n := T_n(Z_n)w(\xi_n)$ (our notation here being different from that in [6]).

For $\epsilon > 0$ and integer $n \ge 1$, let

$$\begin{aligned}
\Gamma_{n,1,\epsilon} &:= \{Z_n + (1-Z_n)\frac{y^2}{\epsilon^2} + iy \ : \ 0 \le y \le \epsilon\},\\
\Gamma_{n,2,\epsilon} &:= \{x + i\epsilon \ : \ -1 \le x \le 1\}\\
\Gamma_{n,\epsilon} &:= \Gamma_{n,1,\epsilon} \cup (-\Gamma_{n,1,\epsilon}) \cup \overline{\Gamma_{n,1,\epsilon}} \cup (-\overline{\Gamma_{n,1,\epsilon}}) \cup \Gamma_{n,2,\epsilon} \cup \overline{\Gamma_{n,2,\epsilon}},
\end{aligned} \tag{17}$$

where, for a set $S \subseteq \mathbb{C}$, $\overline{S} := \{\overline{z} \ : \ z \in S\}$ and $-S := \{-z \ \ z \in S\}$.

We recall the following facts from [6, Chapter 10.1] (cf. the proof of Lemma 10.1.3 there).

Lemma 3.4 *There exists an $\epsilon_0 > 0$ such that the following statements hold for each ϵ, $0 < \epsilon \le \epsilon_0$, where the constants may depend upon ϵ.*

$$\oint_{\Gamma_{n,2,\epsilon}\cup\overline{\Gamma_{n,2,\epsilon}}} \mathcal{E}_n |T_n(z)\exp(-n|z|^2)|^{-1}|dz| \le c_1\exp(-cn), \tag{18}$$

$$\oint_{\Gamma_{n,1,\epsilon}\cup(-\Gamma_{n,1,\epsilon})\cup\overline{\Gamma_{n,1,\epsilon}}\cup(-\overline{\Gamma_{n,1,\epsilon}})} \mathcal{E}_n |T_n(z)\exp(-n|z|^2)|^{-1}|dz| \le cn^{-1/2}, \tag{19}$$

and

$$\left|\frac{Z_n^2 - \xi^2}{x - \xi}\right| \le c, \qquad x \in \mathbb{R},\ \xi \in \Gamma_{n,\epsilon}. \tag{20}$$

Proof. The estimates (18) and (19) are the estimates (10.1.24) and (10.1.25) respectively in [6, P. 260], the estimate (20) is not difficult, and is shown in [6] near the end of page 261. ∎

Proof of Theorem 3.2. We follow the proof of [6, Proposition 10.1.2]. Let $a < b < 1$, and n be chosen large enough so that $Z_n > (1+b)/2$. In this proof only, let $f_{1,n}(x) := f(x)/(Z_n^2 - x^2)$. Then for each $n \ge c$, each $f_{1,n}$ is continuous on $\mathbb{R}$ and $f_{1,n}(x) = 0$ if $|x| \ge a$. Moreover, $\|f_{1,n}\|_{\infty,\mathbb{R}} \le c\|f\|_{\infty,\mathbb{R}}$. We estimate $\|f_{1,n} - \mathcal{G}_m(f_{1,n})\|_{\infty,\mathbb{R}}$. We observe that for each $N \ge 1$, $R(x) := \sum_{k=0}^{N/2} Z_n^{-2k-2}x^{2k} \in \Pi_N$, and

$$\left|\frac{1}{Z_n^2 - x^2} - R(x)\right| \le c_1\exp(-c_2N), \qquad x \in [-b,b].$$

For $x \in [-3,-b] \cup [b,3]$, $|R(x)| \le c(3/b)^N$. In this proof only, let the constant denoted by c in (11) be denoted by γ. We find an integer $\beta > (1/\gamma)\log(4/b)$. Then (11) implies that for $x \in [-3,-b] \cup [b,3]$

$$|f_{1,n}(x) - \mathcal{G}_{\beta N}(f,x)R(x)| = |\mathcal{G}_{\beta N}(f,x)R(x)| \le c(3/4)^N \|f\|_{\infty,\mathbb{R}}. \tag{21}$$

For $|x| \le b$,

$$\left|\frac{f(x)}{Z_n^2 - x^2} - \mathcal{G}_{\beta N}(f,x)R(x)\right| \le c_1 E^*_{cN}(F) + c_2\exp(-c_3 N)\|f\|_{\infty,\mathbb{R}}.$$

Writing $F_{1,n}(t) := f_{1,n}(3\cos(2t))$, this estimate and (21) imply that for any integer $N \ge 1$,

$$\begin{aligned} E^*_{2(\beta+1)N}(F_{1,n}) &\le \|F_{1,n}(3\cos(2\circ)) - \mathcal{G}_{\beta N}(f, 3\cos(2\circ))R(3\cos(2\circ))\|^* \\ &\le c_1 E^*_{cN}(F) + c_2\exp(-c_3 N)\|f\|_{\infty,\mathbb{R}}, \end{aligned}$$

or equivalently,

$$E^*_m(F_{1,n}) \le c_1 E^*_{cm}(F) + c_2\exp(-c_3 m)\|f\|_{\infty,\mathbb{R}}, \qquad m \ge 1.$$

