

# Emissionsdichteschätzung bei Time-of-Flight Positronen Emissions–Tomographie

Von der Fakultät Mathematik der Universität Stuttgart  
zur Erlangung der Würde eines Doktors der  
Naturwissenschaften (Dr. rer. nat. ) genehmigte Abhandlung

Vorgelegt von

**Silke Kuball**

aus Neunkirchen/Saar

Hauptberichter: Prof. Dr. Christian H. Hesse  
Mitberichter: Prof. Dr. Harro Walk  
Zweiter Mitberichter: Dr. Ing. habil. Fritz Schmidt

Tag der Einreichung: 5. August 2001  
Tag der mündlichen Prüfung: 14. Februar 2002

Mathematisches Institut A der Universität Stuttgart

2002



# Inhaltsverzeichnis

Bezeichnungen	5
Zusammenfassung	6
Abstract	7
<b>1 Einleitung</b>	<b>16</b>
<b>2 Schätzung der Positronen Emissionsdichte</b>	<b>23</b>
2.1 Modell . . . . .	23
2.2 Dekonvolutionsansatz . . . . .	25
2.3 Konstruktion der Schätzfunktion . . . . .	28
<b>3 Eigenschaften des modifizierten Dekonvolutions–Kernschätzers</b>	<b>35</b>
3.1 Bias und Varianz der Schätzfunktion . . . . .	35
3.2 Fehlerverhalten der Schätzfunktion . . . . .	43
3.2.1 Wahl der Funktionenklasse . . . . .	44
3.2.2 Die Fehlermaße Mittlerer Quadratischer Fehler und Mitt- lerer Integrierter Quadratischer Fehler . . . . .	48
3.2.3 Das Verhalten des Schätzers in der Supremum–Norm . . .	62
3.3 Anmerkungen zu Kapitel 3 . . . . .	73
<b>4 Optimalitätseigenschaften</b>	<b>74</b>
4.1 Ein Verfahren zur Bestimmung der Minimax Konvergenzrate . . . . .	74
4.2 Anwendung auf ToFPET Dichteschätzung . . . . .	83
4.3 Untere Schranke für den MSE . . . . .	85

4.4	Untere Schranke für den MISE . . . . .	95
4.5	Untere Schranke für die Supremum-Norm . . . . .	100
4.6	Einordnung der Ergebnisse in die Literatur . . . . .	103
<b>5</b>	<b>Test des Schätzverfahrens</b>	<b>105</b>
5.1	Simulation von ToFPET-Beobachtungen . . . . .	105
5.1.1	Die Pearson-Typ-II Dichtefunktion . . . . .	106
5.1.2	Erzeugen von ToFPET-Beobachtungspaaren . . . . .	106
5.2	Bestimmung der Funktionenklasse . . . . .	110
5.3	Algorithmische Berechnung der Rekonstruktionsdichte . . . . .	111
5.3.1	Einführende Bemerkungen . . . . .	111
5.3.2	Wahl der Bandweite . . . . .	113
5.4	Darstellung der Rekonstruktionsergebnisse . . . . .	114
5.5	Ausblick . . . . .	124
	<b>Anhang A: Parallele Implementierung des Schätzverfahrens</b>	<b>126</b>
A.1	Einleitung . . . . .	126
A.2	C-Programmcode für einen Parix-Rechner . . . . .	128
	<b>Danksagung</b>	<b>141</b>
	<b>Literaturverzeichnis</b>	<b>142</b>
	<b>Lebenslauf</b>	<b>147</b>

## Bezeichnungen

$\mathbb{N}$ ,	Menge der natürlichen Zahlen ohne 0,
$\mathbb{N}_0$ ,	Menge der natürlichen Zahlen einschließlich 0,
$\mathbb{R}$ ,	Menge der reellen Zahlen,
$\mathbb{R}_0^+$ ,	Menge der positiven reellen Zahlen, einschließlich 0,
$\mathcal{L}^1(\mathbb{R}^2)$	$:= \{f : \mathbb{R}^2 \rightarrow \mathbb{R}, \int_{\mathbb{R}^2}  f(y_1, y_2)  dy_1 dy_2 < \infty\}$ ,
$\mathcal{L}^2(\mathbb{R}^2)$	$:= \{f : \mathbb{R}^2 \rightarrow \mathbb{R}, (\int_{\mathbb{R}^2}  f(y_1, y_2) ^2 dy_1 dy_2)^{\frac{1}{2}} < \infty\}$ ,
$\mathcal{H}^{\alpha,2}$	Sobolev-Raum reellwertiger Ordnung $\alpha$ ,
$\mathbf{X}, \mathbf{x}$	Vektoren in $\mathbb{R}^2$ werden in Fettschrift bezeichnet, Zufallsvektoren werden mit großen Buchstaben, deren Realisierung mit kleinen Buchstaben bezeichnet,
$\ \mathbf{x}\ _p$	$:= (\sum_{i=1}^m  x_i ^p)^{\frac{1}{p}}$ , $\mathbf{x} \in \mathbb{R}^m$ , die $p$ -Norm im $\mathbb{R}^m$ , ist $p$ nicht spezifiziert, so ist immer die 2-Norm gemeint,
$\mathbf{x}' \mathbf{y}$	$:= x_1 y_1 + x_2 y_2$ , euklidisches Skalarprodukt,
$\mathcal{C}^0(\mathbb{R}^2)$	Raum der stetigen Funktionen über $\mathbb{R}^2$ ,
$\mathcal{C}_\infty^0(\mathbb{R}^2)$	Raum der unendlich oft stetig differenzierbaren Funktionen mit kompaktem Träger in $\mathbb{R}^2$ ,
$\mathcal{B}$	$:= \{\mathbf{y} \text{ mit } \ \mathbf{y}\  \leq 1\}$ ,
$\mathcal{D}$	$:= \{\mathbf{y} \text{ mit } \ \mathbf{y}\  = 1\}$ ,
$f$	Emissionsdichte,
$\hat{f}_n(\mathbf{y})$	modifizierter Dekonvolutions-Kernschätzer von $f$ ,
$\lambda_n$	Bandweiteparameter des Schätzers $\hat{f}_n(\mathbf{y})$ ,
$\mathcal{N}(\mu, \sigma^2)$	die Normalverteilung mit Erwartungswert $\mu$ und Varianz $\sigma^2$ ,
$\mathcal{U}([a, b])$	die stetige gleichmäßige Verteilung auf dem Intervall $[a, b]$ ,
$E[\mathbf{X}]$	Erwartungswert von $\mathbf{X}$ ,
$E[\mathbf{X} \mathbf{y}]$	bedingter Erwartungswert von $\mathbf{X}$ gegeben $\mathbf{Y} = \mathbf{y}$ ,
$var(\mathbf{X})$	Varianz von $\mathbf{X}$ ,
$MSE$	Mittlerer Quadratischer Fehler,
$MISE$	Mittlerer Integrierter Quadratischer Fehler,
$\Phi_w(\mathbf{t})$	Fouriertransformierte einer Funktion $w$ ,
$\angle(\mathbf{x}, \mathbf{y})$	Winkel zwischen zwei Vektoren $\mathbf{x}$ und $\mathbf{y}$ ,
$f * h$	Faltung zweier Funktionen $f$ und $h$ ,
$[f(x)]_a^b$	$:= f(b) - f(a)$ ,
f. ü.	fast überall,
f. s.	fast sicher, mit Wahrscheinlichkeit 1,
u. i. v.	unabhängig identisch verteilt,
$C$	eine generische Konstante.

# Zusammenfassung

In dieser Arbeit wird ein nichtparametrisches Schätzverfahren entwickelt, um die Dichtefunktion des Positronen-Emissionsprozesses bei Time-of-Flight Positronen-Emissions-Tomographie (ToFPET) über einem zweidimensionalen Querschnitt eines untersuchten Objektes zu schätzen. Es wird das Beispiel der Gehirntomographie zur Modellierung gewählt, die Ergebnisse sind jedoch auf alle anderen Fälle von ToFPET zu übertragen. ToFPET unterscheidet sich von herkömmlicher PET durch das Vorliegen zusätzlicher Meßinformation (Time of Flight), die es ermöglicht die Beobachtungsdaten als Punkte im Querschnitt des Gehirnes anzusehen, statt, wie bei PET, als Messungen auf einem Detektor außerhalb des Kopfes.

Es werden idealisierte Annahmen über den experimentellen Ablauf vorgenommen. Als Klasse der Emissionsdichten wird  $\mathcal{H}^{\alpha,2}$ , ein Sobolev-Raum der Ordnung  $\alpha \in \mathbb{R}_0^+$  mit  $\alpha > 1$  gewählt. Damit werden nur geringe Anforderungen an die Glattheit der Emissionsdichte gestellt, was dem Auftreten hoher Kontraste bei Tomographie-Bildern entspricht. Ein Dekonvolutionsverfahren mit gewichteten Beobachtungen wird eingesetzt, um eine Schätzfunktion für die gesuchte Emissionsdichte zu entwickeln. Der entwickelte Schätzer ist asymptotisch unverfälscht. Es werden obere Schranken für die Konvergenzraten des Schätzers bezüglich der Fehlermaße Mittlerer Quadratischer Fehler (MSE), Mittlerer Integrierter Quadratischer Fehler (MISE) und Supremum-Norm bestimmt. Es ergeben sich algebraische Konvergenzordnungen der Form  $O(n^{-\frac{2(\alpha-1)}{2\alpha+1}})$  für den Mittleren Quadratischen Fehler sowie den Mittleren Integrierten Quadratischen Fehler und  $O(n^{-\frac{\alpha-1}{2\alpha+1}} \ln(n)^{\frac{\alpha-1}{2\alpha+1}})$  für die Supremum-Norm.

Die erreichten Raten werden auf Optimalität untersucht. Es ergibt sich, daß unter dem MSE und dem MISE, der entwickelte Schätzer eine Minimax-Konvergenzrate erreicht, und daß unter der Supremum-Norm die Konvergenzrate des Schätzers bis auf einen logarithmischen Faktor übereinstimmt mit einer gleichmäßigen unteren Schranke über dem Funktionenraum  $\mathcal{H}^{\alpha,2}$ , die für jedes denkbare Rekonstruktionsverfahren gültig ist. Es gibt also kein Rekonstruktionsverfahren, welches hinsichtlich Konvergenz dem hier entwickelten modifizierten Dekonvolutionsschätzer überlegen ist.

Die Implementierung der Schätzfunktion wird anhand eines simulierten Datensatzes demonstriert. Die experimentellen Ergebnisse sind mit den theoretischen Resultaten im Einklang. Eine mögliche Parallelisierung der Implementierung wird untersucht.

# Abstract

Medicine benefits increasingly from the routine-use of computer-tomography techniques such as Positron-Emission-Tomography (PET). PET is used for example in the examination of brain functionality and represents a source of information to be used in the study of diseases such as Alzheimers or Shizophrenia. It works via an injection of a radia-pharmaceutical into the organ under study. We consider the case of brain-tomography. Radioactive (positron emitting) labelled glucose is injected and taken up by the brain cells. The amount of absorbed glucose is related to the area's metabolism. After a short while, the positron emission process starts. Once emitted, a positron annihilates almost instantly with a free electron and in this process, a pair of high-energy photons is generated, which fly off in opposite directions, the flight angle of a photon pair to a horizontal axis is randomly distributed. If both photons of a pair are captured by a surrounding detector-ring, the emission is counted. The actual PET measurements consist of flight-paths of photon-pairs.

Reconstruction algorithms exist to infer an estimate of the positron emission distribution from these measurements.

The use of PET could be much supported by techniques to shorten the time it takes to evaluate the PET images. Shorter evaluation times mean higher cost-effectiveness. This can be achieved by enhancing the image quality. One possibility to enhance image quality is by the use of additional measurement data. Time-of-Flight PET (ToFPET) yields such additional data by monitoring not just the emitted photons' flight-path but in addition the flight-time difference two photons in a pair hit the detector-ring with. This additional information allows to not only detect flight-paths but to calculate approximate emission locations in the brain.

It is the scope of this work to make use of this additional information by building it into a reconstruction algorithm for the positron emission density function.

We assume an idealized experimental setup for ToFPET, see also Johnstone und Silverman (1990) [31], around which we develop our reconstruction algorithm. This setup is given in Figure 6.1. Hereby  $\mathbf{X}_j$  is the observed emission location,  $\mathbf{Y}_j$  is the actual location of a positron emission,  $\mathbf{Z}_j$  is the observation error and can be written as  $r_j \Theta_{\varphi_j}$ , with  $\varphi_j$  the flight-angle of the j-th photon pair.

The following model is used to model the occurrence of positron emissions.

$$\mathbf{X}_j = \mathbf{Y}_j + \mathbf{Z}_j = \mathbf{Y}_j + r_j \Theta_{\varphi_j}, \quad j = 1, \dots, n.$$

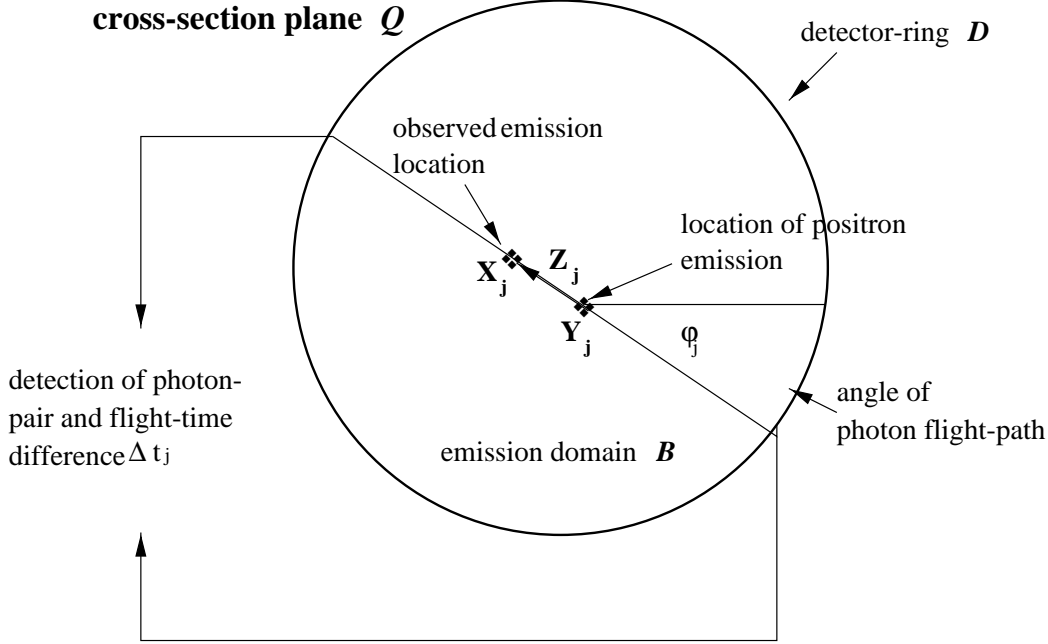


Figure 0.1: Idealized ToFPET experimental setup.

$n$  is large, up to an order of magnitude of  $10^6$ . The model assumptions are as follows

1.  $\mathbf{Y}_j, j = 1, \dots, n$  i. i. d. random variables with a density function  $f$  on  $\mathcal{B}$ .
2.  $r_j, j = 1, \dots, n$  i. i. d.  $\mathcal{N}(0, \sigma^2)$  random variables.
3.  $\varphi_j, j = 1, \dots, n$  i. i. d.  $\mathcal{U}([0, \pi])$  random variables.
4.  $r_j$  und  $\varphi_j$  are independent random variables.
5.  $\mathbf{Y}_j$  und  $\mathbf{Z}_j$  are independent random variables.

More detail on the model and the assumptions is given in the full thesis, chapter 2. We consider the problem of estimating an emission density from the space  $\mathcal{H}^{\alpha,2}(\mathcal{B})$ , a Sobolev space of real order with  $\alpha > 1$ . This choice poses only little assumptions on the smoothness of the emission density. The algorithm we use is a deconvolution kernel density estimator with an additional weight function to emphasize particularly informative measurements, see also Zhang (1992) [61] who treats this estimation problem over spaces of functions with Hölder-continuous classical derivatives, a function class which poses more restrictive assumptions on



the emission density's smoothness properties.

The introduced estimator is called modified deconvolution kernel density estimator and is defined as in Definition 6.1.

Hereby let  $K(x_1, x_2) : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a kernel function with

1.  $K(x_1, x_2) = K_1(x_1) \cdot K_2(x_2)$ .

For  $i = 1, 2$ :

2.  $K_i(x_i) = K_i(-x_i)$ ,

3.  $|K_i(x_i)| \leq C < \infty$ ,

4.  $\int_{\mathbb{R}} K_i(x_i) dx_i = 1$ ,

5.  $\int_{\mathbb{R}} x_i \cdot K_i(x_i) dx_i = 0$ .

Furthermore let  $G(\mathbf{t}, \Theta_\varphi)$  be a weight function taking on values in  $]0, 1]$ . This function is monotone decreasing with increasing  $|\mathbf{t}'\Theta|$  and possesses a finite integral. More detail in chapter 2 of this thesis.

**Definition 0.1 (Modified Dekonvolution Kernel Density Estimator)**

Given  $n$  i. i. d. observations  $(\mathbf{X}_j, \varphi_j)$ ,  $j = 1, \dots, n$  for the above described estimation problem. Given a sequence of parameters  $(\lambda_n)_{n \in \mathbb{N}}$  in  $\mathbb{R}^+$  with  $\lambda_n \rightarrow 0$  and  $n\lambda_n^2 \rightarrow \infty$  for  $n \rightarrow \infty$ . Let the function  $M(\mathbf{t})$  be defined as  $M(\mathbf{t}) := E[e^{i\mathbf{t}'\Theta_\varphi} G(\mathbf{t}, \Theta_\varphi)]$ . Then

$$\hat{f}_n(\mathbf{y}) := \frac{1}{(2\pi)^{2n}} \sum_{j=1}^n \int_{\mathbf{t} \in \mathbb{R}^2} e^{-i\mathbf{t}'(\mathbf{y}-\mathbf{X}_j)} \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt \quad \forall \mathbf{y} \in \mathcal{B} \quad (0.1)$$

is called **modified deconvolution kernel density estimator** for  $f(\mathbf{y})$ .

Hereby  $\lambda_n$  is called bandwidth parameter. The choice of  $\lambda_n$  influences the properties of the estimator. What the optimal choice for  $\lambda_n$  is depends on the smoothness properties of the density to be estimated. In this thesis we investigate the properties of this estimator under the assumption  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$ ,  $\alpha > 1$ . We show that the estimator is asymptotically unbiased. We investigate the estimator's convergence behaviour under the following distance measures.

1. Mean Squared Error (MSE) of  $\hat{f}_n$  at  $\mathbf{y}$ :

$$MSE(\hat{f}_n(\mathbf{y})) := E \left[ (\hat{f}_n(\mathbf{y}) - f(\mathbf{y}))^2 \right].$$

2. Mean Integrated Squared Error (MISE) of  $\hat{f}_n$ :

$$MISE(\hat{f}_n) := \int_{\mathcal{B}} MSE(\hat{f}_n(\mathbf{y})) d\mathbf{y}.$$

3. The supremum norm of  $\hat{f}_n - f$ :

$$\|\hat{f}_n(\mathbf{y}) - f(\mathbf{y})\|_{\infty} := \sup_{\mathcal{B}} | \hat{f}_n(\mathbf{y}) - f(\mathbf{y}) |.$$

We start by investigating upper bounds of the speed of convergence of  $\hat{f}_n$  towards  $f$  under the above distance measures for  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$ . To obtain these, kernel and weight function need to be specified. The following choices for the kernel function and the weight function are made.

$$K(\mathbf{y}) := \frac{\sin(y_1)}{\pi y_1} \frac{\sin(y_2)}{\pi y_2}.$$

This function is called Fourierintegral kernel, see for example Stefanski (1990) [50] or Bullock-Davis [9], [10]. For the weight function we choose

$$G(\mathbf{t}, \Theta_{\varphi}) := e^{-(\mathbf{t}' \Theta_{\varphi})^2 \frac{\sigma^2}{2}}.$$

Hereby  $\sigma$  is the standard deviation of  $r_j$ , the observation error size. We show that with respect to the estimator's variance, this choice of weight function is optimal. We obtain the following results for the rates of convergence of the modified deconvolution density estimator.