Using Theorem 3.3 again with $f_{1,n}$ in place of f, we conclude that

$$\|f_{1,n} - \mathcal{G}_m(f_{1,n})\|_{\infty,\mathbb{R}} \le c_1 E^*_{cm}(F) + c_2\exp(-c_3 m)\|f\|_{\infty,\mathbb{R}}, \qquad m \ge 1,\ n \ge c, \tag{22}$$

and

$$\|\mathcal{G}_m(f_{1,n})\|_{\infty,[-3,-b]\cup[b,3]} \le c_1\exp(-cm)\|f\|_{\infty,\mathbb{R}}. \tag{23}$$

In view of Bernstein's inequality [2, Theorem 2.2, Chapter 4], we may then choose an $\epsilon > 0$ such that

$$|\mathcal{G}_{\lfloor\sqrt{n}\rfloor}(f_{1,n},z)| \le c_1\exp(-c\sqrt{n})\|f\|_{\infty,\mathbb{R}}, \quad z \in \Gamma_{n,1,\epsilon} \cup (-\Gamma_{n,1,\epsilon}) \cup \overline{\Gamma_{n,1,\epsilon}} \cup (-\overline{\Gamma_{n,1,\epsilon}}), \tag{24}$$

and in addition, the conclusions of Lemma 3.4 hold. In this proof only, let

$$\begin{aligned} g(z) &:= \mathcal{G}_{\lfloor\sqrt{n}\rfloor}(f_{1,n},z), \\ g_n(z) &:= \exp(nz^2)(Z_n^2 - z^2)g(z), \qquad z \in \mathbb{C}, \\ h_n(x) &:= \begin{cases} g(x)(Z_n^2 - x^2), & \text{if } x \in (-Z_n, Z_n), \\ 0, & \text{if } x \in \mathbb{R}\setminus(-Z_n, Z_n), \end{cases} \end{aligned} \tag{25}$$

and L_n be polynomial interpolating g_n at the zeros of T_n. As in [6, p. 261], we obtain for $x \in \mathbb{R}$, $x \ne \pm Z_n$,

$$h_n(x)\exp(nx^2) - L_n(x) = \frac{T_n(x)}{2\pi i}\oint_{\Gamma_{n,\epsilon}} \frac{g_n(\xi)}{T_n(\xi)(\xi - x)}d\xi. \tag{26}$$

Hence,

$$
\begin{aligned}
&|h_n(x) - \exp(-nx^2)L_n(x)| \\
&\leq \quad c\|w(\sqrt{n}(\circ))T_n\|_{\infty,\mathbb{R}} \oint_{\Gamma_{n,\epsilon}} \frac{|g(\xi)|}{\exp(-n|\xi|^2)|T_n(\xi)|} \frac{|Z_n^2 - \xi^2|}{|x-\xi|}|d\xi|. \qquad (27)
\end{aligned}
$$

Taking into account the fact that $\|w(\sqrt{n}(\circ))T_n\|_{\infty,\mathbb{R}} = \mathcal{E}_n$, Lemma 3.4 implies that

$$
\begin{aligned}
|h_n(x) - \exp(-nx^2)L_n(x)| &\leq c_1\exp(-cn) \max_{\xi\in\Gamma_{n,2,\epsilon}\cup\overline{\Gamma_{n,2,\epsilon}}} |g(\xi)| \\
+ \quad c_2 n^{-1/2} &\max_{\xi\in\Gamma_{n,1,\epsilon}\cup(-\Gamma_{n,1,\epsilon})\cup\overline{\Gamma_{n,1,\epsilon}}\cup(-\overline{\Gamma_{n,1,\epsilon}})} |g(\xi)|. \qquad (28)
\end{aligned}
$$

In view of the Bernstein inequality [2, Theorem 2.2, Chapter 4],

$$
\max_{\xi\in\Gamma_{n,2,\epsilon}\cup\overline{\Gamma_{n,2,\epsilon}}} |g(\xi)| \leq c\exp(c_2\sqrt{n})\|f\|_{\infty,\mathbb{R}}.
$$

Along with (24) and (28), this implies for $x \in \mathbb{R}$, $x \neq \pm Z_n$,

$$
|h_n(x) - \exp(-nx^2)L_n(x)| \leq c_1\exp(-c\sqrt{n})\|f\|_{\infty,\mathbb{R}}. \qquad (29)
$$

Since the functions involved are continuous, (29) holds for all $x \in \mathbb{R}$.

In particular, in view of (23), we have

$$
|\exp(-nx^2)L_n(x)| \leq c_1\exp(-c\sqrt{n})\|f\|_{\infty,\mathbb{R}}, \qquad x \in [-3,-b]\cup[b,3]. \qquad (30)
$$

For $|x| \leq b$, we use (22) to conclude that

$$
\begin{aligned}
|h_n(x) - f(x)| \quad &= \quad |g(x)(Z_n^2 - x^2) - f(x)| \leq c|\mathcal{G}_{\lfloor\sqrt{n}\rfloor}(f_{1,n}, x) - f_{1,n}(x)| \\
&\leq \quad c_3 E^*_{c\sqrt{n}}(F) + c_1\exp(-c_2\sqrt{n})\|f\|_{\infty,\mathbb{R}}.
\end{aligned}
$$

Along with (29) and (30), this implies that for $|x| \leq 3$,

$$
|f(x) - \exp(-nx^2)L_n(x)| \leq c_3 E^*_{c\sqrt{n}}(F) + c_1\exp(-c_2\sqrt{n})\|f\|_{\infty,\mathbb{R}}.
$$

Necessarily, $|\exp(-nx^2)L_n(x)| \leq c_1\exp(-cn)\|f\|_{\infty,\mathbb{R}}$ if $|x| \geq 3$. This completes the proof of the theorem. ■

Next, we apply Theorem 3.2 and a simultaneous approximation theorem [6, Theorem 4.1.7] to arrive at an estimate on the degree of approximation of smooth functions and their derivatives.

Theorem 3.5 *Let $m \geq 1$, $\beta, r \geq 0$, φ be an infinitely often differentiable function on $\mathbb{R}$, supported on $[-1,1]$, and $\phi(x) := \varphi(x/m)$. There exists $P \in \Pi_{2m^2}$ such that*

$$
\|\phi - wP\|_{\infty,r,\mathbb{R}} \leq c(\beta, r, \varphi)m^{-\beta}. \qquad (31)
$$

Proof. In this proof only, let $g(x) = \exp(x^2)\phi(x)$, and $\|\varphi\|_{\infty,r,\mathbb{R}} = 1$. We prove first that for every $\gamma > 0$ there exists a polynomial $P_1 \in \Pi_{2m^2-1}$ such that

$$\|(g' - P_1)w\|_{\infty,\mathbb{R}} \le cm^{-\gamma} \tag{32}$$

We apply Theorem 3.2 with $\varphi((\sqrt{2m^2-2}/m)\circ)$ (respectively $\varphi'((\sqrt{2m^2-1}/m)\circ))$ in place of f, and make an obvious change of variables to obtain $P_2 \in \Pi_{2m^2-2}$, $R_2 \in \Pi_{2m^2-1}$ such that

$$\|\phi - P_2 w\|_{\infty,\mathbb{R}} \le cm^{-\gamma-1}, \quad \|\phi' - R_2 w\|_{\infty,\mathbb{R}} \le cm^{-\gamma}. \tag{33}$$

Since $wg' = \phi' + 2x\phi$, the polynomial $P_1 := R_2 + 2xP_2 \in \Pi_{2m^2-1}$ satisfies

$$\|(g' - P_1)w\|_{\infty,\mathbb{R}} \le \|\phi' - R_2 w\|_{\infty,\mathbb{R}} + 2\|x\phi - xP_2 w\|_{\infty,\mathbb{R}}. \tag{34}$$

Since $\|P_2 w\|_{\infty,\mathbb{R}} \le c$, $\|xP_2 w\|_{\infty,[-4m,4m]} \le cm$. The infinite-finite range inequality implies that $\|xP_2 w\|_{\infty,\mathbb{R}\setminus[-4m,4m]} \le c_1 \exp(-cm^2)$. Since ϕ is supported on $[-m,m]$, this implies that

$$\|x\phi - xP_2 w\|_{\infty,\mathbb{R}} \le cm\|\phi - P_2 w\|_{\infty,\mathbb{R}} + c_1 \exp(-cm^2) \le cm^{-\gamma}.$$

Together with (33), this leads to (32).

Again, let $P \in \Pi_{cm^2}$, $c \ge 2$, be any polynomial such that

$$\|\phi - wP\|_{\infty,\mathbb{R}} = \|(g - P)w\|_{\infty,\mathbb{R}} \le cm^{-\beta-r}. \tag{35}$$

In view of [6, Theorem 4.1.7] and the estimate (32) applied with $\gamma = \beta + r - 1$, we obtain

$$\|(g' - P')w\|_{\infty,\mathbb{R}} \le cm^{-\beta-r+1}. \tag{36}$$

Arguing as before, we see that $\|x\phi - xPw\|_{\infty,\mathbb{R}} \le cm^{-\beta-r+1}$. Therefore, (36) leads to

$$\|\phi' - (wP)'\|_{\infty,\mathbb{R}} \le cm^{-\beta-r+1}.$$

By induction, we conclude that for any $P \in \Pi_{cm^2}$ for which (35) holds,

$$\|\phi^{(k)} - (wP)^{(k)}\|_{\infty,\mathbb{R}} \le cm^{-\beta-r+k}.$$

The existence of such a polynomial being guaranteed by Theorem 3.2 (applied with $\varphi(\sqrt{c}\circ)$ in place of f), the proof is now complete. ∎

4 Gaussian networks

The basis of our proof of Theorem 2.1 is the following proposition. In the sequel, the various constants, denoted by $c, c_1, \cdots$ or by capital letters depend upon the values of those parameters among A, C, r, p, s that are present in the context. We will explicitly indicate their dependence on other quantities when required.

Proposition 4.1 *Let $\varphi : \mathbb{R} \to [0,\infty)$ be an infinitely many times differentiable function such that $\varphi(x) = 0$ if $x \notin [-1,1]$. Let $m \ge 1$ and $\phi(\mathbf{x}) := \prod_{k=1}^s \varphi(x_k/m)$, $\mathbf{x} \in \mathbb{R}^s$. Let $r \ge 1$ be an integer, $g \in \mathbb{B}(A,C;m,s)$. Then for any $\beta,\gamma > 0$,*

$$\|\phi g\|_{p,r,\mathbb{R}^s} \le c_1 m^r \|\phi g\|_{p,\mathbb{R}^s} + c_2 m^{-\beta}\|g\|_{p,r,[-cm,cm]^s} + c_3\exp(-\gamma^2 m^2)\|g\|_{p,\mathbb{R}^s}, \quad (37)$$

where c_1, c_2, c_3 are positive constants depending on φ,β,γ.