**Result 0.1** *The optimal bandwidth choice for the modified deconvolution kernel density estimator with Fourierintergral kernel and weight function  $G(\mathbf{t}, \Theta_{\varphi}) = e^{-(\mathbf{t}' \Theta_{\varphi})^2 \frac{\sigma^2}{2}}$  under MSE and MISE and for  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$  is given by*

$$\lambda_n^{MSE} = C n^{-\frac{1}{2\alpha+1}}.$$

Hereby,  $\alpha$  is the parameter indicating the order of the chosen Sobolev space of emission densities. Upper bounds on the MSE and MISE are then given by

$$MSE(\hat{f}_n(\mathbf{y})) \leq C n^{-\frac{2(\alpha-1)}{2\alpha+1}}.$$

$$MISE(\hat{f}_n) \leq C n^{-\frac{2(\alpha-1)}{2\alpha+1}}.$$

For the supremum norm we obtain the following result.

**Result 0.2** *Optimal bandwidth choice for the modified deconvolution kernel density estimator under the supremum norm and for  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$ :*

$$\lambda_n^{sup} = C \left( \frac{\ln(n)}{n} \right)^{\frac{1}{2\alpha+1}}.$$

*Upper bound:*

$$\sup_{\mathcal{B}} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})| = O \left( n^{-\frac{\alpha-1}{2\alpha+1}} \ln(n)^{\frac{\alpha-1}{2\alpha+1}} \right) \quad \text{with probability one.}$$

The achieved rates are then examined for their optimality. This means we ask the question: is it possible that there exists another estimator, which achieves, uniformly over  $\mathcal{H}^{\alpha,2}$  better rates of convergence?

Let  $\mathcal{T}_n$  be the class of all possible estimators for  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$  and  $\hat{T}_n \in \mathcal{T}_n$ .  $L(\hat{T}_n, f)$  stands for any of the distance measures examined, such as MSE, MISE. The convergence rate of

$$\inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} L(\hat{T}_n, f) \tag{0.2}$$

is called minimax rate of convergence for estimation of  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$ . There is no estimator that achieves a better rate of convergence. We show that the modified deconvolution kernel density estimator achieves a minimax rate of convergence both for MSE and MISE, and that the upper bound inferred for the convergence of the estimator under the supremum norm complies up to a logarithmic factor with a uniform lower bound for the supremum norm, which holds for any reconstruction technique.

We calculate minimax rates of convergence for  $\hat{f}_n(\mathbf{y})$  under MSE and MISE by adapting to our estimation problem a technique introduced by Donoho und Liu (1991) [12].

**Result 0.3**

$$\inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} MSE(\hat{T}_n(\mathbf{y}_M)) \geq C n^{-\frac{2(\alpha-1)}{2\alpha+1}}.$$

$$\inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} MISE(\hat{T}_n) \geq C n^{-\frac{2(\alpha-1)}{2\alpha+1}}.$$

A comparison with the upper bounds for MSE and MISE shows that indeed for the examined estimation problem, the modified deconvolution kernel density estimate achieves the - under the minimax criterium - optimal rate of convergence under MSE and MISE.

In chapter 4 we introduce a definition for the lower bound of an estimator's convergence rate under the supremum norm, which is similar to definitions introduced by Stone (1980, 1983) [52], [54] and Donoho and Liu (1991).

**Definition 0.2** *Let  $(b_n)$  be a sequence of positiv numbers.*

*The sequence  $(b_n)$  is called lower bound for the rate of convergence in the estimation of  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$  under the supremum norm, if for all  $C > 0$*

$$\liminf_{n \rightarrow \infty} \inf_{T_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(\sup_{\mathcal{B}} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| > C b_n) > 0.$$

The following result is derived.

$$\liminf_{n \rightarrow \infty} \inf_{T_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(\sup \lim its_{\mathcal{B}} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq C n^{-\frac{\alpha-1}{2\alpha+1}}) > 0.$$

This implies that the lower bound achieved for the supremum norm, a quality of the estimation problem is - up to a logarithmic factor - identical with the upper bound achieved by the modified deconvolution kernel density estimator.

We conclude that the modified deconvolution kernel density estimator shows - uniform over  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  and under the observed distance measures - optimal convergence behaviour. There is no other reconstruction technique, which - under the observed distance measures - achieves better convergence properties than the modified deconvolution kernel density estimator uniformly over the assumed function space  $\mathcal{H}^{\alpha,2}$ .

In chapter 5 an implementation of the modified deconvolution kernel density estimator is described and discussed and the estimator is tested on simulated data sets. Data are generated as samples from a Pearson–Type–II density and observation noise is added according to the model assumptions.

Bandwidth selection is performed through a simple semi-automatic technique, which approximates the true value for the optimal bandwidth by considering a grid of possible values for the bandwidth, gradually refining this grid and encircling the optimal bandwidth choice. Reconstruction is performed over a grid of 400 points. Figures 6.2-6.5 show two examples of results obtained for the reconstruction of a Pearson–Type–II density with parameters  $m = 2, \sigma_P^2 = 8$  based on

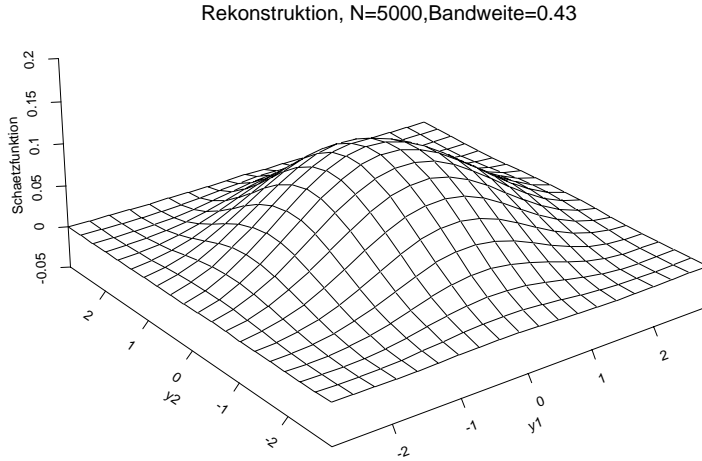


Figure 0.2: Reconstruction of the Pearson–Type–II density  $f_{2,8}(\mathbf{y})$  from 5000 data points.

5000 and 75000 simulated ToFPET data pairs respectively. This density is shown to belong to a Sobolev space  $\mathcal{H}^{\alpha,2}$  with  $\alpha < 2.5$ .

Figures 6.2 and 6.4 show the reconstruction itself on a scale from -0.05 to 0.2, figures 6.3 and 6.5 show the absolute value of the reconstruction error on a scale from 0 to 0.006.

When passing from 5000 to 75000 data points, the value  $\max_{i,j=0,\dots,19} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}})|$  reduces from ca. 0.00586 to 0.00243. Hereby  $\mathbf{y}_{M_{ij}}$  is the centre of the grid segment with index  $(ij)$ . For a series of 4 reconstruction results, we assess  $\max_{i,j=0,\dots,19} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}})|$ , the maximum of the absolute value of the reconstruction error over the grid points. This yields an approximation of the supremum-norm with increasing sample size. Plotting this value on a logarithmic scale over the sample size allows us to derive an experimental convergence rate for the approximated supremum norm. The experimental rate is in accordance with the theoretical results derived in the first chapters of this thesis.

In the Appendix, the potential for building a parallel algorithm for the calculation of the estimator is studied. An implementation of the estimation algorithm on a Parix parallel computer is given in the Appendix.

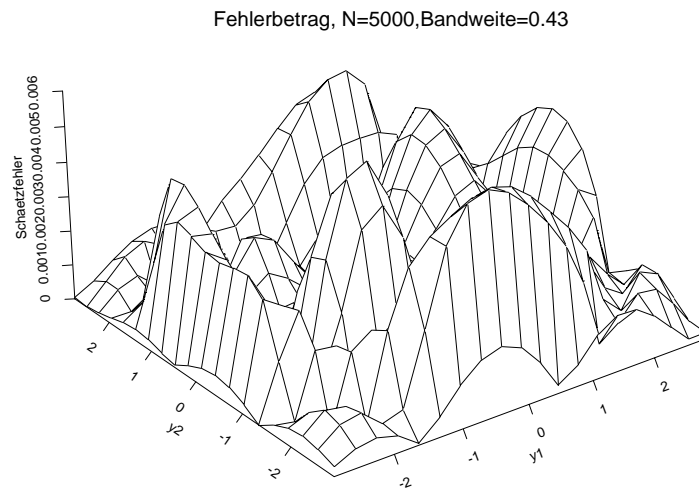


Figure 0.3: Reconstruction error  $|\hat{f}_n(\mathbf{y}) - f_{2,8}(\mathbf{y})|$  for  $n = 5000$ .

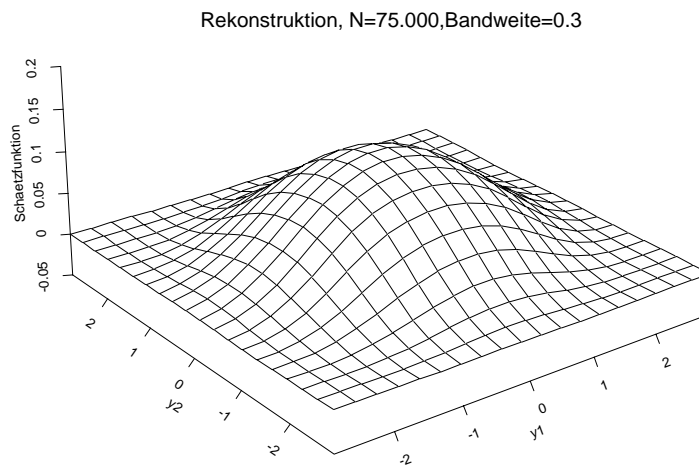


Figure 0.4: Reconstruction of the Pearson-Type-II density  $f_{2,8}(\mathbf{y})$  from 75,000 data points.

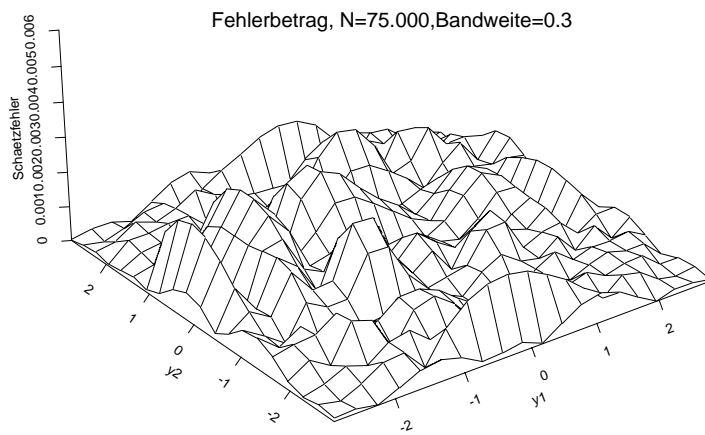


Figure 0.5: Reconstruction error  $|\hat{f}_n(\mathbf{y}) - f_{2,8}(\mathbf{y})|$  for  $n = 75,000$ .

# Kapitel 1

## Einleitung

Die medizinische Wissenschaft profitiert in zunehmendem Maße vom Routine-Einsatz moderner Computer-Tomographie Verfahren. Man unterscheidet zwischen Emissions-Tomographie Verfahren wie der **P**ositronen-**E**missions-**T**omographie (**PET**) und Transmissions-Tomographie Verfahren wie der Röntgen-Tomographie.

Emissions-Tomographie wird zur Untersuchung der Funktionsweise von Körperorganen verwendet, ohne daß ein direkter Eingriff in den Körper erfolgt. Das Verfahren ermittelt die Verteilung eines Indikatormittels im Körper aus dem Auftreten radioaktiver Emissionen. Daraus wird die Intensität der Durchblutung verschiedener interessierender Körper-Regionen abgeleitet. PET-Bilder werden z.B. zur Untersuchung der Wirksamkeit von Medikamenten bei der Behandlung von Suchterkrankungen herangezogen, als Hilfsmittel zur Diagnose bestimmter Gehirnerkrankungen verwendet sowie auch zur Erforschung der generellen Funktionsweise des Gehirnes eingesetzt.

Transmissions-Tomographie wird zur Bestimmung der Anatomie von Körperorganen eingesetzt. Dabei wird ein Röntgenstrahl durch den Körper gesendet und die Verminderung der Intensität zwischen Ein- und Austritt gemessen. Dies ermöglicht eine Darstellung der Gewebedichte.

Die Kombination von Emissions- mit Transmissions-Tomographie Bildern zur Erstellung einer Diagnose gewinnt zunehmend an Bedeutung in der Medizin. Emissions-Tomographie Verfahren sind jedoch kostspielig und daher bisher hauptsächlich in grossen Universitätskliniken oder Forschungsinstituten zu finden. Um den Einsatz dieser Verfahren zu fördern, wird nach Möglichkeiten gesucht, die Effizienz der Anwendung zu steigern und damit die Kosten zu senken. Dies beinhaltet unter anderem, die Qualität der Tomographie-Bilder zu erhöhen, um die



Auswertungszeit der Bilder zu verkürzen. Zudem kann bei besserer Bildqualität die Auswertung der Bilder vom Experten auf einen Arzt mit weniger Expertise verlagert werden, was Kosten spart. Eine Erhöhung der Genauigkeit der Bild-darstellung würde es ermöglichen, daß Tomographie-Verfahren in der Zukunft schmerzhaft und risikobehaftete Eingriffe wie die Entnahme von Biopsien ersetzen können. Zu diesem Zweck müssen verbesserte Rekonstruktionsverfahren entwickelt werden. Verbesserte Rekonstruktionsverfahren benötigen weniger Meß-daten und verringern damit die Strahlenbelastung für den Patienten.

Das Rekonstruktionsergebnis kann darüberhinaus durch die Einarbeitung von Zusatzinformationen verbessert werden. Dies kann Vorabinformation über das zu erwartende Bild oder zusätzliche Meßinformation sein, die bei der Registrierung der Daten anfällt. Ein PET-Verfahren, bei dem dies möglich ist, ist die **Time-of-Flight Positronen-Emissions-Tomographie (ToFPET)**. In dieser Arbeit wird ein Verfahren entwickelt, um aus ToFPET-Meßdaten Bildinformation hoher Qualität in Form einer Intensitätsverteilung zu gewinnen. Die Güte dieses Verfahrens wird untersucht.

Im Folgenden wird zunächst die Vorgehensweise bei der Positronen-Emissions-Tomographie und der spezielle Fall der ToFPET am Beispiel der Untersuchung eines menschlichen Gehirnes erläutert. Der klinische Verlauf der PET kann wie folgt gegliedert werden:

### **Positronen-Emissions-Tomographie:**

#### 1. Vorbereitung

- Herstellung eines positronen-emittierenden Isotops in einem Zyklo-tron.
- Versetzen eines Trägerstoffes mit dem Isotop („Labeling“) und somit Gewinnung des Indikatormittels.

#### 2. Durchführung

- Injektion des Indikatormittels in die Blutbahn des Patienten.
- Detektion von Photonenemissionen auf einem Detektorring.

#### 3. Nachbereitung

- Abbildung des Gehirnquerschnittes mit farbiger Darstellung der Durch-blutungsaktivität verschiedener Regionen (Bild-Rekonstruktion). Bild-bearbeitung mittels Computerverfahren.

– Bildinterpretation, Diagnose.

Das Ziel bei der Anwendung von Gehirn-PET ist es, Informationen über die Intensität der Durchblutung, d.h. die metabolische Aktivität, verschiedener Gehirnregionen zu erhalten. Daraus können Aussagen über die Funktionsweise des Gehirnes abgeleitet werden und eventuelle Fehlfunktionen, wie sie bei Schizophrenie, Alzheimer'scher Krankheit, Epilepsie und anderen Gehirnerkrankungen auftreten, diagnostiziert werden.

Dem Patienten wird durch eine Injektion ein Trägerstoff verabreicht, der mit einem radioaktiven Substrat versetzt ist. Der Trägerstoff ist so gewählt, daß er die physiologischen und chemischen Prozesse des Körpers nicht beeinträchtigt. Die Ausbreitung des Trägerstoffes im Körper kann indirekt durch die Messung radioaktiver Emissionen (Emission von Positronen) beobachtet werden. Bei Gehirnuntersuchungen wird oft Glukose als Trägerstoff verwendet, da sie als Hauptenergielieferant des Gehirnes fungiert und leicht von den Gehirnzellen aufgenommen wird. Mit dem Blutkreislauf gelangt die radioaktiv versetzte Glukose in das Gehirn, wo sie von den Zellen aufgenommen wird. Der Positronen-Emissions-Prozess setzt dann fast unmittelbar ein. Die Anzahl emittierter Positronen in einem Gehirnbereich ist proportional zur Menge aufgenommener Glukose und damit ein Maß für die metabolische Aktivität dieser Region. Ziel ist es, die Verteilung des Auftretens von Positronen-Emissionen im Gehirn darzustellen. Dies wird mittels einer Wahrscheinlichkeitsdichte erreicht, welche als *Emissionsdichte* bezeichnet wird. In Vardi et al. (1985) [58] sowie Johnstone und Silverman (1990) [31] wird das Auftreten von Positronen-Emissionen als raum-zeitlicher Poisson-Prozeß modelliert. Die räumliche Intensitätsfunktion des Poisson-Prozesses beschreibt dann die Glukose-Verteilung im Gehirn.

Die Emission von Positronen im Gehirn ist nicht direkt beobachtbar. Es erfolgt eine indirekte Ermittlung des Ortes einer Positronen-Emission unter Verwendung von Richtungs- und Zeitinformation über das zugehörige Photonenpaar. Der Messung liegt das folgende physikalische Phänomen zugrunde. Ein emittiertes Positron trifft fast unmittelbar auf ein freies Elektron. Bei der dann stattfindenden Annihilierung von Elektron und Positron entsteht ein Röntgenphotonen-Paar. Die Photonen des Paares werden in zueinander entgegengesetzten Richtungen ausgesandt. Der Kopf des Patienten ist von einem Detektorring umgeben. Treffen beide Photonen eines Paares auf Detektoren des Ringes, wird das Paar gezählt und eine Emission registriert. Da die Photonen immer nur paarweise registriert werden, und sich nicht alle in der von den Detektoren definierten Ebene bewegen, bleibt ein gewisser Anteil der Photonen undetektiert. Die Beobachtungen bestehen aus Flugbahnen von Photonen. Aus diesen Daten kann mittels bekann-

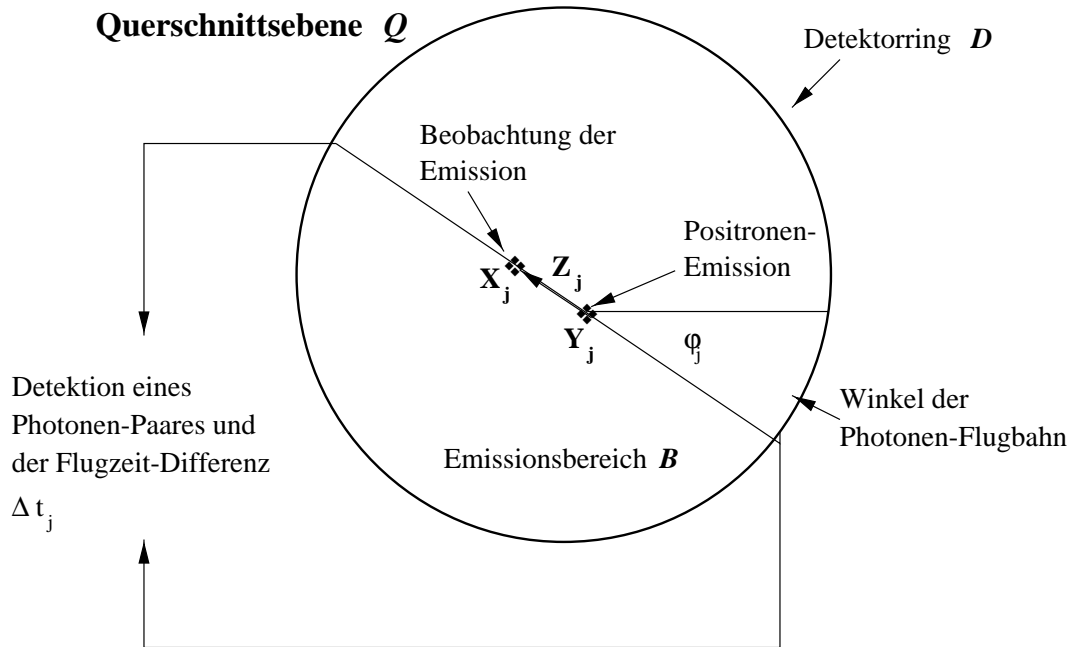


Abbildung 1.1: Schematische Darstellung des Detektionsvorganges.

ter Rekonstruktionsverfahren die Verteilung der Positronen-Emissionen geschätzt werden. Häufig Verwendung findende Rekonstruktionsverfahren sind zum Beispiel die „Algebraic Reconstruction Technique“ (ART) von Gordon (1974) [22], der EM-Algorithmus und das Convolved Backprojection Verfahren, Vardi et al. (1985) [58].