We begin by reviewing a number of results from [7], and/or proving them in a somewhat modified manner. Thus, results which are not proved here can be found in [7]. In particular, the next proposition is a reformulation of [7, Proposition 3.3].

Proposition 4.2 *Let $1 \le p \le \infty$, $m \ge 1$, $A, C > 0$. There exist positive constants c, B_1 with the following property. If $g \in \mathbb{B}(A,C;m,s)$, then*

$$c\|g\|_{p,\mathbb{R}^s} \le \sum_{k=1}^N |a_k| \le c_1 \exp(B_1^2 m^2)\|g\|_{p,\mathbb{R}^s}. \quad (38)$$

Corollary 4.3 *Let $1 \le p \le \infty$, $m \ge 1$, $A, C > 0$, $1 \le N \le C\exp(Am^2)$ be an integer, $\mathbf{x}_1, \cdots, \mathbf{x}_N$ be points in $\mathbb{R}^s$, $g = \sum_{k=1}^N a_k \exp(-\|\circ - \mathbf{x}_k\|^2) \in \mathbb{G}_{N,\infty,m,s}$, and $S \subseteq \{1,\cdots,N\}$. Then*

$$\left\|\sum_{k\in S} a_k \exp(-\|\circ - \mathbf{x}_k\|^2)\right\|_{p,r,\mathbb{R}^s} \le c\exp(B_1^2 m^2)\|g\|_{p,\mathbb{R}^s}. \quad (39)$$

In particular,

$$\|g\|_{p,r,\mathbb{R}^s} \le c\exp(B_1^2 m^2)\|g\|_{p,\mathbb{R}^s}. \quad (40)$$

The following proposition (cf. [7, Proposition 3.4]) estimates the norm on a cube of a part of a Gaussian network whose centers are away from the cube.

Proposition 4.4 *Let $1 \le p \le \infty$, $A, C > 0$, $m \ge 1$, $g := \sum_{k=1}^N a_k \exp(-\|\circ - \mathbf{x}_k\|^2) \in \mathbb{B}(A,C;m,s)$, $\mathcal{A} \subset \mathbb{R}^s$, $r \ge 0$ be an integer, and $b > 0$. Let $L_{\mathcal{A}} := \{k : \mathrm{dist}\,(\mathbf{x}_k, \mathcal{A}) \ge m\sqrt{2B_1^2 + 2b^2}\}$, and $h_{\mathcal{A}} := \sum_{k\in L_{\mathcal{A}}} a_k \exp(-\|\circ - \mathbf{x}_k\|^2)$. Then*

$$\|h_{\mathcal{A}}\|_{p,r,\mathcal{A}} \le c_3 \exp(-b^2 m^2)\|g\|_{p,\mathbb{R}^s}. \quad (41)$$

In particular, if $a > 0$, there exists a constant B_2 depending only on a, b, (but not on g) with the following property: Let $L := \{k : \|\mathbf{x}_k\| \ge B_2 m\}$, and $h := \sum_{k\in L} a_k \exp(-\|\circ - \mathbf{x}_k\|^2)$. Then

$$\|h\|_{p,r,[-am,am]^s} \le c_3 \exp(-b^2 m^2)\|g\|_{p,\mathbb{R}^s}. \quad (42)$$

Before proving Proposition 4.4, we need to introduce some further notation. For $x > 0$, let $\Pi_{x,s}$ be the class of all polynomials in s real variables with coordinatewise degree not exceeding x. The symbol $W\Pi_{x,s}$ denotes the class of all functions of the form $\mathbf{x} \to \exp(-\|\mathbf{x}\|^2/2)P(\mathbf{x})$, $P \in \Pi_{x,s}$, $\mathbf{x} \in \mathbb{R}^s$. We will make extensive use of the classical Hermite polynomials $\{h_k\}$, defined formally by the generating function (cf. [10, formula (5.5.7)])

$$\exp(2xt - t^2) =: \pi^{1/4} \sum_{k=0}^{\infty} \frac{h_k(x)}{\sqrt{k!}} (\sqrt{2}t)^k, \tag{43}$$

or by means of the Rodrigues' formula (cf. [10, formula (5.5.3)]):

$$\exp(-x^2)h_k(x) = \frac{(-1)^k}{\pi^{1/4}2^{k/2}\sqrt{k!}} \left(\frac{d}{dx}\right)^k \exp(-x^2). \tag{44}$$

The polynomial h_k is of precise degree k, and satisfies (cf. [10, formula (5.5.1)])

$$\int_{\mathbb{R}} h_k(x)h_j(x)\exp(-x^2)dx = \begin{cases} 1, & \text{if } k = j,\ k, j = 0, 1, \cdots, \\ 0, & \text{otherwise.} \end{cases} \tag{45}$$

For a multi-integer $\mathbf{k}$, we define

$$h_{\mathbf{k}}(\mathbf{x}) = \prod_{j=1}^{s} h_{k_j}(x_j). \tag{46}$$

Writing $\mathbf{k}! := \prod_{j=1}^{s}(k_j!)$, and using standard multivariate notation, we have

$$\exp(2\mathbf{x}\cdot\mathbf{w} - \|\mathbf{w}\|^2) = \pi^{s/4} \sum_{\mathbf{k}\geq 0} \frac{h_{\mathbf{k}}(\mathbf{x})}{\sqrt{\mathbf{k}!}} (\sqrt{2}\mathbf{w})^{\mathbf{k}}, \qquad \mathbf{x}, \mathbf{w} \in \mathbb{R}^s, \tag{47}$$

and for $\mathbf{k}, \mathbf{j} \geq 0$,

$$\int_{\mathbb{R}^s} h_{\mathbf{k}}(\mathbf{x})h_{\mathbf{j}}(\mathbf{x})\exp(-\|\mathbf{x}\|^2)d\mathbf{x} = \begin{cases} 1, & \text{if } \mathbf{k} = \mathbf{j}, \\ 0, & \text{otherwise.} \end{cases} \tag{48}$$

Proof of Proposition 4.4. We may assume that $\|g\|_{p,\mathbb{R}^s} = 1$. Since $g \in \mathbb{B}(A, C; m, s)$, we conclude from (38) that

$$\sum_{k=1}^{N} |a_k| \leq c_1 \exp(B_1^2 m^2). \tag{49}$$

For $\mathbf{x} \in \mathcal{A}$ and $k \in L_{\mathcal{A}}$, we have from (44) that for $|\mathbf{j}| \leq r$,

$$\begin{aligned} |\mathbf{D}^{\mathbf{j}} h_{\mathcal{A}}(\mathbf{x})| &\leq \sum_{k\in L_{\mathcal{A}}} |a_k| \{\pi^{s/4} 2^{|\mathbf{j}|/2} \sqrt{\mathbf{j}!}\} |h_{\mathbf{j}}(\mathbf{x} - \mathbf{x}_k)| \exp(-\|\mathbf{x} - \mathbf{x}_k\|^2) \\ &\leq c \exp\left(-(B_1^2 + b^2)m^2\right) \sum_{k\in L_{\mathcal{A}}} |a_k||h_{\mathbf{j}}(\mathbf{x} - \mathbf{w}_k)| \exp(-\|\mathbf{x} - \mathbf{w}_k\|^2/2) \\ &\leq c \exp\left(-(B_1^2 + b^2)m^2\right) \sum_{k=1}^{N} |a_k||h_{\mathbf{j}}(\mathbf{x} - \mathbf{w}_k)| \exp(-\|\mathbf{x} - \mathbf{w}_k\|^2/2). \end{aligned}$$

Since $\|h_{\mathbf{j}}\exp(-\|\circ\|^2/2)\|_{\infty,\mathbb{R}^s}\le c$, we obtain from (49) that

$$|\mathbf{D}^{\mathbf{j}}h_{\mathcal{A}}(\mathbf{x})|\le c\exp(-b^2m^2),$$

which is (41) with $\|g\|_{p,\mathbb{R}^s}=1$. The remaining part of the proposition follows by setting $B_2=\sqrt{s}a+\sqrt{2B_1^2+2b^2}$ and $\mathcal{A}:=[-am,am]^s$. ■

The following analogue of [7, Proposition 3.5] estimates the rate of approximation of Gaussian networks with weighted polynomials.

Proposition 4.5 *Let $C_1,C_2>0$, $1\le p\le\infty$. There exists a constant C_3 depending only on C_1, C_2, with the following property. For any $g\in\mathbb{G}_{\infty,C_1m,m,s}$, there exists a polynomial $P_g\in\Pi_{C_3m^2,s}$ such that*

$$\|g-P_g\exp(-\|\circ\|^2)\|_{p,r,\mathbb{R}^s}\le c_1\exp(-C_2^2m^2)\|g\|_{p,\mathbb{R}^s}. \tag{50}$$

As in [7], the proof of this proposition is immediate from Proposition 4.2, and the following lemma regarding the approximation of basic Gaussians by weighted polynomials.

Lemma 4.6 *For integer $n\ge1$, $\mathbf{w}\in\mathbb{R}^s$, let*

$$P_n(\mathbf{x},\mathbf{w}):=\pi^{s/4}\sum_{0\le|\mathbf{k}|\le n}\frac{h_{\mathbf{k}}(\mathbf{x})\exp(-\|\mathbf{x}\|^2)}{\sqrt{\mathbf{k}!}}(\sqrt{2}\mathbf{w})^{\mathbf{k}}. \tag{51}$$

Then for any p, $1\le p\le\infty$,

$$\|\exp(-\|\circ-\mathbf{w}\|^2)-P_n(\circ,\mathbf{w})\|_{p,r,\mathbb{R}^s}\le c_1n^c\frac{(\sqrt{2s}\|\mathbf{w}\|)^{n+1}\exp(s\|\mathbf{w}\|^2)}{\sqrt{n!}}. \tag{52}$$

Proof. Let $|\mathbf{j}|\le r$. In view of the Rodrigues' formula (44),

$$\mathbf{D}^{\mathbf{j}}\left[h_{\mathbf{k}}(\mathbf{x})\exp(-\|\mathbf{x}\|^2)\right]=(-\sqrt{2})^{|\mathbf{j}|}\frac{\sqrt{(\mathbf{k}+\mathbf{j})!}}{\sqrt{\mathbf{k}!}}h_{\mathbf{k}+\mathbf{j}}(\mathbf{x})\exp(-\|\mathbf{x}\|^2).$$

Therefore, using (47), we obtain

$$\begin{aligned}&\mathbf{D}^{\mathbf{j}}\left(\exp(-\|\mathbf{x}-\mathbf{w}\|^2)-P_n(\mathbf{x},\mathbf{w})\right)\\ &=\ \pi^{s/4}\sum_{|\mathbf{k}|\ge n}(-\sqrt{2})^{|\mathbf{j}|}\frac{\sqrt{(\mathbf{k}+\mathbf{j})!}}{\mathbf{k}!}h_{\mathbf{k}+\mathbf{j}}(\mathbf{x})\exp(-\|\mathbf{x}\|^2)(\sqrt{2}\mathbf{w})^{\mathbf{k}}.\end{aligned} \tag{53}$$