Time-of-Flight Positronen-Emissions-Tomographie (ToFPET) ist ein verfeinertes PET-Verfahren, bei dem zusätzlich die Zeitdifferenz  $\Delta t_j$  gemessen wird, mit der die Photonen des  $j$ -ten Paares auf Detektoren des Ringes auftreffen. Während bei PET lediglich Flugbahnen beobachtet werden, kann bei ToFPET aus der Kenntnis der fehlerbehafteten Flugzeitdifferenz und der Flugbahn der Emissionsort approximativ bestimmt werden. Die Messung der Zeitdifferenz erreicht man durch den Einsatz spezieller Szintillationsdetektoren. Obwohl sich die der Time-of-Flight PET zugrundeliegende Idee bis 1966 zurückverfolgen läßt, war es bis vor wenigen Jahren nicht möglich, Detektoren zu konstruieren, die die Flugzeitdifferenz mit ausreichender Genauigkeit messen, um den Fehler bei der Berechnung des Emissionsortes im Rahmen der Größenordnung von 1-2 cm zu halten. ToFPET befindet sich daher zur Zeit noch im Versuchsstadium.

Aus der bei ToFPET gewonnenen Richtungs- und Zeitinformation über die Flugbahnen der Photonenpaare kann anschließend die Emissionsdichte der Positronen-Emissionen rekonstruiert werden. In Snyder et al. (1981) [49] findet man einen ersten generellen Ansatz für ToFPET-Rekonstruktionsalgorithmen. In einer späteren Arbeit von Snyder and Politte (1983) [48] wird ein Maximum-Likelihood Schätzverfahren mit dem EM-Algorithmus kombiniert. In beiden Fällen werden keine quantitativen Untersuchungen zur Güte der Schätzverfahren vorgenommen.

Eine weitreichende mathematische Analyse der Positronen-Emissions-Tomographie (ohne Time-of-Flight Information) findet man in Johnstone und Silverman (1990) [31]. Dort wird Gehirn-PET als inverses Problem behandelt. Die Modellierung des PET-Problems geht von einer idealisierten Beschreibung des experimentellen Ablaufes aus. Diese Idealisierung erweist sich auch für die mathematische Behandlung des ToFPET-Rekonstruktionsproblems als geeignet und wird in dieser Arbeit teilweise übernommen.

Das Modell und seine Annahmen sind im Folgenden erläutert (siehe auch Abbildung 1.1).  $\mathcal{Q} \subset \mathbb{R}^2$  bezeichnet die Ebene eines Querschnittes durch den Kopf des Patienten. Es wird ein Detektorring  $\mathcal{D}$  betrachtet, der als Kreis in der Ebene des Querschnittes dargestellt wird und die Fläche  $\mathcal{B}$  begrenzt. Die tatsächlichen Emissionsorte  $\mathbf{Y}_j$  werden dann allgemein als Punkte in  $\mathcal{B}$  modelliert, die Flugbahn der Photonen als Gerade in  $\mathcal{B}$ . Die Detektoren werden als beliebig hoch auflösend angenommen, so daß die Orte der Registrierung als Punkte betrachtet und exakt ermittelt werden können. Es werden die Winkel  $\varphi_j$  der Flugbahnen der Photonen, sowie die Zeitdifferenz  $\Delta t_j$  des Auftreffens beider Photonen eines Paares auf dem Detektorring ermittelt. Daraus wird ein ungefährer Emissionsort  $\mathbf{X}_j \in \mathcal{B}$  berechnet (siehe Abbildung 1.1). Es wird angenommen, daß  $n$  unabhängige Emissionen registriert werden. Ein Datensatz besteht dann aus den Punkten  $\mathbf{X}_j \in \mathcal{B}$ ,  $j = 1, \dots, n$ , den beobachteten Emissionsorten, und den Winkeln  $\varphi_j \in [0, \pi]$ ,  $j = 1, \dots, n$ . Die Punkte  $\mathbf{X}_j$  werden als Realisierungen von Zufallsvektoren betrachtet, die aus den eigentlichen Emissionsorten  $\mathbf{Y}_j$  durch Überlagerung mit einem Zufallsfehler  $\mathbf{Z}_j$  hervorgehen.  $\mathbf{Z}_j$  modelliert den Fehler bei der Rekonstruktion des Ortes der Positronen-Emission aus Richtungs- und Zeitinformation. Die Fehler  $\mathbf{Z}_1, \dots, \mathbf{Z}_n$  sind wegen mangelnder Genauigkeit bei der Messung von  $\Delta t_j$  nicht vernachlässigbar. Sie werden bei der Rekonstruktion der Emissionsdichte berücksichtigt. Andere Faktoren wie Streuung der Photonen, Bewegung der Positronen vor der Annihilation usw. sind im Vergleich zur Größenordnung des Meßfehlers vernachlässigbar und bleiben unberücksichtigt. Der Einfluß des Meßfehlers kann dazu führen, daß Beobachtungen  $\mathbf{X}_j$  außerhalb von  $\mathcal{B}$  anfallen. Vorabwissen über die Form des Schädels kann zur Vorbereitung des Datensatzes  $\mathbf{X}_1, \dots, \mathbf{X}_n$  verwendet werden, um Datenpunkte - zum Beispiel

außerhalb einer Umgebung der äußeren Schädelgrenze - zu entfernen bevor die Rekonstruktion durchgeführt wird.

In dieser Arbeit wird ein Verfahren entwickelt, um aus den bei ToFPET gewonnenen Meßdaten eine Schätzfunktion für die gesuchte Emissionsdichte zu bestimmen. Die Eigenschaften und die Qualität dieser Schätzfunktion werden untersucht. Wir beweisen, daß das vorgeschlagene Rekonstruktionsverfahren hinsichtlich der Konvergenzraten für wichtige lokale und globale Fehlermaße optimal ist. Dies geschieht indem Minimax-Raten über bestimmte Funktionen-Klassen, in diesem Fall Sobolev-Klassen der Ordnung  $\alpha > 1$ , ermittelt werden (die von keinem Rekonstruktionsverfahren unterschritten werden können) und gezeigt wird, daß das vorgeschlagene Rekonstruktionsverfahren diese optimalen Raten erreicht. In Zhang (1992) [61] wird ToFPET als Dekonvolutionsproblem in Hölder-Klassen von Funktionen behandelt. In der vorliegenden Arbeit wird die Klasse der zulässigen Dichtefunktionen erweitert auf Sobolev-Klassen der Ordnung  $\alpha > 1$ . Auf diese Weise wird eine Reduzierung der Glattheitsanforderungen an die bei ToFPET auftretenden Emissionsdichtefunktionen erreicht, was im Einklang mit praktischen Erfordernissen steht.

Der Schätzalgorithmus, der in dieser Arbeit entwickelt wird, besitzt ideale Voraussetzungen für eine Parallelisierung. Im Anschluß an die theoretischen Untersuchungen wird die Implementierung des Algorithmus beschrieben (sequentiell und parallel). Rekonstruktionsergebnisse für simulierte Datensätze werden dargestellt. Die theoretischen Konvergenzraten werden mit Simulationsergebnissen verglichen.

In **Kapitel 2** wird das mathematische Modell für die Behandlung des Schätzproblems vorgestellt. Ein modifiziertes Dekonvolutions-Kernschätzverfahren für die Rekonstruktion der Emissionsdichtefunktion wird hergeleitet. In **Kapitel 3** werden Eigenschaften der Schätzfunktion hergeleitet. Die Wahl der erforderlichen Modellparameter wird diskutiert. Die Klasse der betrachteten Emissionsdichten wird als Sobolev-Raum der Ordnung  $\alpha > 1$  eingeführt und diese Wahl wird motiviert. Für den so festgelegten Raum der Emissionsdichten wird die Güte des modifizierten Dekonvolutions-Kernschätzverfahrens unter den Fehlermaßen Mittlerer Quadratischer Fehler, Mittlerer Integrierter Quadratischer Fehler und unter der Supremum-Norm untersucht. Es werden obere Schranken für die Konvergenzgeschwindigkeit der Schätzfunktion unter den betrachteten Fehlermaßen ermittelt. In **Kapitel 4** werden prinzipielle - für alle Rekonstruktionsverfahren gültige - untere Schranken für die Konvergenzgeschwindigkeit der rekonstruierten Dichtefunktion unter Verwendung der Fehlermaße von Kapitel 3 untersucht. Diese stimmen mit den oberen Schranken für das vorgeschlagene Rekonstruktionsver-

fahren überein. Damit wird die Optimalität des Rekonstruktionsverfahrens und der in Kapitel 3 erhaltenen Konvergenzraten gezeigt. Eine Einordnung der Ergebnisse in die Literatur zu diesem Gebiet befindet sich am Ende von Kapitel 4. In **Kapitel 5** wird die Umsetzung des Schätzverfahrens in einen in der Praxis einsetzbaren Algorithmus behandelt. Die Implementierung des Schätzverfahrens wird beschrieben. Ein automatisiertes Verfahren für die Wahl des Bandweite-Parameters wird angegeben. Die theoretischen Resultate werden anhand simulierter Testdaten überprüft. Der vollständige parallele C-Programmcode ist im **Anhang** beschrieben.

# Kapitel 2

## Schätzung der Positronen Emissionsdichte

In diesem Kapitel wird basierend auf ToFPET-Meßdaten eine Schätzfunktion für die Positronen-Emissionsdichte konstruiert. Durch die Messung von Photonen-Flugbahn und Flugzeitdifferenz kann der Ort der Photonen-Erzeugung und damit der Positronen-Emission bis auf einen stochastischen Fehler  $\mathbf{Z}$  genau bestimmt werden. Um aus den anfallenden Meßpunkten eine Dichtefunktion für die Positronen-Emission zu bestimmen, entwickeln wir ein Verfahren, das auf Dekonvolutionsprinzipien aufbaut, bei welchem Annahmen über die Verteilung des Beobachtungsfehlers in die Rekonstruktion der Dichtefunktion einfließen. Beispiele zur Anwendung von Dekonvolutions-Dichteschätzverfahren in  $\mathbb{R}$  oder  $\mathbb{R}^2$  finden sich jeweils in [25], [26], [27], [28], [51] und in [61].

### 2.1 Modell

Aus der Versuchsanordnung in Kapitel 1 wird ein mathematisches Modell abgeleitet.

**Definition 2.1** *Es sei  $\mathcal{D}$  der Detektorring,  $\mathcal{B}$  die vom Detektorring begrenzte Querschnittsfläche.*

1.  $\mathcal{D} := \{\mathbf{y} \in \mathbb{R}^2, \|\mathbf{y}\| = 1\}$ .
2.  $\mathcal{B} := \{\mathbf{y} \in \mathbb{R}^2, \|\mathbf{y}\| \leq 1\}$ .

### Bezeichnung 2.1

1. Mit  $\mathbf{Y}$  werde der Ort des Auftretens einer Emission bezeichnet.
2. Mit  $\mathbf{X}$  werde der Ort der Beobachtung einer Emission bezeichnet.
3.  $\mathbf{Z}$  bezeichne den Beobachtungsfehler.

Die Größen  $\mathbf{X}$  und  $\mathbf{Z}$  sind als Zufallsvektoren in  $\mathbb{R}^2$  modelliert und  $\mathbf{Y}$  als Zufallsvektor in  $\mathcal{B}$ . Die Zufallsvektoren  $\mathbf{X}$ ,  $\mathbf{Y}$  und  $\mathbf{Z}$  besitzen Verteilungen mit den Wahrscheinlichkeitsdichten  $g(\mathbf{x})$ ,  $f(\mathbf{y})$  und  $h(\mathbf{z})$ . Die Dichtefunktion  $f(\mathbf{y})$  ist die gesuchte Emissionsdichtefunktion. Der folgende Modellansatz wird für  $\mathbf{X}$  aufgestellt.

$$\mathbf{X} = \mathbf{Y} + \mathbf{Z} \quad (2.1)$$

Wenn mit  $\varphi$  der Flugwinkel der Photonenflugbahn bezeichnet wird und mit  $r$  die Größe (einschließlich Vorzeichen) des Beobachtungsfehlers, so kann der Beobachtungsfehler in der Form

$$\begin{aligned} \mathbf{Z} &= r \cdot (\cos(\varphi), \sin(\varphi)) \\ &=: r\Theta_{\varphi}. \end{aligned} \quad (2.2)$$

dargestellt werden. Unter der Annahme, daß  $n$  unabhängige Messungen durchgeführt werden, besteht eine Stichprobe aus

$$\{(\mathbf{X}_1, \varphi_1), \dots, (\mathbf{X}_n, \varphi_n)\},$$

wobei  $(\mathbf{X}_j, \varphi_j)$  für  $j = 1, \dots, n$  unabhängig identisch verteilt sind mit derselben Verteilung wie  $(\mathbf{X}, \varphi)$ . Hierbei ist

$$\mathbf{X}_j = \mathbf{Y}_j + r_j\Theta_{\varphi_j}, \quad j = 1, \dots, n.$$

Der Umfang  $n$  der Stichprobe liegt bei ToFPET in der Größenordnung von  $10^6$ . Folgende Modellannahmen werden festgelegt.



### Modellannahmen (M1):

1.  $\mathbf{Y}_j, j = 1, \dots, n$  sind u. i. v. Zufallsvariablen mit einer auf  $\mathcal{B}$  beschränkten Dichte  $f$ .
2.  $r_j, j = 1, \dots, n$  sind unabhängige  $\mathcal{N}(0, \sigma^2)$ -verteilte Zufallsvariablen.
3.  $\varphi_j, j = 1, \dots, n$  sind unabhängige  $\mathcal{U}([0, \pi])$ -verteilte Zufallsvariablen.
4.  $r_j$  und  $\varphi_j$  sind stochastisch unabhängig.
5.  $\mathbf{Y}_j$  und  $\mathbf{Z}_j$  sind stochastisch unabhängig.

**Anmerkungen:**  $\mathbf{Y}$  und  $\mathbf{Z}$  werden in (M1) als paarweise stochastisch unabhängig modelliert. Dies ist eine Vereinfachung, die zur Folge hat, daß man auch Beobachtungen außerhalb von  $\mathcal{B}$  zuläßt. Das Ausmaß dieses Effektes hängt von der Größe  $\sigma$ , der Fehlervarianz, ab. Die Fehlervarianz ist durch die Eigenschaften der Meßapparatur bestimmt. Die besten gegenwärtig für ToFPET zur Verfügung stehenden Geräte arbeiten bei Gehirn-Tomographie mit  $\sigma \approx 1 - 2$  cm. Da der Schädel im Querschnitt als etwa vom Radius 1 definiert wird, ergibt sich ein entsprechend skaliertes kleineres  $\sigma$ . Es liegt Vorwissen über die Form der Menge  $\mathcal{B}$  vor, daher kann man annehmen, daß die Stichprobe vorbearbeitet wurde, um Daten außerhalb einer gewissen Umgebung von  $\mathcal{B}$  auszuschließen. Die Modellierung der Flugwinkel als gleichverteilte Zufallsvariable entstammt der physikalisch in sehr guter Näherung bestätigten Annahme, daß keine Richtung im  $\mathbb{R}^3$  bei Aussenden eines Photonenpaares bevorzugt auftritt. Die Modellierung der Größe des Meßfehlers als normalverteilte Zufallsvariable ist eine Annahme, die häufig für die Genauigkeit physikalischer Meßinstrumente gemacht wird.

Die Beobachtung der tatsächlichen Emissionsorte  $\mathbf{Y}_j$  ist mit einem Fehler  $\mathbf{Z}_j$  überlagert, dessen Verteilung als bekannt angenommen wird. Die Beobachtungsdaten  $\mathbf{X}_j$  sind indirekte Messungen der  $\mathbf{Y}_j$ . Die unbekannte Dichte  $f$  der Zufallsvariablen  $\mathbf{Y}_j$  soll aus den Beobachtungen  $(\mathbf{X}_j, \varphi_j), j = 1, \dots, n$  mittels eines modifizierten Dekonvolutionsansatzes geschätzt werden.

## 2.2 Dekonvolutionsansatz

Das Problem,  $f(\mathbf{y})$  aus indirekten Beobachtungen  $\mathbf{X}_j, j = 1, \dots, n$ , zu schätzen, gehört zur Klasse der inversen Probleme. Wegen (2.1) und (M1) kann die Dichte  $g(\mathbf{x})$  als Faltung dargestellt werden:

$$\begin{aligned}
g(\mathbf{x}) &= \int_{\mathcal{B}} f(\mathbf{y}) \cdot h(\mathbf{x} - \mathbf{y}) \, d\mathbf{y} \\
&= (f * h)(\mathbf{x}).
\end{aligned}
\tag{2.3}$$

Die Emissionsdichte  $f$  ist in dieser Faltung identifizierbar, wenn

$$\int_{\mathcal{B}} f(\mathbf{y}) h(\mathbf{x} - \mathbf{y}) d\mathbf{y} = \int_{\mathcal{B}} \tilde{f}(\mathbf{y}) h(\mathbf{x} - \mathbf{y}) d\mathbf{y} \Leftrightarrow f(\mathbf{y}) = \tilde{f}(\mathbf{y}), \quad \text{f. ü.}$$

Dies kann mittels der Fouriertransformierten ausgedrückt werden. Die Fouriertransformierte einer Dichte ist wie folgt definiert.

**Definition 2.2** *Es sei  $w(\mathbf{y}) \in \mathcal{L}^1(\mathbb{R}^2)$ . Die Funktion*

$$\Phi_w(\mathbf{t}) := \int_{\mathbb{R}^2} e^{it' \mathbf{y}} \cdot w(\mathbf{y}) d\mathbf{y}$$

*heißt Fouriertransformierte oder charakteristische Funktion von  $w(\mathbf{y})$  an der Stelle  $\mathbf{t} \in \mathbb{R}^2$ .*

**Bezeichnung 2.2** *Es bezeichne  $w(\mathbf{y})$  von nun an eine beliebige Dichtefunktion in  $\mathcal{L}^1(\mathbb{R}^2)$ , nicht in jedem Fall dieselbe.*

**Bezeichnung 2.3** *Für eine Dichtefunktion  $w(\mathbf{y})$  bezeichnen*

1.  $\hat{w}_n(\mathbf{y})$  eine Schätzfunktion für die Funktion  $w(\mathbf{y})$  aus  $n$  Daten,
2.  $\hat{\Phi}_{w,n}(\mathbf{t})$  eine Schätzfunktion für die charakteristische Funktion von  $w(\mathbf{y})$  aus  $n$  Daten.