Using the arithmetic-geometric inequality, we see that $|\mathbf{w}^{\mathbf{k}}|\le\|\mathbf{w}\|^{|\mathbf{k}|}$. Since (cf. [6, Theorem 6.2.10])

$$\|h_{\mathbf{k}+\mathbf{j}}\exp(-\|\circ\|^2/2)\|_{p,\mathbb{R}^s}\le c_1|\mathbf{k}|^c,$$

we see that $\|h_{\mathbf{k}+\mathbf{j}}\exp(-\|\circ\|^2)\|_{p,\mathbb{R}^s} \le c_1|\mathbf{k}|^c$ as well. Therefore, (53) implies that

$$\|\mathbf{D}^{\mathbf{j}}\left(\exp(-\|\circ-\mathbf{w}\|^2) - P_n(\circ,\mathbf{w})\right)\|_{p,\mathbb{R}^s} \le c_1 \sum_{|\mathbf{k}|\ge n+1} \frac{|\mathbf{k}|^c(\sqrt{2}\|\mathbf{w}\|)^{|\mathbf{k}|}}{\sqrt{\mathbf{k}!}}$$

$$= c_1 \sum_{j=n+1}^{\infty} \frac{j^c(\sqrt{2}\|\mathbf{w}\|)^j}{\sqrt{j!}} \sum_{|\mathbf{k}|=j}\left(\frac{j!}{\mathbf{k}!}\right)^{1/2} \le c_1 \sum_{j=n+1}^{\infty} \frac{j^c(\sqrt{2s}\|\mathbf{w}\|)^j}{\sqrt{j!}}.$$

The estimate (52) now follows from [7, Lemma 3.5]. ■

Proof of Proposition 4.5. This proof is verbatim the same as that of [7, Proposition 3.5], except that we use Lemma 4.6 in place of [7, Lemma 3.4]. We omit the details. ■

Lemma 4.7 *Let $n \ge 1$ be an integer, $1 \le p \le \infty$, and $C_4 > 0$. There exists a positive constant C_5 depending only on C_4 such that for every $P \in \Pi_{n,s}$,*

$$\|P\exp(-\|\circ\|^2)\|_{p,r,\mathbb{R}^s\setminus[-C_5\sqrt{n},C_5\sqrt{n}]^s} \le \exp(-C_4 n)\|P\exp(-\|\circ\|^2)\|_{p,\mathbb{R}^s}. \quad (54)$$

Proof of Proposition 4.1. In this proof, we will freely use the following fact, without referring to it explicitly. If F and G are sufficiently smooth functions, then the Leibniz formula implies that for any measurable subset $S \subset \mathbb{R}^s$,

$$\|FG\|_{p,r,S} \le c\|F\|_{\infty,r,S}\|G\|_{p,r,S}. \quad (55)$$

In this proof, all the constants will depend upon $A, C, \beta, \gamma, r, p, s$, and φ. Without loss of generality, we may assume that $\gamma > B_1$, where B_1 is defined in Proposition 4.2. By a repeated application of Theorem 3.5, we obtain a polynomial $P \in \Pi_{2m^2,s}$ such that

$$\|\phi - P\exp(-\|\circ\|^2)\|_{\infty,r,\mathbb{R}^s} \le cm^{-\beta-r}. \quad (56)$$

We will write $\mathcal{P} := P\exp(-\|\circ\|^2)$. Using Lemma 4.7 with $C_4 = 3\gamma^2$, we find $a > 1$ such that

$$\|\mathcal{P}\|_{\infty,r,\mathbb{R}^s\setminus[-am,am]^s} \le c_1\exp(-3\gamma^2 m^2). \quad (57)$$

Next, we use Proposition 4.4 with $\sqrt{3}\gamma$ in place of b to obtain a number B_2, the set L and the subnetwork h of g such that

$$\|h\|_{p,r,[-am,am]^s} \le c_1\exp(-3\gamma^2 m^2)\|g\|_{p,\mathbb{R}^s}. \quad (58)$$

Clearly, the network $g-h$ contains at most cm^{2s} neurons. In view of (39), we observe that

$$\|g-h\|_{p,r,\mathbb{R}^s} \le c\exp(\gamma^2 m^2)\|g\|_{p,\mathbb{R}^s}. \quad (59)$$

Since $a>1$ and $\phi(\mathbf{x})=0$ outside of $[-m,m]^s$, we see from (58) that

$$\|\phi h\|_{p,r,\mathbb{R}^s} \le c\exp(-3\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}. \tag{60}$$

From (58) and the fact that $\|\mathcal{P}\|_{\infty,r,\mathbb{R}^s}\le c$, we obtain that

$$\|\mathcal{P}h\|_{p,r,[-am,am]^s} \le c\exp(-3\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}. \tag{61}$$

In view of (57) and (40), we see that

$$\|\mathcal{P}g\|_{p,r,\mathbb{R}^s\setminus[-am,am]^s} \le c\exp(-3\gamma^2m^2)\|g\|_{p,r,\mathbb{R}^s} \le c_1\exp(-2\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}. \tag{62}$$

Similarly, (57) and (59) imply that

$$\|(g-h)\mathcal{P}\|_{p,r,\mathbb{R}^s\setminus[-am,am]^s} \le c\exp(-2\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}.$$