Nach Anwenden der Fouriertransformation auf beiden Seiten von (2.3) ergibt sich:

$$\begin{aligned}
\Phi_g(\mathbf{t}) &= \Phi_f(\mathbf{t}) \cdot \Phi_h(\mathbf{t}) \\
\Rightarrow \Phi_f(\mathbf{t}) &= \Phi_g(\mathbf{t}) \cdot \Phi_h(\mathbf{t})^{-1}
\end{aligned}
\tag{2.4}$$

sofern  $\Phi_h(\mathbf{t}) \neq 0$ . Für festes  $\Phi_g(\mathbf{t})$  in (2.4) gilt, daß  $\Phi_f$  identifizierbar ist, wenn

$$\Phi_f(\mathbf{t}) \Phi_h(\mathbf{t}) = \Phi_{\tilde{f}}(\mathbf{t}) \Phi_h(\mathbf{t}) \Leftrightarrow \Phi_f(\mathbf{t}) = \Phi_{\tilde{f}}(\mathbf{t}). \quad (2.5)$$

Damit Bedingung 2.5 erfüllt ist, ist es hinreichend zu zeigen, daß  $\Phi_h(\mathbf{t}) \neq 0$  für alle  $\mathbf{t}$ . Unter Verwendung von (M1) kann die Fouriertransformierte  $\Phi_h(\mathbf{t})$  berechnet werden.

$$\Phi_h(\mathbf{t}) = \int_{\mathbb{R}^2} e^{i\mathbf{t}'\mathbf{z}} h(\mathbf{z}) d\mathbf{z}.$$

Wird  $\mathbf{z}$  in Polarkoordinaten dargestellt, ergibt sich

$$\Phi_h(\mathbf{t}) = \int_{\mathbb{R}} \int_0^\pi e^{i\mathbf{t}'r(\cos(\varphi), \sin(\varphi))'} h(r \cos(\varphi), r \sin(\varphi)) r dr d\varphi.$$

Man erhält mit dem Transformationssatz für Dichten

$$\begin{aligned} h(r \cos(\varphi), r \sin(\varphi)) &= \frac{1}{r} w(r) w(\varphi). \\ &= \frac{1}{r} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{r^2}{2\sigma^2}} \frac{1}{\pi}. \end{aligned}$$

Für die Fouriertransformation  $\Phi_h(\mathbf{t})$  folgt dann:

$$\begin{aligned} \Phi_h(\mathbf{t}) &= \int_{\mathbb{R}} \int_0^\pi e^{i\mathbf{t}'r(\cos(\varphi), \sin(\varphi))'} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{r^2}{2\sigma^2}} \frac{1}{\pi} dr d\varphi \\ &= \frac{1}{\pi} \int_0^\pi e^{-(\mathbf{t}'(\cos(\varphi), \sin(\varphi)))^2 \frac{\sigma^2}{2}} d\varphi \\ &= \frac{1}{\pi} \int_0^\pi e^{-\|\mathbf{t}\|^2 |\cos(\angle(\mathbf{t}, (\cos(\varphi), \sin(\varphi))))|^2 \frac{\sigma^2}{2}} d\varphi \\ &= \frac{1}{\pi} \int_0^\pi e^{-\|\mathbf{t}\|^2 |\cos(\vartheta)|^2 \frac{\sigma^2}{2}} d\vartheta. \end{aligned}$$

Für  $\|\mathbf{t}\| = 0$  ergibt sich  $\Phi_h(\mathbf{t}) = 1$ . Für  $\|\mathbf{t}\| > 0$  wird folgende Abschätzung nach [61] verwendet

$$\Phi_h(\mathbf{t}) \geq 2 \cdot \frac{1}{\pi} \int_0^{\frac{\pi}{2}} e^{-\|\mathbf{t}\|^2 \varphi^2 \frac{\sigma^2}{2}} d\varphi.$$

Mit der Substitution  $u = \varphi \|\mathbf{t}\|$  ergibt sich

$$\Phi_h(\mathbf{t}) \geq \frac{2}{\pi} \|\mathbf{t}\|^{-1} \int_0^{\frac{\pi}{2} \|\mathbf{t}\|} e^{-u^2 \frac{\sigma^2}{2}} du. \quad (2.6)$$

und somit  $\Phi_h(\mathbf{t}) \neq 0$  für alle  $\mathbf{t}$ .

Damit ist die Identifizierbarkeit von  $\Phi_f(\mathbf{t})$  und, wegen der Eindeutigkeit der Zuordnung zwischen charakteristischer Funktion und Dichtefunktion, auch die Identifizierbarkeit von  $f(\mathbf{y})$  gezeigt.  $\Phi_f(\mathbf{t})$  ist über Gleichung (2.4) für alle  $\mathbf{t}$  definiert.

## 2.3 Konstruktion der Schätzfunktion

Um  $\Phi_f(\mathbf{t})$  aus (2.4) zu bestimmen, wird zunächst eine Schätzfunktion für  $\Phi_g(\mathbf{t})$  aus den Beobachtungsdaten bestimmt.

**Definition 2.3 (Kernfunktion)** *Es sei  $K(x_1, x_2) : \mathbb{R}^2 \rightarrow \mathbb{R}$  eine Funktion mit folgenden Eigenschaften:*

1.  $K(x_1, x_2) = K_1(x_1) \cdot K_2(x_2).$

Für  $i = 1, 2$  gelte:

2.  $K_i(x_i) = K_i(-x_i),$

3.  $|K_i(x_i)| \leq C < \infty,$

4.  $\int_{\mathbb{R}} K_i(x_i) dx_i = 1,$

5.  $\int_{\mathbb{R}} x_i \cdot K_i(x_i) dx_i = 0.$

Dann heißt  $K(x_1, x_2)$  Kernfunktion.

**Definition 2.4 (Kernschätzer)**  $\mathbf{X}_1, \dots, \mathbf{X}_n$  bezeichnen  $n$  unabhängige Beobachtungen eines Zufallsvektors  $\mathbf{X} \in \mathbb{R}^2$  mit Dichtefunktion  $g(\mathbf{x})$  und  $(\lambda_n)_{n \in \mathbb{N}}$  sei eine Folge in  $\mathbb{R}^+$  mit  $\lambda_n \rightarrow 0$  und  $n\lambda_n^2 \rightarrow \infty$  für  $n \rightarrow \infty$ .  $K(\mathbf{x})$  sei eine Kernfunktion. Die Funktion

$$\hat{g}_n(\mathbf{x}) := \frac{1}{n\lambda_n^2} \sum_{j=1}^n K\left(\frac{\mathbf{x}-\mathbf{X}_j}{\lambda_n}\right)$$

heißt Kernschätzer für  $g(\mathbf{x})$  mit Bandweite  $\lambda_n$ .

In Silverman (1986) [46] werden Kernschätzer als eine Summe von „Peaks“ über den Beobachtungen beschrieben, wobei die Kernfunktion die Form der „Peaks“ festlegt, die Bandweite deren Ausdehnung. Man findet in [46] Näheres über Theorie und Eigenschaften von Kernschätzern. Weitere interessante Quellen hierzu sind [45] und [51].

Die Fouriertransformierte der Funktion  $\hat{g}_n(\mathbf{x})$  an der Stelle  $\mathbf{t} \in \mathbb{R}^2$  hat folgende Form:

$$\Phi_{\hat{g}_n}(\mathbf{t}) = \frac{1}{n} \sum_{j=1}^n e^{it' \mathbf{X}_j} \cdot \Phi_K(\lambda_n \mathbf{t}).$$

Es ergibt sich dann mit (2.4)

$$\hat{\Phi}_{f_n}^{(1)}(\mathbf{t}) := \frac{1}{n} \sum_{j=1}^n e^{it' \mathbf{X}_j} \cdot \frac{\Phi_K(\lambda_n \mathbf{t})}{\Phi_h(\mathbf{t})} \quad . \quad (2.7)$$

Die Verwendung eines Kernschätzers für  $g(\mathbf{x})$  läßt sich wie folgt motivieren. Wenn  $\mathbf{X}_1, \dots, \mathbf{X}_n$  unabhängige Beobachtungen der Dichte  $g(\mathbf{x})$  sind, dann bezeichnet

$$\frac{1}{n} \sum_{j=1}^n e^{it' \mathbf{X}_j}$$

die **empirische charakteristische Funktion** der  $\mathbf{X}_j$  und stellt einen erwartungstreuen Schätzer für  $\Phi_g(\mathbf{t})$  dar. Unter Verwendung von (2.4) erhält man dann den folgenden Schätzer für die Fouriertransformierte von  $f(\mathbf{y})$ :

$$\hat{\Phi}_{f_n}^{(2)}(\mathbf{t}) := \frac{\frac{1}{n} \sum_{j=1}^n e^{it' \mathbf{X}_j}}{\Phi_h(\mathbf{t})}. \quad (2.8)$$

Die Schätzfunktion  $\hat{f}_n^{(2)}(\mathbf{y})$  ergibt sich durch die Anwendung der inversen Fouriertransformation.

$$\hat{f}_n(\mathbf{y}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \frac{\frac{1}{n} \sum_{j=1}^n e^{it' \mathbf{X}_j}}{\Phi_h(\mathbf{t})} dt. \quad (2.9)$$

Die Existenz des Integrales in (2.9) wird durch Hinzunahme einer Funktion  $\Phi_K(\lambda_n \mathbf{t})$  gesichert, welche das Abklingverhalten der Funktion  $\Phi_h(\mathbf{t})$  ausgleicht. Man nennt diese Funktion *Fensterfunktion*.  $\Phi_K(\lambda_n \mathbf{t})$  ist die Fouriertransformierte einer Funktion  $(1/\lambda_n^2)K(\mathbf{x}/\lambda_n)$ , wobei  $K(\mathbf{x})$  eine Kernfunktion ist und  $\lambda_n$  der in Definition 2.4 eingeführte Bandweiteparameter.

Aus (2.7) ergibt sich folgende Schätzfunktion für  $f(\mathbf{y})$ :

$$\begin{aligned} \hat{f}_n^{(1)}(\mathbf{y}) &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \frac{1}{n} \sum_{j=1}^n e^{it' \mathbf{X}_j} \frac{\Phi_K(\lambda_n \mathbf{t})}{\Phi_h(\mathbf{t})} dt \\ &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \hat{\Phi}_{f_n}^{(1)}(\mathbf{t}) dt. \end{aligned}$$

$\hat{f}_n^{(1)}(\mathbf{y})$  heißt **Dekonvolutions–Kernschätzer** von  $f$  an der Stelle  $\mathbf{y}$ . Beispiele für Dekonvolutions–Kernschätzung können unter anderem in [51] und [61] gefunden werden.

Die Beobachtung der Winkel  $\varphi_j$  beim ToFPET–Verfahren stellt zusätzliche Information dar. Diese Zusatzinformation ermöglicht es nun, das herkömmliche Dekonvolutions–Schätzverfahren zu erweitern. Wie aus (2.9) hervorgeht, fließen die Beobachtungsorte  $\mathbf{X}_j$  über die Faktoren  $e^{it' \mathbf{X}_j}$  in die Berechnung der Schätzfunktion ein. Unter Verwendung von (2.1)-(2.2) ergibt sich

$$e^{it' \mathbf{X}_j} = e^{it' (\mathbf{Y}_j + r \Theta_{\varphi_j})} = e^{it' \mathbf{Y}_j} e^{it' r \Theta_{\varphi_j}}.$$

Für  $\mathbf{t}' \Theta_{\varphi_j} = 0$  ist

$$e^{it' \mathbf{X}_j} = e^{it' \mathbf{Y}_j}.$$

Für einen Punkt  $\mathbf{t} \in \mathbb{R}^2$  liefern daher Beobachtungen  $(\mathbf{X}_j, \varphi_j)$  für die  $\mathbf{t}' \Theta_{\varphi_j} \approx 0$  etwa genau soviel Information bezüglich der empirischen charakteristischen Funktion  $\frac{1}{n} \sum_{j=1}^n e^{it' \mathbf{X}_j}$  an der Stelle  $\mathbf{t}$  wie die fehlerfreien  $\mathbf{Y}_j$ . Für solche  $(\mathbf{X}_j, \varphi_j)$  verschwindet der Beobachtungsfehler  $\mathbf{Z}_j$  aus dem Schätzalgorithmus. Der Einfluß dieser Daten auf das Gesamtergebnis soll höher gewichtet werden als der Einfluß anderer Daten für welche  $\mathbf{t}' \Theta_{\varphi_j} \neq 0$  ist. Daher wird eine Gewichtsfunktion eingesetzt, deren Eigenschaften in der folgenden Definition gegeben sind.

**Definition 2.5 (Gewichtsfunktion)** Es sei  $\mathcal{D}' := \{(\cos(\varphi), \sin(\varphi)), \varphi \in [0, \pi]\}$ . Eine Funktion  $G : \mathbb{R}^2 \times \mathcal{D}' \rightarrow ]0, 1]$  heißt Gewichtsfunktion, wenn für  $\mathbf{t} \in \mathbb{R}^2$ ,  $\Theta \in \mathcal{D}'$  gilt:

- i.  $G(\mathbf{t}, \Theta)$  ist eine monoton fallende Funktion von  $|\mathbf{t}'\Theta|$ ,
- ii.  $\int_{\mathbb{R}^2} \int_{\mathcal{D}'} G(\mathbf{t}, \Theta) dt d\Theta < \infty$ .

**Satz 2.1** Es sei  $\mathbf{t} \in \mathbb{R}^2$ ,  $\Theta_\varphi \in \mathcal{D}'$ .  $G(\mathbf{t}, \Theta_\varphi)$  sei eine Gewichtsfunktion gemäß Definition 2.5. Die Funktion

$$\hat{\Phi}_{f^n}^{(3)}(\mathbf{t}) := \frac{1}{n} \sum_{j=1}^n e^{it' \mathbf{x}_j} \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})}$$

ist ein erwartungstreuer Schätzer der charakteristischen Funktion  $\Phi_f(\mathbf{t})$ , wobei  $M(\mathbf{t})$  definiert ist als

$$M(\mathbf{t}) := E[e^{it' r_{\Theta_\varphi}} G(\mathbf{t}, \Theta_\varphi)].$$

**Beweis:** Zu zeigen ist

$$E \left[ \frac{1}{n} \sum_{j=1}^n \left( e^{it' \mathbf{x}_j} \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \right) \right] = \Phi_f(\mathbf{t}).$$

Da  $(\mathbf{X}_j, \varphi_j)$  für  $j = 1, \dots, n$  unabhängig identisch verteilt sind mit der gleichen Verteilung wie  $(\mathbf{X}, \varphi)$ , gilt:

$$E \left[ \frac{1}{n} \sum_{j=1}^n \left( e^{it' \mathbf{x}_j} \cdot \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \right) \right] = E \left[ e^{it' \mathbf{X}} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} \right]$$

Über die Eigenschaften des bedingten Erwartungswertes erhält man

$$\begin{aligned} E[e^{it' \mathbf{X}} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})}] &= \int_{\mathcal{B}} E[e^{it' \mathbf{X}} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} | \mathbf{y}] f(\mathbf{y}) d\mathbf{y} \\ &= \int_{\mathcal{B}} \int_0^\pi \int_{\mathbb{R}^2} e^{it' \mathbf{x}} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} w(\mathbf{x}, \varphi | \mathbf{y}) f(\mathbf{y}) d\mathbf{x} d\varphi d\mathbf{y}. \end{aligned}$$

Bei gegebenem  $\mathbf{Y} = \mathbf{y}$  sind  $\mathbf{X}$  und  $\varphi$  abhängige Größen. Für  $\mathbf{Y} = \mathbf{y}$  und für festes  $\varphi$ , kann  $\mathbf{X}$  nur Werte auf der Kurve  $L_x := \{\mathbf{x} \in \mathbb{R}^2, \mathbf{x} = \mathbf{y} + r\Theta_\varphi\}$  annehmen.

Es wird angenommen, daß alle hier auftretenden Dichtefunktionen beschränkt sind. Es wird zunächst die Substitution  $\mathbf{X} = \mathbf{y} + \mathbf{Z}$  vorgenommen. Bei gegebenem  $\varphi$  nimmt  $\mathbf{Z}$  Werte auf der Kurve  $L_z := \{\mathbf{z} = (z_1, z_2), \frac{z_2}{z_1} = \tan(\varphi)\}$  an.

$$\begin{aligned} E[e^{it' \mathbf{X}} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})}] &= \int_{\mathcal{B}} \int_0^\pi \int_{L_z} e^{it' \mathbf{y}} e^{it' \mathbf{z}} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} w(\mathbf{z}|\varphi, \mathbf{y}) w(\varphi|\mathbf{y}) f(\mathbf{y}) dz d\varphi d\mathbf{y} \\ &= \int_{\mathcal{B}} \int_0^\pi \int_{\mathbb{R}} e^{it' \mathbf{y}} e^{it_1 z_1} e^{it_2 z_1 \tan(\varphi)} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} w(z_1|\varphi) w(\varphi) f(\mathbf{y}) dz_1 d\varphi d\mathbf{y}. \end{aligned}$$

Die Variablen  $\varphi$  und  $z_1$  sind unabhängig von  $\mathbf{Y}$ . Die Verteilung von  $\varphi$  ist gegeben durch  $w(\varphi) = \frac{1}{\pi}$ . Nun wird die Koordinatentransformation  $z_1 = r \cos(\varphi)$  durchgeführt, sowie die Transformation der Dichten  $w(r \cos(\varphi)|\varphi) = \frac{1}{\cos(\varphi)} w(r|\varphi)$ . Da  $r$  unabhängig von  $\varphi$  ist, gilt  $w(r|\varphi) = w(r)$ .

$$\begin{aligned} E[e^{it' \mathbf{X}} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})}] &= \int_{\mathcal{B}} \int_0^\pi \int_{\mathbb{R}} e^{it' \mathbf{y}} e^{it_1 r \cos(\varphi)} e^{it_2 r \sin(\varphi)} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} w(r) \frac{1}{\pi} f(\mathbf{y}) dr d\varphi d\mathbf{y} \\ &= \int_{\mathcal{B}} \int_0^\pi \int_{\mathbb{R}} e^{it' (\mathbf{y} + r\Theta_\varphi)} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} w(r) \frac{1}{\pi} f(\mathbf{y}) dr d\varphi d\mathbf{y} \\ &= \int_{\mathcal{B}} e^{it' \mathbf{y}} f(\mathbf{y}) d\mathbf{y} \int_{\mathbb{R}} \int_0^\pi e^{it' r\Theta_\varphi} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} w(r) \frac{1}{\pi} dr d\varphi \\ &= \Phi_f(\mathbf{t}) \frac{E[e^{it' r\Theta_\varphi} G(\mathbf{t}, \Theta_\varphi)]}{M(\mathbf{t})} \\ &= \Phi_f(\mathbf{t}). \end{aligned}$$

Die letzte Gleichung folgt aus der Definition von  $M(\mathbf{t})$ .

**q.e.d.**

Die Schätzfunktion für  $f(\mathbf{y})$  motiviert sich nun aus diesen Vorüberlegungen.

### **Definition 2.6 (Modifizierter Dekonvolutions–Kernschätzer)**

Gegeben seien  $n$  u. i. v. Beobachtungen  $(\mathbf{X}_j, \varphi_j)$ ,  $j = 1, \dots, n$  für das Schätzproblem (2.1)-(2.2). Gegeben sei weiterhin eine Parameterfolge  $(\lambda_n)_{n \in \mathbb{N}}$  in  $\mathbb{R}^+$  mit



der Eigenschaft  $\lambda_n \rightarrow 0$  und  $n\lambda_n^2 \rightarrow \infty$  für  $n \rightarrow \infty$ . Es sei  $K(\mathbf{y})$  eine Kernfunktion gemäß Definition 2.3 und  $G(\mathbf{t}, \Theta_\varphi)$  sei eine Gewichtsfunktion gemäß Definition 2.5. Die Funktion  $M(\mathbf{t})$  sei definiert als  $M(\mathbf{t}) := E[e^{i\mathbf{t}'\Theta_\varphi}G(\mathbf{t}, \Theta_\varphi)]$ . Die durch

$$\hat{f}_n(\mathbf{y}) := \frac{1}{(2\pi)^{2n}} \sum_{j=1}^n \int_{\mathbf{t} \in \mathbb{R}^2} e^{-i\mathbf{t}'(\mathbf{y}-\mathbf{x}_j)} \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) d\mathbf{t} \quad \forall \mathbf{y} \in \mathcal{B} \quad (2.10)$$

gegebene Funktion heißt **modifizierter Dekonvolutions–Kernschätzer für  $f(\mathbf{y})$** .

Der modifizierte Dekonvolutions–Kernschätzer stellt eine Rekonstruktion der Positronen–Emissionsdichte  $f(\mathbf{y})$  aus ToFPET–Meßdaten dar. In die Berechnung fließen die Beobachtungen  $(X_j, \varphi_j)$ ,  $j = 1, \dots, n$  ein. Modellannahmen über die Verteilung des Meßfehlers fließen in die Konstruktion der Funktion  $M(\mathbf{t})$  ein. Die Funktion  $\Phi_K(\lambda_n \mathbf{t})$  garantiert bei geeigneter Wahl der Kernfunktion  $K(\mathbf{y})$  und geeigneter Wahl von  $\lambda_n$  die Existenz des obigen Integrales. Der Bandweiteparameter  $\lambda_n$  steuert das Abklingverhalten der Funktion  $\Phi_K(\lambda_n \mathbf{t})$ . Es wird im Laufe der nächsten Kapitel erläutert, daß die optimale Wahl des Bandweiteparameters von großer Bedeutung ist. Die optimale Bandweite hängt ab vom jeweils betrachteten Fehlermaß und von der zu Grunde liegenden Funktionenklasse. Kleine Bandweitenparameter sind geeignet für Funktionen mit geringer Glattheit, größere Bandweitenparameter sind geeignet für Funktionen mit größerer Glattheit [46]. Wird die Bandweite zu klein gewählt, zeigt die resultierende Schätzfunktion für die betrachtete Funktionenklasse zu große Schwankungen. Wird die Bandweite zu groß gewählt, ist die berechnete Schätzfunktion gegenüber der zu schätzenden zu glatt. In beiden Fällen ist der Fehler bei der Rekonstruktion von  $f(\mathbf{y})$  größer als bei der optimalen Wahl von  $\lambda_n$  für ein gegebenes Fehlermaß.