Consequently,

$$\|\mathcal{P}h\|_{p,r,\mathbb{R}^s\setminus[-am,am]^s} \le c\exp(-2\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}.$$

Together with (61), this implies that

$$\|\mathcal{P}h\|_{p,r,\mathbb{R}^s} \le c\exp(-2\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}.$$

Therefore, (60) implies that

$$\|(\phi-\mathcal{P})h\|_{p,r,\mathbb{R}^s} \le c\exp(-2\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}.$$

Since (cf. (56), (62))

$$\|(\phi-\mathcal{P})g\|_{p,r,\mathbb{R}^s} \le c_1m^{-\beta-r}\|g\|_{p,r,[-am,am]^s} + c_2\exp(-2\gamma^2m^2)\|g\|_{p,\mathbb{R}^s},$$

we conclude that

$$\|(\phi-\mathcal{P})(g-h)\|_{p,r,\mathbb{R}^s} \le c_1m^{-\beta-r}\|g\|_{p,r,[-am,am]^s} + c_2\exp(-2\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}. \tag{63}$$

Since $g-h\in\mathbb{G}_{cm^{2s},c_1m,m,s}$, we may use Proposition 4.5 to obtain a polynomial $Q\in\Pi_{cm^2,s}$, such that with $\mathcal{Q}=Q\exp(-\|\circ\|^2)$,

$$\|g-h-\mathcal{Q}\|_{p,r,\mathbb{R}^s} \le c_1\exp(-3\gamma^2m^2)\|g-h\|_{p,\mathbb{R}^s} \le c_2\exp(-2\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}. \tag{64}$$

Since

$$\|\phi g-\mathcal{P}\mathcal{Q}\|_{p,r,\mathbb{R}^s} \le \|\phi h\|_{p,r,\mathbb{R}^s} + \|(\phi-\mathcal{P})(g-h)\|_{p,r,\mathbb{R}^s} + \|(g-h-\mathcal{Q})\mathcal{P}\|_{p,r,\mathbb{R}^s},$$

the estimates (60), (63), (64) lead to

$$\begin{aligned}\|\phi g-\mathcal{P}\mathcal{Q}\|_{p,\mathbb{R}^s} &\le \|\phi g-\mathcal{P}\mathcal{Q}\|_{p,r,\mathbb{R}^s}\\ &\le c_1m^{-\beta-r}\|g\|_{p,r,[-am,am]^s} + c_2\exp(-2\gamma^2m^2)\|g\|_{p,\mathbb{R}^s}.\end{aligned} \tag{65}$$

A repeated application of (8), keeping in mind the fact that the derivative of a weighted polynomial in $W\Pi_{n,s}$ is in $W\Pi_{n+1,s}$, implies that
$\|\mathcal{P}\mathcal{Q}\|_{p,r,\mathbb{R}^s} \le cm^r\|\mathcal{P}\mathcal{Q}\|_{p,\mathbb{R}^s}$. The estimate (37) now follows easily from (65). ■

PROOF OF THEOREM 2.1. We prove the theorem for $1 \le p < \infty$. The same proof works also in the case $p = \infty$, but is simpler. It is enough to prove the theorem for $m \ge c$, where c is some constant depending only on A, C, r, p, s. Let $\varphi : \mathbb{R} \to [0, \infty)$ be an infinitely differentiable function on $\mathbb{R}$, such that $\varphi(x) = 1$ if $|x| \le 1/2$ and $\varphi(x) = 0$ if $|x| \ge 1$. Let

$$\psi(x) := \frac{\varphi(x)}{\sum_{k\in\mathbb{Z}} \varphi(x-k)}$$

and $\phi(\mathbf{x}) := \prod_{j=1}^s \psi(x_j/m)$. Then ϕ is supported on $[-m, m]^s$, $0 < c_1 \le \phi(\mathbf{x}) \le 1$ if $\mathbf{x} \in [-m/2, m/2]^s$, and $\sum_{\mathbf{k}\in\mathbb{Z}^s} \phi(\mathbf{x} - \mathbf{k}m) = 1$. Let $1 \le p < \infty$. Since $D^{\mathbf{j}}g = \sum_{k\in\mathbb{Z}^s} D^{\mathbf{j}}(\phi(\circ - \mathbf{k}m)g)$ for each $\mathbf{j}$, and only finitely many of the supports of $D^{\mathbf{j}}(\phi(\circ - \mathbf{k}m)g)$ may intersect each other (the number being dependent only on s), we conclude that

$$\|g\|^p_{p,r,\mathbb{R}^s} \sim c\sum_{\mathbf{k}\in\mathbb{Z}^s} \|\phi(\circ - \mathbf{k}m)g\|^p_{p,r,\mathbb{R}^s} = c\sum_{\mathbf{k}\in\mathbb{Z}^s} \|\phi g_{\mathbf{k}}\|^p_{p,r,[-m,m]^s}, \tag{66}$$

where each

$$g_{\mathbf{k}} := g(\circ + \mathbf{k}m) \in \mathbb{B}(A, C; m, s). \tag{67}$$