Im Vergleich zur herkömmlichen Dekonvolutions–Dichteschätzung wie sie zum Beispiel in [51] zu finden ist, fließt in den modifizierten Dekonvolutions–Kernschätzer Zusatzinformation über die Richtung  $\varphi$  des Meßfehlers  $\mathbf{Z}$  mittels der Gewichtsfunktion  $G(\mathbf{t}, \Theta_\varphi)$  ein. Diese Funktion gewichtet Beobachtungen hinsichtlich ihrer Qualität, indem solche Paare  $(\mathbf{X}_j, \Theta_{\varphi_j})$  besonders hervorgehoben werden, für die der Richtungsvektor des Beobachtungsfehlers orthogonal zu  $\mathbf{t}$  ist, und für die damit der Beobachtungsfehler  $\mathbf{Z}_j$  im Ausdruck  $e^{i\mathbf{t}'\mathbf{x}_j}$  eliminiert wird.

Im Rahmen dieser Arbeit sollen nun quantitative Eigenschaften des modifizierten Dekonvolutions–Kernschätzers  $\hat{f}_n(\mathbf{y})$  untersucht werden. Zunächst werden im

nächsten Kapitel Aussagen über das Verhalten von Bias und Varianz der Schätzfunktion  $\hat{f}_n(\mathbf{y})$  in Abhängigkeit von den Funktionen  $K(\mathbf{y})$  und  $G(\mathbf{t}, \Theta_\varphi)$  getroffen. Danach werden  $K(\mathbf{y})$  und  $G(\mathbf{t}, \Theta_\varphi)$  fest gewählt, und das Konvergenzverhalten von  $\hat{f}_n(\mathbf{y})$  gegen die tatsächliche Emissionsdichte  $f(\mathbf{y})$  wird für diese Wahl unter verschiedenen Fehlermaßen untersucht.

# Kapitel 3

## Eigenschaften des modifizierten Dekonvolutions–Kernschätzers

In diesem Kapitel werden quantitative Eigenschaften des modifizierten Dekonvolutions–Kernschätzers hergeleitet. Zunächst wird das Verhalten von Bias und Varianz von  $\hat{f}_n$  unter der Annahme  $f \in \mathcal{L}^1(\mathbb{R}^2)$  und bei beliebiger Wahl von Kernfunktion  $K(\mathbf{y})$  und Gewichtsfunktion  $G(\mathbf{t}, \Theta_\varphi)$  untersucht. Danach wird die Funktionenklasse der Emissionsdichten definiert. Kernfunktion und Gewichtsfunktion werden fest gewählt. Für diesen Fall werden obere Schranken für das Verhalten von  $(\hat{f}_n(\mathbf{y}) - f(\mathbf{y}))$  unter verschiedenen Fehlermaßen bestimmt.

### 3.1 Bias und Varianz der Schätzfunktion

Der modifizierte Dekonvolutions–Kernschätzer Schätzer  $\hat{f}_n(\mathbf{y})$  ist definiert als die inverse Fouriertransformierte der Funktion

$$\frac{1}{n} \sum_{j=0}^n e^{it' \mathbf{x}_j} \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t})$$

an der Stelle  $\mathbf{y}$  (siehe Definition 2.6). Die Emissionsdichte  $f(\mathbf{y})$  ist eine reellwertige Funktion, daher wird zunächst überprüft, ob die Schätzfunktion  $\hat{f}_n(\mathbf{y})$  ebenfalls reellwertig ist.

**Bemerkung 3.1**  $\Phi_w(\mathbf{t})$  bezeichne die Fouriertransformation einer beliebigen Dichtefunktion  $w(\mathbf{y})$  wie in Definition 2.2 eingeführt. Dann gilt:

- Parseval's Gleichung:

$$\int_{\mathbb{R}^2} |w(\mathbf{y})|^2 d\mathbf{y} = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} |\Phi_w(\mathbf{t})|^2 d\mathbf{t}.$$

- Ist  $w(\mathbf{y})$  eine beschränkte und gerade reellwertige Funktion, dann ist  $\Phi_w(\mathbf{t})$  reellwertig und gerade.
- Es bezeichne  $\Phi_w^*(\mathbf{t})$  die konjugiert komplexe Funktion zu  $\Phi_w(\mathbf{t})$ . Wenn  $\Phi_w^*(\mathbf{t}) = \Phi_w(-\mathbf{t})$  ist, dann ist  $w(\mathbf{y})$  reellwertig.

**Bemerkung 3.2** Es sei  $K(\mathbf{y})$  eine Kernfunktion,  $G(\mathbf{t}, \Theta_\varphi)$  eine Gewichtsfunktion und  $M(\mathbf{t}) = E[e^{it' r \Theta_\varphi} G(\mathbf{t}, \Theta_\varphi)]$ . Es gilt

i.  $\Phi_K(\mathbf{t})$  ist gerade und reellwertig;

ii.  $G(\mathbf{t}, \Theta_\varphi)$  ist gerade und reellwertig;

iii. Unter den Modellannahmen (M1) ist  $M(\mathbf{t})$  gerade und reellwertig.

**Beweis:** Mit Bemerkung 3.1 folgt i. aus Definition 2.3, ii. folgt aus Definition 2.5 und iii. folgt aus der Definition von  $M(\mathbf{t})$  unter Verwendung von (M1).

**q.e.d.**

**Korollar 3.1** Der modifizierte Dekonvolutions-Kernschätzer  $\hat{f}_n(\mathbf{y})$  ist reellwertig.

**Beweis:** Gemäß Definition 2.6 läßt sich  $\hat{f}_n(\mathbf{y})$  ausdrücken über eine Summe inverser Fouriertransformierter von Funktionen  $\hat{\Phi}_{f_j}(\mathbf{t})$ :

$$\hat{f}_n(\mathbf{y}) = \frac{1}{n} \sum_{j=1}^n \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \hat{\Phi}_{f_j}(\mathbf{t}) d\mathbf{t},$$

mit

$$\hat{\Phi}_{f_j}(\mathbf{t}) = \cos(\mathbf{tX}_j) \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) + i \sin(\mathbf{tX}_j) \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}).$$

Für die konjugiert komplexe Funktion gilt

$$\hat{\Phi}_{f_j}^*(\mathbf{t}) = \cos(\mathbf{tX}_j) \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) - i \sin(\mathbf{tX}_j) \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}).$$

Da  $\frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t})$  nach Bemerkung 3.2 eine gerade, reellwertige Funktion ist, folgt

$$\hat{\Phi}_{f_j}^*(\mathbf{t}) = \hat{\Phi}_{f_j}(-\mathbf{t}) \quad j = 1, \dots, n.$$

Die Aussage folgt dann direkt aus Bemerkung 3.1.

**q.e.d.**

Wie bereits in Kapitel 2 erwähnt, läßt das hier betrachtete Model des Meßfehlers Emissionsbeobachtungen außerhalb von  $\mathcal{B}$  zu. Die Emissionsdichte  $f(\mathbf{y})$  sei wie folgt auf  $\mathbb{R}^2$  fortgesetzt,

$$f(\mathbf{y}) := \begin{cases} f(\mathbf{y}) & \text{für } \mathbf{y} \in \mathcal{B}, \\ 0 & \text{sonst.} \end{cases}$$

Die Schätzfunktion  $\hat{f}_n(\mathbf{y})$  ist eine Zufallsvariable. Zunächst werden Bias und Varianz von  $\hat{f}_n(\mathbf{y})$  untersucht. Der folgende Satz zeigt, daß der modifizierte Dekonvolutions-Kernschätzer asymptotisch unverfälscht ist.

**Satz 3.1** *Für den modifizierten Dekonvolutions-Kernschätzer  $\hat{f}_n(\mathbf{y})$  gilt für alle  $\mathbf{y} \in \mathcal{B}$ :*

$$\lim_{n \rightarrow \infty} \text{Bias}(\hat{f}_n(\mathbf{y})) = 0.$$

**Beweis:** Für den Bias des modifizierten Dekonvolutions-Kernschätzers an einer Stelle  $\mathbf{y} \in \mathcal{B}$  gilt:

$$\text{Bias}(\hat{f}_n(\mathbf{y})) := E[\hat{f}_n(\mathbf{y})] - f(\mathbf{y}).$$

Es wird nun gezeigt, daß

$$\text{Bias}(\hat{f}_n(\mathbf{y})) = \frac{1}{\lambda_n^2} \int_{\mathcal{B}} K\left(\frac{\mathbf{y}-\mathbf{y}_b}{\lambda_n}\right) f(\mathbf{y}_b) d\mathbf{y}_b - f(\mathbf{y}).$$

Es ist

$$E[\hat{f}_n(\mathbf{y})] = E \left[ \frac{1}{n} \sum_{j=1}^n \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{x}_j)} \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt \right].$$

Da  $(\mathbf{X}_1, \varphi_1), \dots, (\mathbf{X}_n, \varphi_n)$  u.i.v. Zufallsvariablen sind, läßt sich dies schreiben als

$$E[\hat{f}_n(\mathbf{y})] = E \left[ \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{X}_1)} \frac{G(\mathbf{t}, \Theta_{\varphi_1})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt \right].$$

Über die Eigenschaften des bedingten Erwartungswertes erhält man

$$\begin{aligned} E[\hat{f}_n(\mathbf{y})] &= \int_{\mathcal{B}} E \left[ \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{X}_1)} \frac{G(\mathbf{t}, \Theta_{\varphi_1})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt \mid \mathbf{y}_b \right] f(\mathbf{y}_b) d\mathbf{y}_b \\ &= \int_{\mathcal{B}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{x})} \frac{G(\mathbf{t}, \Theta_{\varphi})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt w(\mathbf{x}, \varphi | \mathbf{y}_b) d\mathbf{x} d\varphi f(\mathbf{y}_b) d\mathbf{y}_b. \end{aligned}$$

Wie zuvor im Beweis zu Satz 2.1 kann bei gegebenem  $\varphi$  und  $\mathbf{y}_b$ ,  $\mathbf{x}$  nur Werte auf der Geraden  $L_x := \{\mathbf{x} \in \mathbb{R}^2, \mathbf{x} = \mathbf{y}_b + \mathbf{z}, \mathbf{z} = (z_1, z_2), \frac{z_2}{z_1} = \tan(\varphi)\}$  annehmen.

$$E[\hat{f}_n(\mathbf{y})] =$$

$$\int_{\mathcal{B}} \int_{L_x} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{x})} \frac{G(\mathbf{t}, \Theta_{\varphi})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt w(\mathbf{x}, \varphi | \mathbf{y}_b) d\mathbf{x} d\varphi f(\mathbf{y}_b) d\mathbf{y}_b.$$

Es wird die Koordinatentransformation  $\mathbf{x} = \mathbf{y}_b + \mathbf{z}$  für festes  $\mathbf{y}_b$  durchgeführt. Transformation der Wahrscheinlichkeitsdichten  $w(\mathbf{x}, \varphi | \mathbf{y}_b) = w(\mathbf{z}, \varphi | \mathbf{y}_b)$  ergibt die nächste Gleichung. Als neuer Integrationsbereich ergibt sich  $L_z := \{(z_1, z_2) \in \mathbb{R}^2, \frac{z_2}{z_1} = \tan(\varphi)\}$ . Es werden weiterhin die Beziehungen  $w(\mathbf{z} | \varphi, \mathbf{y}_b) = w(\mathbf{z} | \varphi)$  und  $w(\varphi | \mathbf{y}_b) = \frac{1}{\pi}$  genutzt.

$$\begin{aligned} E[\hat{f}_n(\mathbf{y})] &= \int_{\mathcal{B}} \int_{L_z} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{y}_b)} e^{it_1 z_1} e^{it_2 z_2} \frac{G(\mathbf{t}, \Theta_{\varphi})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt \\ &\quad w(\mathbf{z} | \varphi) \frac{1}{\pi} dz d\varphi f(\mathbf{y}_b) d\mathbf{y}_b, \\ &= \int_{\mathcal{B}} \int_{\mathbb{R}} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{y}_b)} e^{it_1 z_1} e^{it_2 z_1 \tan(\varphi)} \frac{G(\mathbf{t}, \Theta_{\varphi})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt \\ &\quad w(z_1 | \varphi) \frac{1}{\pi} dz_1 d\varphi f(\mathbf{y}_b) d\mathbf{y}_b. \end{aligned}$$

Koordinatentransformation  $z_1 = r \cos(\varphi)$  mit  $dz_1 = \cos(\varphi) dr$  sowie die Transformation von Wahrscheinlichkeitsdichten  $w(z_1|\varphi) = \frac{1}{\cos(\varphi)}w(r|\varphi)$  ergibt

$$E[\hat{f}_n(\mathbf{y})] = \int_{\mathcal{B}} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{y}_b)} e^{it_1 r \cos(\varphi)} e^{it_2 r \sin(\varphi)} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) d\mathbf{t} \\ w(r|\varphi) \frac{1}{\pi} dr d\varphi f(\mathbf{y}_b) d\mathbf{y}_b.$$

Die Größen  $r$  und  $\varphi$  sind unabhängige Zufallsvariablen. Also ist  $w(r|\varphi) = w(r)$ . Nach Definition 2.3 der Kernfunktion ist  $K(y_1, y_2) = K_1(y_1) \cdot K_2(y_2)$ . Daher ist  $\Phi_K(\lambda_n \mathbf{t}) = \Phi_{K_1}(\lambda_n t_1) \cdot \Phi_{K_2}(\lambda_n t_2)$ .

$$E[\hat{f}_n(\mathbf{y})] = \frac{1}{(2\pi)^2} \int_{\mathcal{B}} \int_{\mathbb{R}} e^{-it'(\mathbf{y}-\mathbf{y}_b)} \Phi_{K_1}(\lambda_n t_1) \Phi_{K_2}(\lambda_n t_2) \int_{\mathbb{R}^2} e^{it' r \Theta_\varphi} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} \\ w(r) \frac{1}{\pi} dr d\varphi d\mathbf{t} f(\mathbf{y}_b) d\mathbf{y}_b.$$

Aus der Definition von  $M(\mathbf{t})$  folgt, daß  $\int_{\mathbb{R}^2} \int_{\mathbb{R}} e^{it' r \Theta_\varphi} \frac{G(\mathbf{t}, \Theta_\varphi)}{M(\mathbf{t})} w(r) \frac{1}{\pi} dr d\varphi = 1$ . Damit gilt:

$$E[\hat{f}_n(\mathbf{y})] = \frac{1}{(2\pi)^2} \int_{\mathcal{B}} \int_{\mathbb{R}} e^{-it_1(y_1-y_{b1})} \Phi_{K_1}(\lambda_n t_1) dt_1 \int_{\mathbb{R}} e^{-it_2(y_2-y_{b2})} \Phi_{K_2}(\lambda_n t_2) dt_2 f(\mathbf{y}_b) d\mathbf{y}_b.$$

Aus den Eigenschaften der Fouriertransformation folgt

$$E[\hat{f}_n(\mathbf{y})] = \int_{\mathcal{B}} \frac{1}{\lambda_n^2} K_1\left(\frac{y_1-y_{b1}}{\lambda_n}\right) K_2\left(\frac{y_2-y_{b2}}{\lambda_n}\right) f(\mathbf{y}_b) d\mathbf{y}_b \\ = \frac{1}{\lambda_n^2} \int_{\mathcal{B}} K\left(\frac{\mathbf{y}-\mathbf{y}_b}{\lambda_n}\right) f(\mathbf{y}_b) d\mathbf{y}_b.$$

Daraus folgt

$$\text{Bias}(\hat{f}_n(\mathbf{y})) = \frac{1}{\lambda_n^2} \int_{\mathcal{B}} K\left(\frac{\mathbf{y}-\mathbf{y}_b}{\lambda_n}\right) f(\mathbf{y}_b) d\mathbf{y}_b - f(\mathbf{y}). \quad (3.1)$$

Die Emissionsdichtefunktion  $f(\mathbf{y})$  wurde außerhalb von  $\mathcal{B}$  durch  $f(\mathbf{y}) = 0$  fortgesetzt. Daher kann das obige Integral über ganz  $\mathbb{R}^2$  betrachtet werden.

$$\text{Bias}(\hat{f}_n(\mathbf{y})) = \frac{1}{\lambda_n^2} \int_{\mathbb{R}^2} K\left(\frac{\mathbf{y}-\mathbf{y}_b}{\lambda_n}\right) f(\mathbf{y}_b) d\mathbf{y}_b - f(\mathbf{y}).$$

Die Substitution  $\mathbf{s} = \frac{\mathbf{y}-\mathbf{y}_b}{\lambda_n}$  wird durchgeführt.

$$\text{Bias}(\hat{f}_n(\mathbf{y})) = \int_{\mathbb{R}^2} K(\mathbf{s}) f(\mathbf{y} - \lambda_n \mathbf{s}) d\mathbf{s} - f(\mathbf{y}).$$

Aus Definition 2.6 folgt, daß der Bandweiteparameter  $\lambda_n \rightarrow 0$  für  $n \rightarrow \infty$ .

$$\lim_{n \rightarrow \infty} \left( \int_{\mathbb{R}^2} K(\mathbf{s}) f(\mathbf{y} - \lambda_n \mathbf{s}) d\mathbf{s} \right) - f(\mathbf{y}) = f(\mathbf{y}) \int_{\mathbb{R}^2} K(\mathbf{s}) d\mathbf{s} - f(\mathbf{y}).$$

Nach Definition 2.3 ist  $\int_{\mathbb{R}^2} K(\mathbf{s}) d\mathbf{s} = 1$ . Daher ist

$$\lim_{n \rightarrow \infty} \text{Bias}(\hat{f}_n(\mathbf{y})) = 0.$$

**q.e.d.**

Gleichung (3.1) zeigt das Bias-Verhalten des modifizierten Dekonvolutions-Kernschätzers  $\hat{f}_n(\mathbf{y})$ . Dies wird mit dem Bias eines Kernschätzers, der auf direkten Beobachtungen basiert verglichen. Es seien  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  u. i. v. direkte Beobachtungspunkte in  $\mathcal{B}$ . Der Kernschätzer an der Stelle  $\mathbf{y}_0 \in \mathcal{B}$ , der aus diesen direkten Beobachtungen gebildet werden kann, ist nach [46], [45]

$$\hat{f}_n^{dir}(\mathbf{y}_0) = \frac{1}{n\lambda_n^2} \sum_{j=1}^n K\left(\frac{1}{\lambda_n}(\mathbf{y}_0 - \mathbf{Y}_j)\right).$$

Hierbei ist  $K(\mathbf{y})$  eine Kernfunktion, welche den Bedingungen in Definition 2.3 genügt. Für den Erwartungswert, dieses Kernschätzers gilt

$$E[\hat{f}_n^{dir}(\mathbf{y}_0)] = \frac{1}{n\lambda_n^2} \sum_{j=1}^n E\left[K\left(\frac{1}{\lambda_n}(\mathbf{y}_0 - \mathbf{Y}_j)\right)\right].$$



Da  $\mathbf{Y}_j, j = 1, \dots, n$  u. i. v. sind, ist dies gleich

$$\frac{1}{\lambda_n^2} E \left[ K\left(\frac{1}{\lambda_n}(\mathbf{y}_0 - \mathbf{Y}_1)\right) \right] = \frac{1}{\lambda_n^2} \int_{\mathcal{B}} K\left(\frac{1}{\lambda_n}(\mathbf{y}_0 - \mathbf{y})\right) f(\mathbf{y}) d\mathbf{y}.$$

Daher ist der Bias des Kernschätzers aus direkten Beobachtungen

$$Bias(\hat{f}_n^{dir}(\mathbf{y}_0)) = \frac{1}{\lambda_n^2} \int_{\mathcal{B}} K\left(\frac{1}{\lambda_n}(\mathbf{y}_0 - \mathbf{y})\right) f(\mathbf{y}) d\mathbf{y} - f(\mathbf{y}_0).$$

Dies ist gleich dem Ausdruck in (3.1), woraus folgt, daß der Bias des Dekonvolutions-Kernschätzers mit gewichteten Beobachtungen sich genau so verhält wie der Bias des Kernschätzers aus direkten Beobachtungsdaten. Das Fehlen direkter Beobachtungen hat bei Verwendung des modifizierten Dekonvolutions-Kernschätzers keinen Einfluß auf das Verhalten des Bias der Schätzfunktion. Als nächstes ermitteln wir eine obere Schranke für die Varianz von  $\hat{f}_n(\mathbf{y})$ . Die Varianz von  $\hat{f}_n(\mathbf{y})$  ist gegeben durch

$$var(\hat{f}_n(\mathbf{y})) := E \left[ (\hat{f}_n(\mathbf{y}) - E[\hat{f}_n(\mathbf{y})])^2 \right].$$

**Satz 3.2** Für die Varianz des modifizierten Dekonvolutions-Kernschätzers an einer beliebigen Stelle  $\mathbf{y} \in \mathcal{B}$  ergibt sich die Abschätzung

$$var(\hat{f}_n(\mathbf{y})) \leq C_{bed} \frac{1}{n\lambda_n^2} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \Phi_K(\mathbf{t})^2 \frac{E[G(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi)^2]}{|M(\frac{\mathbf{t}}{\lambda_n})|^2} d\mathbf{t} + (C/n) E^2[\hat{f}_n(\mathbf{y})].$$

**Beweis:** Hier und an anderen Stellen in dieser Arbeit bezeichnet  $C$  eine generische Konstante, nicht in jedem Fall die gleiche.

Die Varianz des modifizierten Dekonvolutions-Kernschätzers ist gegeben durch

$$var(\hat{f}_n(\mathbf{y})) = var \left( \frac{1}{n} \frac{1}{(2\pi)^2} \sum_{j=0}^n \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{X}_j)} \frac{G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) d\mathbf{t} \right).$$

Da die Beobachtungen  $(\mathbf{X}_1, \varphi_1), \dots, (\mathbf{X}_n, \varphi_n)$  u.i.v. sind, ist

$$\begin{aligned}
\text{var}(\hat{f}_n(\mathbf{y})) &= \frac{1}{n} \frac{1}{(2\pi)^4} \text{var}\left(\int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{X}_1)} \frac{G(\mathbf{t}, \Theta_{\varphi_1})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt\right) \\
&= \frac{1}{n} \frac{1}{(2\pi)^4} E \left[ \left( \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{X}_1)} \frac{G(\mathbf{t}, \Theta_{\varphi_1})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt \right)^2 \right] \\
&\quad + (C/n) E^2[\hat{f}_n(\mathbf{y})].
\end{aligned}$$

Die Dichtefunktion  $w(\mathbf{x}, \varphi) := w(\mathbf{x}|\varphi) \frac{1}{\pi}$  sei beschränkt. Es gilt daher für alle Paare  $(\mathbf{x}, \varphi)$ , daß die bedingte Wahrscheinlichkeitsdichte  $w(\mathbf{x}|\varphi) \leq C_{bed}$ .

$$\begin{aligned}
\text{var}(\hat{f}_n(\mathbf{y})) &= \frac{1}{n} \frac{1}{(2\pi)^4} E \left[ \int_{\mathbb{R}^2} \left( \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{x})} \frac{G(\mathbf{t}, \Theta_{\varphi})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt \right)^2 w(\mathbf{x}|\varphi) d\mathbf{x} \right] \\
&\quad + (C/n) E^2[\hat{f}_n(\mathbf{y})]. \\
&\leq \frac{1}{n} \frac{1}{(2\pi)^4} C_{bed} E \left[ \int_{\mathbb{R}^2} \left( \int_{\mathbb{R}^2} e^{-it'(\mathbf{y}-\mathbf{x})} \frac{G(\mathbf{t}, \Theta_{\varphi})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) dt \right)^2 d\mathbf{x} \right] \\
&\quad + (C/n) E^2[\hat{f}_n(\mathbf{y})].
\end{aligned}$$

Es sei  $\tilde{f}(\mathbf{t}) := e^{-it' \mathbf{y}} \frac{G(\mathbf{t}, \Theta_{\varphi})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t})$ .

$$\text{var}(\hat{f}_n(\mathbf{y})) \leq \frac{1}{n} \frac{1}{(2\pi)^2} C_{bed} E \left[ \int_{\mathbb{R}^2} \frac{1}{(2\pi)^2} \left| \int_{\mathbb{R}^2} e^{it' \mathbf{x}} \tilde{f}(\mathbf{t}) dt \right|^2 d\mathbf{x} \right] + (C/n) E^2[\hat{f}_n(\mathbf{y})].$$

Hierbei ist  $\int_{\mathbb{R}^2} e^{it' \mathbf{x}} \tilde{f}(\mathbf{t}) dt$  die reellwertige Fouriertransformation der Funktion  $\tilde{f}(\mathbf{t})$  an der Stelle  $\mathbf{x}$ . Mit Parsevals's Gleichung ist

$$\int_{\mathbb{R}^2} \left| \int_{\mathbb{R}^2} e^{it' \mathbf{x}} \tilde{f}(\mathbf{t}) dt \right|^2 d\mathbf{x} = (2\pi)^2 \int_{\mathbb{R}^2} |\tilde{f}(\mathbf{t})|^2 dt.$$

Daher ist

$$\begin{aligned}
\text{var}(\hat{f}_n(\mathbf{y})) &\leq \frac{1}{n} \frac{1}{(2\pi)^2} C_{bed} E \left[ \int_{\mathbb{R}^2} |\tilde{f}(\mathbf{t})|^2 dt \right] + (C/n) E^2[\hat{f}_n(\mathbf{y})]. \\
&= \frac{1}{n} \frac{1}{(2\pi)^2} C_{bed} E \left[ \int_{\mathbb{R}^2} |e^{-it' \mathbf{y}}|^2 \frac{|G(\mathbf{t}, \Theta_\varphi)|^2}{|M(\mathbf{t})|^2} |\Phi_K(\lambda_n \mathbf{t})|^2 dt \right] \\
&\quad + (C/n) E^2[\hat{f}_n(\mathbf{y})] \\
&= \frac{1}{n} \frac{1}{(2\pi)^2} C_{bed} \left[ \int_{\mathbb{R}^2} \frac{E[G(\mathbf{t}, \Theta_\varphi)^2]}{|M(\mathbf{t})|^2} |\Phi_K(\lambda_n \mathbf{t})|^2 dt \right] + (C/n) E^2[\hat{f}_n(\mathbf{y})].
\end{aligned}$$

Substitution  $\mathbf{s} = \lambda_n \mathbf{t}$  ergibt:

$$\text{var}(\hat{f}_n(\mathbf{y})) \leq \frac{1}{n\lambda_n^2} \frac{1}{(2\pi)^2} C_{bed} \int_{\mathbb{R}^2} \Phi_K(\mathbf{s})^2 \frac{E[G(\frac{\mathbf{s}}{\lambda_n}, \Theta_\varphi)^2]}{|M(\frac{\mathbf{s}}{\lambda_n})|^2} d\mathbf{s} + (C/n) E^2[\hat{f}_n(\mathbf{y})].$$

**q.e.d.**

Die letzte Ungleichung gilt für alle  $\mathbf{y} \in \mathcal{B}$ . Gleichung (3.1), sowie Satz 3.2 zeigen, daß das Verhalten von Bias und Varianz des modifizierten Dekonvolutions-Kernschätzers von dem Bandweiteparameter  $\lambda_n$  beeinflusst wird. Der Bias wird zudem von der Kernfunktion  $K(\mathbf{y})$  und der Dichtefunktion  $f(\mathbf{y})$  bestimmt. Die Varianz wird zusätzlich beeinflusst von der Kernfunktion, der Gewichtsfunktion  $G(\mathbf{t}, \Theta_\varphi)$  sowie der Verteilung des Beobachtungsfehlers  $r\Theta_\varphi$ , welche in die Funktion  $M(\mathbf{t})$  einfließt.

## 3.2 Fehlerverhalten der Schätzfunktion

Die Rekonstruktion von  $f(\mathbf{y})$  durch  $\hat{f}_n(\mathbf{y})$  ist fehlerbehaftet. Dieser Rekonstruktionsfehler kann mittels verschiedener Fehlermaße beschrieben werden. Definition 3.1 führt die in dieser Arbeit verwendeten Fehlermaße ein.

**Definition 3.1** *Die folgenden Fehlermaße werden zur Beurteilung des modifizierten Dekonvolutions-Kernschätzers  $\hat{f}_n(\mathbf{y})$  eingeführt.*

1. Der Mittlere Quadratische Fehler (MSE) von  $\hat{f}_n$  an der Stelle  $\mathbf{y}$ :

$$MSE(\hat{f}_n(\mathbf{y})) := E \left[ (\hat{f}_n(\mathbf{y}) - f(\mathbf{y}))^2 \right].$$

2. Der Mittlere Integrierte Quadratische Fehler (MISE) von  $\hat{f}_n$ :

$$MISE(\hat{f}_n) := \int_{\mathcal{B}} MSE(\hat{f}_n(\mathbf{y})) \, d\mathbf{y}.$$

3. Die Supremum-Norm von  $\hat{f}_n - f$ :

$$\|\hat{f}_n(\mathbf{y}) - f(\mathbf{y})\|_{\infty} := \sup_{\mathcal{B}} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})|.$$

Im weiteren Verlauf der Arbeit wird das Verhalten des modifizierten Dekonvolutions-Kernschätzers unter den Fehlermaßen aus Definition 3.1 für große  $n$  untersucht. Zunächst wird die Funktionenklasse definiert und motiviert, aus welcher  $f$  gewählt wird.

### 3.2.1 Wahl der Funktionenklasse

#### Die Sobolev-Räume

In der Literatur, die sich mit der Rekonstruktion von Intensitätsfunktionen aus Tomographie-Daten beschäftigt, werden häufig Sobolev-Räume von Funktionen verwendet, um die Eigenschaften der Intensitätsfunktionen zu beschreiben, so zum Beispiel in [37], [36], [31]. Dies basiert auf der Verwendung der Radontransformation als Operator zwischen Emissionsdichte und Detektionsdichte. Es sei  $f \in \mathcal{L}^2(\mathbb{R}^2)$ , es sei  $\Theta$  ein Einheitsvektor aus  $\mathcal{D}$  (siehe Definition 2.1).  $\Theta_s^{\perp}$ ,  $s \in \mathbb{R}_0^+$ , bezeichnet die Gerade in  $\mathcal{B}$  senkrecht zu  $\Theta$  im Abstand  $s$  zum Ursprung. Der Schnittpunkt dieser Geraden mit  $\mathcal{D}$  definiert eindeutig die Beobachtungsorte in  $\mathcal{D}$ . Falls keine Flugzeitinformation vorliegt ist die Detektionsdichte auf  $\mathcal{D}$  definiert. Sie hängt mit der Emissionsdichte  $f$  zusammen wie folgt.

**Definition 3.2** *Das Integral*

$$(Rf)[\Theta, s] := \int_{\Theta_s^{\perp}} f(s\Theta + \mathbf{y}) \, d\mathbf{y}$$

heißt Radon-Transformation der Funktion  $f$  an der Stelle  $(\Theta, s)$ .

Es bezeichne  $d(\Theta, s)$  die Detektionsdichte in  $\mathcal{D}$ . Dann ist

$$d(\Theta, s) = (Rf)[\Theta, s].$$

$f(\mathbf{y})$  kann aus der inversen Radontransformation von  $d(\Theta, s)$  bestimmt werden. In [37] wird erläutert, daß die inverse Radontransformation als Operator zwischen

$\mathcal{L}^2$ -Funktionsräumen nicht stetig ist, jedoch stetig ist als Operator zwischen geeigneten Sobolev-Räumen. Daher werden zum Beispiel in [37], [36] und [31] die Sobolev-Räume als Funktionsräume für Emissionsdichte und Detektionsdichte vorgeschlagen. Definition 3.3 bis Definition 3.5 führen den Begriff des Sobolev-Raumes ein.

**Definition 3.3 (Schwache Ableitung)**

Es sei  $f \in \mathcal{L}^2(\mathbb{R}^2)$  und  $u \in \mathcal{C}_0^\infty(\mathcal{B})$ . Es bezeichne  $\beta$  einen Mehrfachindex,  $\beta = (\beta_1, \beta_2)$ ,  $\beta_1, \beta_2 \in \mathbb{N}_0$  mit  $|\beta| := \beta_1 + \beta_2$ . Der Operator  $D^\beta f$  heißt schwache Ableitung von  $f$  zur Ordnung  $|\beta|$ , wenn

$$\int_{\mathbb{R}^2} (D^\beta f)(\mathbf{y}) u(\mathbf{y}) d\mathbf{y} = (-1)^{|\beta|} \int_{\mathbb{R}^2} f(\mathbf{y}) (\mathbf{D}^\beta u)(\mathbf{y}) d\mathbf{y}, \quad \forall u \in \mathcal{C}_0^\infty.$$

Hierbei bezeichnet  $\mathbf{D}^\beta u$  die klassische Ableitung von  $u$  zur Ordnung  $|\beta|$ .

Die schwache Ableitung  $(D^\beta f) \in \mathcal{L}^2(\mathbb{R}^2)$  ist bis auf Nullmengen eindeutig bestimmt.

**Definition 3.4 (Sobolev-Raum ganzzahliger Ordnung)**

Für ein  $k \in \mathbb{N}_0$  heißt die Menge

$$\mathcal{W}^{k,2}(\mathbb{R}^2) := \{f \in \mathcal{L}^2(\mathbb{R}^2), D^\beta f \in \mathcal{L}^2(\mathbb{R}^2), |\beta| \leq k\}$$

Sobolev-Raum der Ordnung  $k$ . Die Norm

$$\|f\|_{\mathcal{W}^{k,2}} := \left( \int_{\mathbb{R}^2} \sum_{|\beta| \leq k} |D^\beta f(\mathbf{y})|^2 d\mathbf{y} \right)^{\frac{1}{2}}.$$

heißt Sobolev-Norm.

Die Räume  $\mathcal{W}^{k,2}(\mathbb{R}^2)$  werden zu Sobolev Räumen reeller Ordnung erweitert.

**Definition 3.5 (Sobolev-Raum reellwertiger Ordnung)**

Es sei  $f \in \mathcal{L}^2(\mathbb{R}^2)$ ,  $\alpha \in \mathbb{R}_0^+$ . Der Sobolev-Raum  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  ist die Menge

$$\mathcal{H}^{\alpha,2}(\mathbb{R}^2) = \{f \in \mathcal{L}^2(\mathbb{R}^2), \|f\|_{\mathcal{H}^{\alpha,2}(\mathbb{R}^2)} < \infty\},$$

mit der Norm

$$\|f\|_{\mathcal{H}^{\alpha,2}(\mathbb{R}^2)} := \int_{\mathbb{R}^2} (1 + \|\mathbf{t}\|^2)^\alpha |\Phi_f(\mathbf{t})|^2 d\mathbf{t}.$$

Für  $f \in \mathcal{L}^2(\mathbb{R}^2)$  ist die Fouriertransformierte definiert als

$$\Phi_f(\mathbf{t}) = \lim_{a \rightarrow \infty} \int_{\|\mathbf{y}\| \leq a} e^{it' \mathbf{y}} f(\mathbf{y}) d\mathbf{y}.$$

Da die Emissionsdichte  $f$  außerhalb von  $\mathcal{B}$  durch  $f(\mathbf{y}) = 0$  fortgesetzt wurde, entspricht in diesem Fall die Fouriertransformierte von  $f$  in  $\mathcal{L}^2(\mathbb{R}^2)$  der in Definition 2.2 eingeführten Fouriertransformierten von  $f$  in  $\mathcal{L}^1(\mathcal{B})$ . Die Räume  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  werden auch als *Bessel-Potential Räume* bezeichnet, um sie von den Räumen  $\mathcal{W}^{k,2}(\mathbb{R}^2)$  zu unterscheiden. Für geradzahliges  $\alpha$  stimmen die Räume  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  mit den Räumen  $\mathcal{W}^{\alpha,2}(\mathbb{R}^2)$  überein. In dieser Arbeit sind immer die Räume  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  gemeint, wenn der Begriff des Sobolev-Raumes verwendet wird.

**Anmerkung 3.1** *Emissionsdichte  $f$  und Detektionsdichte  $g$  seien Elemente aus einem Sobolev-Raum  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  der Ordnung  $\alpha \in \mathbb{R}_0^+$ .*

Die Einführung eines Sobolev-Raumes als Funktionenraum für Emissionsdichte und Detektionsdichte impliziert Glattheitseigenschaften von  $f$  und  $g$ . Je größer  $\alpha$ , desto höher die Glattheitsanforderungen. Wir begründen nun die Wahl der Sobolev-Räume als Funktionenräume und die Wahl eines geeigneten Sobolev-Parameters.

### Wahl eines Sobolev-Raumes geeigneter Ordnung

ToFPET-Rekonstruktionen weisen hohe Kontraste auf. Es zeichnen sich Gebiete hoher Emissionsaktivität dunkel vor einem helleren Hintergrund ab. Die Gebietsgrenzen sind glatte Kurven, zum Beispiel Ellipsen. Die Gebietsränder sind wegen der geringen Auflösung des Verfahrens verrauscht. Bei einer Emissionsdichte können daher große Steigungen auftreten, jedoch - wegen der geringen Bildauflösung - keine Unstetigkeiten. In Zhang (1992) [61] wird die Rekonstruktion einer Emissionsdichtefunktion aus ToFPET-Daten über dem Raum der Hölder-Stetigen Funktionen  $\mathcal{C}^{k,\mu}(\mathbb{R}^2)$  mit  $k > 0$  durchgeführt. Diese Räume sind wie folgt definiert.

**Definition 3.6** ( $\mathcal{C}^{0,\mu}(\mathbb{R}^2)$ ,  $\mathcal{C}^{k,\mu}(\mathbb{R}^2)$ )

- $\mathcal{C}^{0,\mu}(\mathbb{R}^2) := \{f, \text{ wobei für } \mathbf{y}, \tilde{\mathbf{y}} \in \mathbb{R}^2, |f(\mathbf{y}) - f(\tilde{\mathbf{y}})| \leq C \|\mathbf{y} - \tilde{\mathbf{y}}\|^\mu\}$ ,  
 $0 < \mu \leq 1$ , heißt Raum der Hölder-Stetigen Funktionen.
- Es bezeichne  $\beta = (\beta_1, \beta_2)$  einen Mehrfachindex. Es sei  $k \in \mathbb{N}_0$ . Dann ist  $\mathcal{C}^{k,\mu}(\mathbb{R}^2) := \{f \in \mathcal{C}^{0,\mu}(\mathbb{R}^2) \text{ mit } \mathbf{D}^\beta f \in \mathcal{C}^{0,\mu}(\mathbb{R}^2), 0 \leq |\beta| \leq k\}$ .

Die Räume  $\mathcal{C}^{k,\mu}(\mathbb{R}^2)$  stellen eine Verallgemeinerung der Räume  $\mathcal{C}^k(\mathbb{R}^2)$  der  $k$ -fach stetig differenzierbaren Funktionen dar. In Zhang (1992) [61] wird damit von der Emissionsdichte  $f$  die Existenz einer klassischen Ableitung gefordert. In dieser Arbeit wollen wir die Klasse der zulässigen Emissionsdichten ausdehnen auf Dichtefunktionen, welche nicht unbedingt eine klassische Ableitung besitzen, aber schwache Ableitungen. Dies führt zunächst zu den Räumen  $\mathcal{W}^{k,2}(\mathbb{R}^2)$ ,  $k \in \mathbb{N}_0$  aus Definition 3.4, welche als Verallgemeinerung der Räume  $\mathcal{C}^{k,\mu}$  interpretiert werden können. Einführung des Bessel-Potentiales ermöglicht es, die Sobolev-Räume ganzzahliger Ordnung zu erweitern auf die Sobolev-Räume reellwertiger Ordnung,  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$ ,  $\alpha \in \mathbb{R}_0^+$  aus Definition 3.5.

Folgendes Theorem nach Sobolev [2] zeigt Einbettungs-Eigenschaften des Sobolev-Raumes  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  in die Räume  $\mathcal{W}^{k,2}(\mathbb{R}^2)$ ,  $\mathcal{C}^{k,\mu}(\mathbb{R}^2)$  und den Raum  $\mathcal{C}^0(\mathbb{R}^2)$ .

**Theorem 3.1 (nach S.L. Sobolev)**

- $\mathcal{H}^{\alpha,2}(\mathbb{R}^2) \subset \mathcal{H}^{\alpha_2,2}(\mathbb{R}^2)$ , wenn  $\alpha > \alpha_2$ .
- Sei  $k \in \mathbb{N}_0$  und  $\epsilon > 0$  aus  $\mathbb{R}$ . Dann ist

$$\mathcal{H}^{k+\epsilon,2}(\mathbb{R}^2) \subset \mathcal{W}^{k,2}(\mathbb{R}^2) \subset \mathcal{H}^{k-\epsilon,2}(\mathbb{R}^2).$$

- $\mathcal{H}^{\alpha,2} \subset \mathcal{C}^0(\mathbb{R}^2)$ , wenn  $\alpha > 1$ .
- $\mathcal{H}^{\alpha,2} \subset \mathcal{C}^{k,\mu}(\mathbb{R}^2)$ , wenn  $\alpha > 1 + \mu + k$ .

Aus Theorem 3.1 ergibt sich, daß Existenz einer Hölder-stetigen Ableitung für Funktionen aus dem Sobolev-Raum  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  für  $\alpha > 2 + \mu$  garantiert ist. Wir wählen die Klasse der Emissionsdichten und Detektionsdichten nun wie folgt.