Now, let $\gamma^2 > 2A$, and in this proof only, B_1 be the constant defined in Proposition 4.2, $b = \gamma\sqrt{3/p}$, and $B_3 := \sqrt{s} + \sqrt{2B_1^2 + 2b^2}$ (cf. Proposition 4.4) . In this proof only, let $\mathcal{C}$ be the set of all the centers of g,

$$\mathbb{Z}^s \setminus S := \{\mathbf{k} \in \mathbb{Z}^s \ : \ \|\mathbf{w} - \mathbf{k}m\| \ge B_3 m, \ \mathbf{w} \in \mathcal{C}\}, \quad \mathcal{A} := \bigcup_{\mathbf{k}\in\mathbb{Z}^s\setminus S} (\mathbf{k}m + [-m, m]^s) .$$

With our choice of $\mathcal{A}$ and B_3, it is clear that the subnetwork $h_{\mathcal{A}}$ defined in Proposition 4.4 is the whole network g. We observe also that only finitely many cubes among all the cubes comprising $\mathcal{A}$ intersect each other, the number depending only on s. Therefore, (41) implies that

$$\sum_{\mathbf{k}\in\mathbb{Z}^s\setminus S} \|\phi g_{\mathbf{k}}\|^p_{p,r,\mathbb{R}^s} \sim \|g\|^p_{p,r,\mathcal{A}} \le c\exp(-3\gamma^2 m^2)\|g\|^p_{p,\mathbb{R}^s}. \tag{68}$$

In view of (66) and (68), we obtain

$$\|g\|^p_{p,r,\mathbb{R}^s} \le c\sum_{\mathbf{k}\in S} \|\phi g_{\mathbf{k}}\|^p_{p,r,[-m,m]^s} + c_1\exp(-3\gamma^2 m^2)\|g\|^p_{p,\mathbb{R}^s}. \tag{69}$$

In view of the fact that the total number of centers in g is at most $C\exp(Am^2)$, the cardinality of S satisfies

$$|S| \le cm^{2s}\exp(Am^2) \le c_1\exp(2Am^2) \le c_2\exp(\gamma^2 m^2). \tag{70}$$

We now use Proposition 4.1 with each $g_{\mathbf{k}}$, $\mathbf{k}\in S$, and with some $\beta>0$ to obtain (using (69) and (70)) that

$$\|g\|_{p,r,\mathbb{R}^s}^p \le c\left\{m^{rp}\sum_{\mathbf{k}\in S}\|\phi g_{\mathbf{k}}\|_{p,\mathbb{R}^s}^p + m^{-\beta p}\sum_{\mathbf{k}\in S}\|g_{\mathbf{k}}\|_{p,r,[-cm,cm]^s}^p + \|g\|_{p,\mathbb{R}^s}^p\right\}. \tag{71}$$

Clearly, $\sum_{\mathbf{k}\in S}\|\phi g_{\mathbf{k}}\|_{p,\mathbb{R}^s}^p \le c\|g\|_{p,\mathbb{R}^s}^p$. Since only finitely many cubes of the form $\mathbf{k}m+[-cm,cm]^s$ intersect each other, we also conclude that $\sum_{\mathbf{k}\in S}\|g_{\mathbf{k}}\|_{p,r,[-cm,cm]^s}^p \le c_1\|g\|_{p,r,\mathbb{R}^s}^p$. So, (71) implies that

$$\|g\|_{p,r,\mathbb{R}^s}^p \le c\left\{m^{rp}\|g\|_{p,\mathbb{R}^s}^p + m^{-\beta p}\|g\|_{p,r,\mathbb{R}^s}^p\right\},$$

and hence, for sufficiently large m,

$$\|g\|_{p,r,\mathbb{R}^s}^p \le \frac{cm^{rp}}{1-cm^{-\beta p}}\|g\|_{p,\mathbb{R}^s}^p \le c_1 m^{rp}\|g\|_{p,\mathbb{R}^s}^p.$$

■

Acknowledgement: The research of this author was supported, in part, by grant DMS-0204704 from the National Science Foundation and grant W911NF-04-1-0339 from the U.S. Army Research Office.

References

[1] D. Benko, Approximation by weighted polynomials, J. Approx. Theory, **120** (2003), 153–182.

[2] R. A. DeVore and G. G. Lorentz, *Constructive approximation*, Springer Verlag, Berlin, 1993.

[3] T. Erdélyi, Sharp Bernstein-type inequalities for linear combinations of shifted Gaussians with applications, www.math.tamu.edu/ tamas.erdelyi/papers-online/list.html.

[4] D. Gaier, Polynomial approximation of piecewise analytic functions, J. Anal., **4** (1996), 67–79.

[5] R. Labahn, On an interesting property of binomial coefficients, Rostock Math. Kolloq., **29** (1986), 21–24.

[6] H. N. Mhaskar, *Introduction to the theory of weighted polynomial approximation*, World Scientific, Singapore, 1996.

[7] H. N. Mhaskar, When is approximation by Gaussian networks necessarily a linear process?, Neural Networks, **17** (2004), 989–1001.

[8] H. N. Mhaskar and J. Prestin, On a sequence of fast decreasing polynomial operators, Applications and computation of orthogonal polynomials (Oberwolfach, 1998), 165–178, Internat. Ser. Numer. Math., 131, Birkhäuser, Basel, 1999.

[9] F. Narcowich and J. D. Ward, Norm estimates for the inverses of a genral class of scattered data radial function interpolation matrices, Journal of Approximation Theory, **69** (1992), 84–109.

[10] G. Szegö, *Orthogonal polynomials*, Amer. Math. Soc. Colloq. Publ. **23**, Amer. Math. Soc., Providence, 1975.

Department of Mathematics
California State University
Los Angeles, California, 90032, U.S.A.
Email address: hmhaska@calstatela.edu