**Anmerkung 3.2** Die Emissionsdichte  $f$  und die Detektionsdichte  $g$  seien Elemente eines Sobolev-Raumes  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  der Ordnung  $\alpha > 1$ .

Wie aus Theorem 3.1 hervorgeht, garantiert diese Wahl des Funktionenraumes Stetigkeit in den Dichtefunktionen sowie die Existenz schwacher Ableitungen mindestens zur Ordnung 1, fordert aber nicht notwendigerweise die Existenz einer klassischen Ableitung. Die Wahl des Raumes  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  mit  $\alpha > 1$  als Funktionenraum ist im Einklang mit den Anforderungen an ToFPET Emissionsdichten und ist weniger restriktiv als die Wahl des Raumes  $\mathcal{C}^{k,\mu}(\mathbb{R}^2)$  mit  $k > 0$  aus [61].

In der Literatur findet man für Tomographie–Verfahren mit höherer Bildauflösung die Wahl  $f \in \mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  mit  $\alpha < \frac{1}{2}$ , so zum Beispiel in [37]. Hier sind also Unstetigkeiten in der Intensitätsfunktion zugelassen.

Ähnlich wie bei den Räumen  $k$ -fach stetig differenzierbarer Funktionen,  $\mathcal{C}^k(\mathbb{R}^2)$ , besteht bei Sobolev–Räumen ganzzahliger Ordnung ein Zusammenhang zwischen der Existenz schwacher Ableitungen  $D^\beta$ ,  $|\beta| \leq k$  und der Existenz eines Integrals mit quadrierter Norm der Fouriertransformation  $|\Phi_f(\mathbf{t})|^2$  und polynomialem Faktor  $(1 + \|\mathbf{t}\|^2)^k$ . Näheres über die Theorie von Sobolev–Räumen ganzzahliger und reeller Ordnung findet man zum Beispiel in [1], [2], [56], [60].

Nachdem nun die Klasse der Emissionsdichten gewählt ist, soll das Verhalten des modifizierten Dekonvolutions–Kernschätzers unter den in Definition 3.1 eingeführten Fehlermaßen untersucht werden.

### 3.2.2 Die Fehlermaße Mittlerer Quadratischer Fehler und Mittlerer Integrierter Quadratischer Fehler

In diesem Abschnitt wird die Schätzfunktion  $\hat{f}_n(\mathbf{y})$  unter den Fehlermaßen Mittlerer Quadratischer Fehler (MSE) und Mittlerer Integrierter Quadratischer Fehler (MISE) untersucht. Das asymptotische Verhalten von  $MSE(\hat{f}_n(\mathbf{y}))$  und  $MISE(\hat{f}_n)$  für  $n \rightarrow \infty$  ist von Interesse, insbesondere die Konvergenzrate, da sie als Grundlage für einen Vergleich der Güte des hier entwickelten Schätzverfahrens mit der Güte anderer Schätzverfahren dient.

Aus der Definition des Mittleren Quadratischen Fehlers,

$$MSE(\hat{f}_n(\mathbf{y})) := E \left[ (\hat{f}_n(\mathbf{y}) - f(\mathbf{y}))^2 \right],$$

folgt

$$MSE(\hat{f}_n(\mathbf{y})) = Bias^2(\hat{f}_n(\mathbf{y})) + var(\hat{f}_n(\mathbf{y})). \quad (3.2)$$

Für  $Bias(\hat{f}_n(\mathbf{y}))$  gilt gemäß Gleichung (3.1) :

$$Bias(\hat{f}_n(\mathbf{y})) = \frac{1}{\lambda_n^2} \int_{\mathcal{B}} K \left( \frac{\mathbf{y} - \mathbf{s}}{\lambda_n} \right) f(\mathbf{s}) \, d\mathbf{s} - f(\mathbf{y}).$$

Da  $f$  keine klassische Ableitung besitzen muß, kann man den obigen Ausdruck nicht mittels einer Taylorentwicklung von  $f$  approximieren, wie es zum Beispiel



in [46] für  $f \in \mathcal{C}^\infty$  vorgeschlagen wird. Der Bias von  $\hat{f}_n$  an der Stelle  $\mathbf{y}$  kann allerdings mittels der Fouriertransformierten von  $f$  an der Stelle  $\mathbf{y}$  ausgedrückt werden. Für die Fouriertransformierte von  $f$  aus  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  gilt

$$\int_{\mathbb{R}^2} (1 + \|\mathbf{t}\|^2)^\alpha (|\Phi_f(\mathbf{t})|^2) d\mathbf{t} < \infty.$$

Die Fouriertransformierte des Integrales

$$\frac{1}{\lambda_n^2} \int_{\mathcal{B}} K\left(\frac{\mathbf{y} - \mathbf{s}}{\lambda_n}\right) f(\mathbf{s}) d\mathbf{s}$$

ist gleich dem Produkt  $\Phi_K(\lambda_n \mathbf{t}) \Phi_f(\mathbf{t})$  ist. Damit ist

$$\text{Bias}(\hat{f}_n(\mathbf{y})) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-i\mathbf{t}'\mathbf{y}} \Phi_K(\lambda_n \mathbf{t}) \Phi_f(\mathbf{t}) d\mathbf{t} - \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-i\mathbf{t}'\mathbf{y}} \Phi_f(\mathbf{t}) d\mathbf{t}. \quad (3.3)$$

Definition 3.7 und Anmerkung 3.3 beschreiben die Wahl der Kernfunktion  $K(\mathbf{y})$  für  $\hat{f}_n(\mathbf{y})$ .

**Definition 3.7** *Die Funktion*

$$K(\mathbf{y}) := \frac{\sin(y_1)}{\pi y_1} \frac{\sin(y_2)}{\pi y_2}.$$

heißt **Fourierintegral–Kernfunktion**.

**Bemerkung 3.3** *Die Fourierintegral–Kernfunktion erfüllt die Bedingungen von Definition 2.3.*

**Beweis:** Ergibt sich sofort aus Definition 3.7.

**Anmerkung 3.3** *Die Kernfunktion des modifizierten Dekonvolutions–Kernschätzers wird gewählt als*

$$K(\mathbf{y}) = \frac{\sin(y_1)}{\pi y_1} \frac{\sin(y_2)}{\pi y_2}.$$

Die Fourierintegral–Kernfunktion tritt in der Literatur zum Beispiel bei Stefanski (1990) [50] oder in [9], [10] auf. Für die Fouriertransformierte der Fourierintegral–Kernfunktion ergibt sich:

$$\Phi_K(\mathbf{t}) = \Phi_{K_1}(t_1) \cdot \Phi_{K_2}(t_2),$$

$$\Phi_{K_i}(t_i) = \begin{cases} 1 & \text{für } |t_i| \leq 1 \\ 0 & \text{sonst} \end{cases} \quad i = 1, 2. \quad (3.4)$$

Mit Anmerkung 3.3 und Gleichung (3.4) folgt für den Bias des modifizierten Dekonvolutions–Kernschätzers:

$$\begin{aligned} \text{Bias}(\hat{f}_n(\mathbf{y})) &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \Phi_f(\mathbf{t}) \Phi_K(\lambda_n \mathbf{t}) - \Phi_f(\mathbf{t}) \, dt \\ &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \Phi_f(\mathbf{t}) (\Phi_K(\lambda_n \mathbf{t}) - 1) \, dt. \end{aligned}$$

Es bezeichne  $\overline{\mathcal{T}}$  das Komplement der Menge  $[-\frac{1}{\lambda_n}, \frac{1}{\lambda_n}] \times [-\frac{1}{\lambda_n}, \frac{1}{\lambda_n}]$  im  $\mathbb{R}^2$ . Mit Gleichung (3.4) ergibt sich

$$\begin{aligned} \text{Bias}(\hat{f}_n(\mathbf{y})) &= -\frac{1}{(2\pi)^2} \int_{\overline{\mathcal{T}}} e^{-it' \mathbf{y}} \Phi_f(t_1, t_2) \, dt_2 \, dt_1 \\ \Rightarrow |\text{Bias}(\hat{f}_n(\mathbf{y}))| &\leq \frac{1}{(2\pi)^2} \int_{\overline{\mathcal{T}}} |\Phi_f(t_1, t_2)| \, dt_2 \, dt_1. \end{aligned} \quad (3.5)$$

Der Bandweite–Parameter  $\lambda_n$  bestimmt den Integrationsbereich in Gleichung (3.5). Der Wert des Integrales in (3.5) wird außerdem beeinflusst vom Verhalten von  $|\Phi_f(t_1, t_2)|$  für betragsgroße  $\mathbf{t}$ . Je langsamer  $|\Phi_f(t_1, t_2)|$  abklingt mit wachsender Norm von  $\mathbf{t}$ , desto langsamer wird  $|\text{Bias}(\hat{f}_n(\mathbf{y}))|$  konvergieren. Das Abklingverhalten von  $|\Phi_f(\mathbf{t})|$  bestimmt außerdem, von welcher Ordnung der Sobolev–Raum ist, zu dem die Funktion  $f(\mathbf{y})$  gehört. Die Glattheits–Eigenschaften von  $f(\mathbf{y})$  beeinflussen also das Konvergenzverhalten von  $\text{Bias}(\hat{f}_n(\mathbf{y}))$ . Der folgende Satz gibt eine Abschätzung für den Bias von  $\hat{f}_n(\mathbf{y})$  an.

**Satz 3.3** *Es sei  $f(\mathbf{y})$  eine Funktion in  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  mit  $\alpha > 1$ . Es bezeichne  $\hat{f}_n(\mathbf{y})$  den modifizierten Dekonvolutions–Kernschätzer mit Fourierintegral–Kernfunktion.*

*Dann gilt*

$$|\text{Bias}(\hat{f}_n(\mathbf{y}))|^2 \leq C \lambda_n^{2\alpha-2}.$$

**Beweis:** Da  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$  ist, gilt  $\int_{\mathbb{R}^2} (1 + \|\mathbf{t}\|^2)^\alpha |\Phi_f(\mathbf{t})|^2 d\mathbf{t} < \infty$ .

Dies impliziert, daß

$$|\Phi_f(\mathbf{t})| \leq C(1 + \|\mathbf{t}\|^2)^s, \quad (3.6)$$

mit  $s < \frac{-1-\alpha}{2}$ . Für  $Bias(\hat{f}_n(\mathbf{y}))$  gilt dann

$$|Bias(\hat{f}_n(\mathbf{y}))| \leq C \int_{\overline{\mathcal{T}}} (1 + \|\mathbf{t}\|^2)^{\frac{-1-\alpha}{2}} dt_1 dt_2.$$

Wegen der Symmetrie des Integranden bezüglich  $\mathbf{t}$ , erfolgt die Integration über dem Quadranten  $[0, \frac{1}{\lambda_n}] \times [0, \frac{1}{\lambda_n}]$ .

Die Koordinatentransformation  $(t_1, t_2) = (r \cos(\varphi), r \sin(\varphi))$  wird durchgeführt. Die Integrationsgrenzen für  $r$  und  $\varphi$  sind

- a)  $\varphi \in [0, \frac{\pi}{4}] \Rightarrow r \in [1/(\lambda_n \cos(\varphi)), \infty]$
- b)  $\varphi \in [\frac{\pi}{4}, \frac{\pi}{2}] \Rightarrow r \in [1/(\lambda_n \sin(\varphi)), \infty]$ .

Damit ergibt sich für den Bias

$$\begin{aligned} Bias(\hat{f}_n(\mathbf{y})) &\leq 4 \left( \int_0^{\frac{\pi}{4}} \int_{\frac{1}{\lambda_n \cos(\varphi)}}^{\infty} (1 + r^2)^{-\frac{1}{2}(\alpha+1)} r dr d\varphi \right) \\ &\quad + 4 \left( \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \int_{\frac{1}{\lambda_n \sin(\varphi)}}^{\infty} (1 + r^2)^{-\frac{1}{2}(\alpha+1)} r dr d\varphi \right). \end{aligned}$$

Hierbei gilt

$$\int_0^{\frac{\pi}{4}} \int_{\frac{1}{\lambda_n \cos(\varphi)}}^{\infty} (1 + r^2)^{-\frac{1}{2}(\alpha+1)} r dr d\varphi = \int_0^{\frac{\pi}{4}} \left[ \frac{(1+r^2)^{-(\alpha+1)/2+1}}{2 \cdot (1-(\alpha+1)/2)} \right]_{\frac{1}{\lambda_n \cos(\varphi)}}^{\infty} d\varphi. \quad (3.7)$$

Das Integral in (3.7) existiert, falls  $-\frac{(\alpha+1)}{2} + 1 < 0$ . Dies ist erfüllt, da als Funktionenklasse ein Sobolev-Raum  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  mit Parameter  $\alpha$  größer als 1 gewählt wurde. Das Integral in (3.7) wird dann zu

$$\begin{aligned}
\int_0^{\frac{\pi}{4}} \int_{\frac{1}{\lambda_n \cos(\varphi)}}^{\infty} (1+r^2)^{-\frac{1}{2}(\alpha+1)} r dr d\varphi &= \int_0^{\frac{\pi}{4}} C \left(1 + \frac{1}{\lambda_n^2 \cos(\varphi)^2}\right)^{-\left(\frac{\alpha+1}{2}\right)+1} d\varphi \\
&\leq C \frac{1}{\lambda_n^2}^{-\frac{\alpha+1}{2}+1} \\
&= C \lambda_n^{\alpha-1}.
\end{aligned}$$

Analog erhält man

$$\begin{aligned}
\left( \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \int_{\frac{1}{\lambda_n \sin(\varphi)}}^{\infty} (1+r^2)^{-\frac{1}{2}(\alpha+1)} r dr d\varphi \right) &= \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \left(1 + \frac{1}{\lambda_n^2 \sin(\varphi)^2}\right)^{-\left(\frac{\alpha+1}{2}\right)+1} d\varphi \\
&\leq C \lambda_n^{\alpha-1}.
\end{aligned}$$

Es ergibt sich dann

$$|Bias(\hat{f}_n(\mathbf{y}))|^2 \leq C \lambda_n^{2\alpha-2}$$

**q.e.d.**

Da  $\lambda_n \rightarrow 0$  mit  $n \rightarrow \infty$  gemäß Definition 2.6, gilt also, daß der modifizierte Dekonvolutions-Kernschätzer mit Fourierintegral-Kernfunktion asymptotisch erwartungstreu ist. Aus Satz 3.3 folgt, daß mit wachsendem Glattheitsparameter  $\alpha$ , die Konvergenzrate des Bias gegen 0 zunimmt.

Als nächstes wird die Varianz von  $\hat{f}_n(\mathbf{y})$  untersucht. Nach Satz 3.2 ist für alle  $\mathbf{y} \in \mathcal{B}$

$$var(\hat{f}_n(\mathbf{y})) \leq \frac{1}{n\lambda_n^2} \frac{1}{(2\pi)^2} C_{bed} \int_{\mathbb{R}^2} \Phi_K(\mathbf{t})^2 \frac{E[G(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi)^2]}{|M(\frac{\mathbf{t}}{\lambda_n})|^2} d\mathbf{t} + C/n E^2[\hat{f}_n(\mathbf{y})].$$

Wie in (3.3) kann  $E^2[\hat{f}_n(\mathbf{y})]$  folgendermaßen dargestellt werden

$$\begin{aligned}
E^2[\hat{f}_n(\mathbf{y})] &= \left( \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-i\mathbf{t}'\mathbf{y}} \Phi_K(\lambda_n \mathbf{t}) \Phi_f(\mathbf{t}) d\mathbf{t} \right)^2 \\
&\leq \frac{1}{(2\pi)^4} \left( \int_{\mathbb{R}^2} |\Phi_K(\lambda_n \mathbf{t})| |\Phi_f(\mathbf{t})| d\mathbf{t} \right)^2.
\end{aligned}$$

Unter Verwendung von (3.4) folgt

$$E^2[\hat{f}_n(\mathbf{y})] \leq \frac{1}{(2\pi)^4} \left( \int_{-\frac{1}{\lambda_n}}^{\frac{1}{\lambda_n}} \int_{-\frac{1}{\lambda_n}}^{\frac{1}{\lambda_n}} |\Phi_f(t_1, t_2)| dt_1 dt_2 \right)^2.$$

Da  $f(\mathbf{y}) \in \mathcal{H}^{\alpha,2}(\mathcal{B})$  ist, kann man (3.6) anwenden.

$$E^2[\hat{f}_n(\mathbf{y})] \leq \frac{1}{(2\pi)^4} \left( C \int_{-\frac{1}{\lambda_n}}^{\frac{1}{\lambda_n}} \int_{-\frac{1}{\lambda_n}}^{\frac{1}{\lambda_n}} (1 + \|\mathbf{t}\|^2)^{\frac{-1-\alpha}{2}} dt_1 dt_2 \right)^2.$$

Hierbei ist  $\alpha > 1$  und  $\frac{-1-\alpha}{2} < -1$ . Mit der Koordinatentransformation  $t_1 = r \cos(\varphi)$ ,  $t_2 = r \sin(\varphi)$  ergibt sich

$$\begin{aligned} E^2[\hat{f}_n(\mathbf{y})] &\leq \frac{1}{(2\pi)^4} C \left( 4 \int_0^{\frac{\pi}{4}} \int_0^{\frac{1}{\lambda_n \cos(\varphi)}} (1+r^2)^{\frac{-1-\alpha}{2}} r dr d\varphi \right. \\ &\quad \left. + 4 \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \int_0^{\frac{1}{\lambda_n \cos(\varphi)}} (1+r^2)^{\frac{-1-\alpha}{2}} r dr d\varphi \right)^2 \\ &= \frac{1}{(2\pi)^4} C \left( \int_0^{\frac{\pi}{4}} \left[ \frac{(1+r^2)^{\frac{(-1-\alpha)/2+1}{2}}}{2^{\frac{(-1-\alpha)/2+1}{2}}} \right]_0^{\frac{1}{\lambda_n \cos(\varphi)}} d\varphi \right. \\ &\quad \left. + \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \left[ \frac{(1+r^2)^{\frac{(-1-\alpha)/2+1}{2}}}{2^{\frac{(-1-\alpha)/2+1}{2}}} \right]_0^{\frac{1}{\lambda_n \sin(\varphi)}} d\varphi \right)^2. \end{aligned}$$

Es bezeichne  $s := \frac{-1-\alpha}{2} + 1 < 0$  und  $q := 2 \left( \frac{-1-\alpha}{2} + 1 \right)$ , wobei  $s < 0$  und  $q < 0$ . Dann ist

$$E^2[\hat{f}_n(\mathbf{y})] \leq C \left( \int_0^{\frac{\pi}{4}} q \left[ \left(1 + \frac{1}{\lambda_n \cos(\varphi)}\right)^s - 1 \right] + \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} q \left[ \left(1 + \frac{1}{\lambda_n \sin(\varphi)}\right)^s - 1 \right] \right)^2. \quad (3.8)$$

Für  $n \rightarrow \infty$  konvergiert  $\lambda_n$  gegen 0 und damit konvergiert (3.8) von unten gegen eine positive Konstante C. Damit gilt für  $n \rightarrow \infty$ :

$$\sup_{f \in \mathcal{H}^{\alpha,2}(\mathcal{B})} E^2[\hat{f}_n(\mathbf{y})] \leq C.$$

Die Varianz des modifizierten Dekonvolutions–Kernschätzers an der Stelle  $\mathbf{y}$  ist somit für alle  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$  nach oben beschränkt durch

$$\text{var}(\hat{f}_n(\mathbf{y})) \leq \frac{1}{n\lambda_n^2} \frac{1}{(2\pi)^2} C_{bed} \int_{\mathbb{R}^2} \Phi_K(\mathbf{t})^2 \frac{E[G(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi)^2]}{|M(\frac{\mathbf{t}}{\lambda_n})|^2} d\mathbf{t} + C/n. \quad (3.9)$$

Nur der erste Summand in (3.9) ist abhängig von der Wahl der Gewichtsfunktion  $G(\mathbf{t}, \Theta_\varphi)$ . Es sollen nun zwei Möglichkeiten für die Wahl einer Gewichtsfunktion betrachtet werden.

1.  $G_1(\mathbf{t}, \Theta_\varphi) := \begin{cases} 1, & |\mathbf{t}'\Theta_\varphi| \leq \delta, \quad \delta > 0 \\ 0, & \text{sonst.} \end{cases}$
2.  $G_2(\mathbf{t}, \Theta_\varphi) := e^{-(\mathbf{t}'\Theta_\varphi)^2 \frac{\sigma^2}{2}}.$

**Satz 3.4** *Es sei  $\hat{f}_n(\mathbf{y})$  der modifizierte Dekonvolutions–Kernschätzer mit Gewichtsfunktion  $G(\mathbf{t}, \Theta_\varphi) = G_1(\mathbf{t}, \Theta_\varphi)$ . Dann ist für große  $n$  eine obere Schranke für  $\text{var}(\hat{f}_n(\mathbf{y}))$  gegeben durch*

$$\text{var}(\hat{f}_n(\mathbf{y})) \leq C_{bed} \frac{1}{n\lambda_n^3} \frac{e^{\delta^2 \sigma^2}}{2\delta\pi}.$$

Hierbei ist  $C_{bed} \geq \sup_{\mathbf{x}} \sup_{\varphi} w(\mathbf{x}|\varphi)$  wie im Beweis zu Satz 3.2.

**Beweis:** Es bezeichne

$$B_\delta(\frac{\mathbf{t}}{\lambda_n}) = \{\varphi \in [0, \pi], \|\frac{\mathbf{t}}{\lambda_n}\| \|\Theta_\varphi\| |\cos(\angle(\mathbf{t}, \Theta_\varphi))| \leq \delta\},$$

$$\mu(B_\delta(\frac{\mathbf{t}}{\lambda_n})) := \text{Lebesgue–Ma\ss der Menge } B_\delta(\frac{\mathbf{t}}{\lambda_n}).$$

Dann ist

$$E[G_1(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi)^2] = \int_{B_\delta(\frac{\mathbf{t}}{\lambda_n})} \frac{1}{\pi} d\varphi = \frac{1}{\pi} \mu(B_\delta(\frac{\mathbf{t}}{\lambda_n}))$$

und

$$M(\frac{\mathbf{t}}{\lambda_n}) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{\pi} \int_{\mathbb{R}} \int_0^\pi e^{ir \frac{\mathbf{t}}{\lambda_n}' \Theta_\varphi} G_1(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi) e^{-\frac{r^2}{2\sigma^2}} dr d\varphi.$$

Vertauschen der Integrationsreihenfolge ergibt

$$\begin{aligned}
M\left(\frac{\mathbf{t}}{\lambda_n}\right) &= \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{\pi} \int_0^\pi G_1\left(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi\right) \int_{\mathbb{R}} e^{i\frac{\mathbf{t}}{\lambda_n} \Theta_\varphi r} e^{-\frac{r^2}{2\sigma^2}} dr d\varphi \\
&= \frac{1}{\pi} \int_0^\pi G_1\left(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi\right) e^{-\left(\frac{\mathbf{t}}{\lambda_n} \Theta_\varphi\right)^2 \frac{\sigma^2}{2}} d\varphi.
\end{aligned}$$

Die letzte Gleichung folgt aus der Berechnung der Fouriertransformierten von  $e^{-\frac{r^2}{2\sigma^2}}$ . Setzt man nun für  $G_1(\mathbf{t}, \Theta_\varphi)$  ein, so ergibt sich

$$M\left(\frac{\mathbf{t}}{\lambda_n}\right) = \frac{1}{\pi} \int_{B_\delta\left(\frac{\mathbf{t}}{\lambda_n}\right)} e^{-\left(\frac{\mathbf{t}}{\lambda_n} \Theta_\varphi\right)^2 \frac{\sigma^2}{2}} d\varphi. \quad (3.10)$$

Die folgende Bedingung an  $\mathbf{t}' \Theta_\varphi$  muß erfüllt sein:

$$\begin{aligned}
0 &\leq \left| \frac{\mathbf{t}}{\lambda_n} \Theta_\varphi \right|^2 \leq \delta^2 \\
\Rightarrow 1 &\geq e^{-\left(\frac{\mathbf{t}}{\lambda_n} \Theta_\varphi\right)^2} \geq e^{-\delta^2} \\
\Rightarrow M\left(\frac{\mathbf{t}}{\lambda_n}\right) &\geq \frac{1}{\pi} \cdot e^{-\delta^2 \frac{\sigma^2}{2}} \cdot \mu(B_\delta\left(\frac{\mathbf{t}}{\lambda_n}\right)) > 0 \\
\Rightarrow \frac{1}{|M\left(\frac{\mathbf{t}}{\lambda_n}\right)|^2} &\leq \pi^2 e^{\delta^2 \sigma^2} \frac{1}{\mu(B_\delta\left(\frac{\mathbf{t}}{\lambda_n}\right))^2} \\
\Rightarrow \frac{E[G_1\left(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi\right)^2]}{|M\left(\frac{\mathbf{t}}{\lambda_n}\right)|^2} &\leq \pi e^{\delta^2 \sigma^2} \frac{1}{\mu(B_\delta\left(\frac{\mathbf{t}}{\lambda_n}\right))}.
\end{aligned}$$

Es gilt für  $\varphi$ :

$$\begin{aligned}
\varphi \in B_\delta\left(\frac{\mathbf{t}}{\lambda_n}\right) &\iff (|\mathbf{t}|/\lambda_n) |\cos(\angle(\mathbf{t}, \Theta_\varphi))| \leq \delta \\
&\iff |\cos(\angle(\mathbf{t}, \Theta_\varphi))| \leq (\delta\lambda_n)/\|\mathbf{t}\|.
\end{aligned}$$

**1. Fall:**

$$\begin{aligned}
\text{Wenn } (\delta\lambda_n)/\|\mathbf{t}\| \geq 1 &\text{ folgt } B_\delta\left(\frac{\mathbf{t}}{\lambda_n}\right) = [0, \pi] \\
&\Rightarrow \mu(B_\delta\left(\frac{\mathbf{t}}{\lambda_n}\right)) = \pi.
\end{aligned}$$

**2. Fall:**

Wenn  $0 < (\delta\lambda_n)/\|\mathbf{t}\| < 1$  dann ergibt sich die Forderung

$$\angle(\mathbf{t}, \Theta_\varphi) \in U_\Delta(\frac{\pi}{2})$$

$$\Delta = \frac{\pi}{2} - \arccos(\delta\lambda_n/\|\mathbf{t}\|).$$

$U_\Delta(\frac{\pi}{2})$  ist die  $\Delta$ -Umgebung von  $\frac{\pi}{2}$ . Es sei  $\mathbf{t} \perp \Theta_{\varphi_0}$  für ein  $\varphi_0 \in [0, \pi]$ . Dann ist

$$B_\delta(\frac{\mathbf{t}}{\lambda_n}) = [\varphi_0 - \Delta, \varphi_0 + \Delta].$$

Es folgt

$$\mu(B_\delta(\frac{\mathbf{t}}{\lambda_n})) = 2\Delta = \pi - 2 \arccos(\delta\lambda_n/\|\mathbf{t}\|).$$

Mit der Beziehung  $0 < \arccos(x) \leq -\tan(\frac{\pi}{4}) \cdot x + \frac{\pi}{2}$  folgt für  $\mu(B_\delta(\frac{\mathbf{t}}{\lambda_n}))$ :

$$\begin{aligned} \mu(B_\delta(\frac{\mathbf{t}}{\lambda_n})) &\geq \pi - 2 \cdot (-\tan(\frac{\pi}{4})(\delta\lambda_n)/\|\mathbf{t}\| + \frac{\pi}{2}) \\ &= 2 \tan(\frac{\pi}{4})(\delta\lambda_n)/\|\mathbf{t}\|. \end{aligned}$$

Dann ist folgende Ungleichung erfüllt

$$\frac{E[G_1^2(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi)]}{|M(\frac{\mathbf{t}}{\lambda_n})|^2} \leq \frac{\pi}{2} e^{\delta^2 \sigma^2} \|\mathbf{t}\| / (\delta\lambda_n).$$

Als obere Schranke für  $var(\hat{f}_n(\mathbf{y}))$  ergibt sich

$$\begin{aligned} var(\hat{f}_n(\mathbf{y})) &\leq \frac{1}{n\lambda_n^2} \frac{1}{(2\pi)^2} C_{bed} \pi e^{\delta^2 \sigma^2} \left( \int_{\|\mathbf{t}\| \leq \delta\lambda_n} \Phi_K^2(\mathbf{t}) \frac{1}{\pi} d\mathbf{t} \right. \\ &\quad \left. + \int_{\|\mathbf{t}\| > \delta\lambda_n} \Phi_K^2(\mathbf{t}) \frac{1}{2} \frac{\|\mathbf{t}\|}{\delta\lambda_n} d\mathbf{t} \right) + (C/n). \end{aligned}$$

Die Definition für  $\Phi_K(\mathbf{t})$  aus Gleichung (3.4) wird eingesetzt.



$$\begin{aligned}
\text{var}(\hat{f}_n(\mathbf{y})) &\leq \frac{1}{n\lambda_n^2} \frac{1}{4\pi} C_{bed} e^{\delta^2\sigma^2} \left( \delta^2\lambda_n^2 + 2 \int_0^1 \int_0^1 \frac{\|\mathbf{t}\|}{\lambda_n\delta} dt_1 dt_2 \right. \\
&\quad \left. - 2 \int_0^{\frac{\pi}{2}} \int_0^{\delta\lambda_n} \frac{s^2}{\lambda_n\delta} ds d\vartheta \right) + (C/n) \\
&\leq \frac{1}{n\lambda_n^2} \frac{e^{\delta^2\sigma^2} C_{bed}}{4\pi} \left( \delta^2\lambda_n^2 + \frac{2}{\lambda_n\delta} - \frac{\delta^2\lambda_n^2\pi}{3} \right) + (C/n).
\end{aligned}$$

Die obere Schranke der Varianz ist damit

$$\text{var}(\hat{f}_n(\mathbf{y})) \leq \frac{e^{\delta^2\sigma^2} C_{bed}}{4\pi} \left( \frac{\delta^2}{n} + \frac{2}{n\lambda_n^3\delta} - \frac{\delta^2\pi}{3n} \right) + (C/n). \quad (3.11)$$

Für großes  $n$  wird (3.11) dominiert von  $\frac{e^{\delta^2\sigma^2}}{2\pi\delta} C_{bed} \frac{1}{n\lambda_n^3}$ . Daraus folgt die Behauptung.

**q.e.d.**

Es wird deutlich, daß  $\lambda_n$  so gewählt werden muß, daß  $\frac{1}{n\lambda_n^3}$  mit  $n \rightarrow \infty$  gegen 0 konvergiert.

**Satz 3.5** *Es bezeichne  $\hat{f}_n(\mathbf{y})$  den modifizierten Dekonvolutions-Kernschätzer mit Gewichtsfunktion  $G(\mathbf{t}, \Theta_\varphi) = G_2(\mathbf{t}, \Theta_\varphi)$ . Dann ergibt sich für große  $n$  als obere Schranke für  $\text{var}(\hat{f}_n(\mathbf{y}))$*

$$\text{var}(\hat{f}_n(\mathbf{y})) \leq C_{bed} \frac{1}{n\lambda_n^3} \frac{\sigma}{\sqrt{\pi^3}}.$$

**Beweis:** Aus dem Beweis zu Satz 3.4 ist bekannt, daß

$$\begin{aligned}
M\left(\frac{\mathbf{t}}{\lambda_n}\right) &= \frac{1}{\pi} \int_0^\pi e^{-\left(\frac{\mathbf{t}}{\lambda_n}\right)' \Theta_\varphi)^2 \sigma^2 / 2} G_2\left(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi\right) d\varphi \\
&= \frac{1}{\pi} \int_0^\pi e^{-\left(\frac{\mathbf{t}}{\lambda_n}\right)' \Theta_\varphi)^2 \sigma^2} d\varphi.
\end{aligned} \quad (3.12)$$

Mit der Gleichung

$$E[G_2^2\left(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi\right)] = \int_0^\pi e^{-\left(\frac{\mathbf{t}}{\lambda_n}\right)' \Theta_\varphi)^2 \sigma^2} \frac{1}{\pi} d\varphi = |M\left(\frac{\mathbf{t}}{\lambda_n}\right)|$$

ergibt sich folgende Abschätzung für  $\text{var}(\hat{f}_n(\mathbf{y}))$

$$\text{var}(\hat{f}_n(\mathbf{y})) \leq \frac{1}{(2\pi)^2} C_{bed} \frac{1}{n\lambda_n^2} \int_{\mathbb{R}^2} \Phi_K(\mathbf{t})^2 \frac{1}{|M(\frac{\mathbf{t}}{\lambda_n})|} d\mathbf{t} + C/n.$$

Eine Abschätzung für  $M(\frac{\mathbf{t}}{\lambda_n})$  muß gefunden werden.

$$\begin{aligned} M(\frac{\mathbf{t}}{\lambda_n}) &= \frac{1}{\pi} \int_0^\pi e^{-(\sigma^2 \|\frac{\mathbf{t}}{\lambda_n}\|^2 \cos^2(\angle(\mathbf{t}, \Theta_\varphi)))} d\varphi \\ &= \frac{1}{\pi} \int_0^\pi e^{-(\sigma^2 \|\frac{\mathbf{t}}{\lambda_n}\|^2 \cos(\varphi)^2)} d\varphi. \end{aligned}$$

Es gilt, daß

$$\frac{1}{\pi} \int_0^\pi e^{-(\sigma^2 \|\frac{\mathbf{t}}{\lambda_n}\|^2 \cos(\varphi)^2)} = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} e^{-(\sigma^2 \|\frac{\mathbf{t}}{\lambda_n}\|^2 \sin(\varphi)^2)}.$$

Mit der Taylor-Entwicklung von  $\sin^2(\varphi)$  folgt, daß

$$M(\frac{\mathbf{t}}{\lambda_n}) \geq \frac{2}{\pi} \int_0^{\frac{\pi}{2}} e^{-(\sigma^2 \|\frac{\mathbf{t}}{\lambda_n}\|^2 \varphi^2)} d\varphi. \quad (3.13)$$

Man führt folgende Substitution durch

$$\|\frac{\mathbf{t}}{\lambda_n}\| \varphi := u, \quad \frac{d\varphi}{du} = \frac{1}{\|\frac{\mathbf{t}}{\lambda_n}\|}.$$

Mit  $\|\frac{\mathbf{t}}{\lambda_n}\| \rightarrow \infty$  für  $n \rightarrow \infty$  folgt

$$\begin{aligned} M(\frac{\mathbf{t}}{\lambda_n}) &\geq \frac{2}{\pi} \|\frac{\mathbf{t}}{\lambda_n}\|^{-1} \int_0^\infty e^{-\sigma^2 u^2} du \\ &= \frac{2}{\pi} \frac{\sqrt{\pi}}{2\sigma} \|\frac{\mathbf{t}}{\lambda_n}\|^{-1}. \end{aligned}$$

Damit ergibt sich für  $\text{var}(\hat{f}_n(\mathbf{y}))$ :

$$\begin{aligned} \text{var}(\hat{f}_n(\mathbf{y})) &\leq \frac{1}{(2\pi)^2} C_{bed} \frac{1}{n\lambda_n^2} \int_{\mathbb{R}^2} \sqrt{\pi}\sigma \frac{\|\mathbf{t}\|}{\lambda_n} \Phi_K(\mathbf{t})^2 d\mathbf{t} + C/n. \\ &= \frac{1}{n\lambda_n^3} \frac{\sigma}{\sqrt{\pi^3}} C_{bed} \int_0^1 \int_0^1 \|\mathbf{t}\| dt_1 dt_2 + C/n \\ &\leq \frac{\sigma C_{bed}}{\sqrt{\pi^3}} \frac{1}{n\lambda_n^3} + C/n. \end{aligned}$$

Dies wird für große  $n$  vom ersten Summanden dominiert und es folgt die Behauptung.

**q.e.d.**

Um einen Vergleich zwischen der oberen Schranke von  $var(\hat{f}_n(\mathbf{y}))$  für  $G(\mathbf{t}, \Theta_\varphi) = G_1(\mathbf{t}, \Theta_\varphi)$  und für  $G(\mathbf{t}, \Theta_\varphi) = G_2(\mathbf{t}, \Theta_\varphi)$  herbeizuführen, werden die Konstanten in Satz 3.4 und Satz 3.5 verglichen. Die Gewichtsfunktion ist als eine Funktion von  $|\mathbf{t}'\Theta_\varphi|$  modelliert. Es sei  $G_1(\mathbf{t}, \Theta_\varphi) = g_1(b)$  und  $G_2(\mathbf{t}, \Theta_\varphi) = g_2(b)$  wobei  $b := |\mathbf{t}'\Theta_\varphi|$  ist. Stellt man die Forderung, daß das Integral jeder Gewichtsfunktion  $g(b)$  über dem Bereich  $b \in [-\infty, \infty]$  den gleichen Wert annimmt, so muß die Bedingung

$$\int_{\mathbb{R}} g_1(b) db = \int_{\mathbb{R}} g_2(b) db$$

erfüllt sein und es folgt

$$2\delta = \frac{\sqrt{2\pi}}{\sigma}.$$

Für die obere Schranke der Varianz bei Wahl von  $G_1(\mathbf{t}, \Theta_\varphi)$  ergibt sich hiermit

$$var(\hat{f}_n(\mathbf{y})) \leq var_1 := C_{bed} \frac{1}{n\lambda_n^3} \frac{e^{\frac{\pi}{2}}\sigma}{\sqrt{2}\sqrt{\pi^3}}$$

und bei Verwendung von  $G_2(\mathbf{t}, \Theta_\varphi)$  ist

$$var(\hat{f}_n(\mathbf{y})) \leq var_2 := C_{bed} \frac{1}{n\lambda_n^3} \frac{\sigma}{\sqrt{\pi^3}} < var_1.$$

Damit ist bei Wahl der Gewichtsfunktion  $G_2$  die obere Schranke für die Varianz des Schätzers kleiner als bei Wahl von  $G_1$ . Es kann sogar gezeigt werden, daß die Gewichtsfunktion  $G_2$  die obere Schranke der Varianz minimiert und daher die optimale Wahl darstellt: es bezeichne  $G(\mathbf{t}, \Theta_\varphi)$  eine beliebige Gewichtsfunktion. Setzt man in Satz 3.2 für  $M(\mathbf{t})$  ein, so ergibt sich

$$var(\hat{f}_n(\mathbf{y})) \leq C_{bed} \frac{1}{n\lambda_n^2} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \Phi_K(\lambda_n \mathbf{t}) \frac{E[G(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi)^2]}{|E[e^{-(\mathbf{t}'\Theta_\varphi)^2 \frac{\sigma^2}{2}} G(\mathbf{t}, \Theta_\varphi)]|^2} d\mathbf{t} + (C/n).$$

Der zweite Summand ist - unabhängig von der Wahl von  $G(\mathbf{t}, \Theta_\varphi)$  - von der Ordnung  $\frac{1}{n}$ . Der erste Summand konvergiert langsamer gegen 0 als  $\frac{1}{n\lambda_n^2}$  und dominiert also die obere Schranke der Varianz. Um die kleinste obere Schranke zu erhalten, muß  $G(\mathbf{t}, \Theta_\varphi)$  so gewählt werden, daß das Integral

$$\int_{\mathbb{R}^2} \Phi_K(\lambda_n \mathbf{t}) \frac{E[G(\frac{\mathbf{t}}{\lambda_n}, \Theta_\varphi)^2]}{|E[e^{-(\mathbf{t}'\Theta_\varphi)^2 \frac{\sigma^2}{2}} G(\mathbf{t}, \Theta_\varphi)]|^2} d\mathbf{t}$$

minimal wird. Die Cauchy–Schwartz Ungleichung ergibt

$$\frac{1}{|E[e^{-(\mathbf{t}'\Theta_\varphi)^2 \frac{\sigma^2}{2}} G(\mathbf{t}, \Theta_\varphi)]|^2} \geq \frac{1}{E[e^{-(\mathbf{t}'\Theta_\varphi)^2 \sigma^2}] E[G^2(\mathbf{t}, \Theta_\varphi)]},$$

wobei das Gleichheitszeichen gilt, wenn  $G(\mathbf{t}, \Theta_\varphi) = e^{-(\mathbf{t}'\Theta_\varphi)^2 \frac{\sigma^2}{2}}$  ist, und damit liefert die Wahl  $G(\mathbf{t}, \Theta_\varphi) = G_2(\mathbf{t}, \Theta_\varphi)$  die kleinste obere Schranke für die Varianz des modifizierten Dekonvolutions–Kernschätzers.

**Anmerkung 3.4** Für den modifizierten Dekonvolutions–Kernschätzer wird die Gewichtsfunktion

$$G(\mathbf{t}, \Theta_\varphi) = e^{-(\mathbf{t}'\Theta_\varphi)^2 \frac{\sigma^2}{2}}$$

gewählt.

Eine obere Schranke für die Konvergenzrate des Mittleren Quadratischen Fehlers liefert der folgende

**Satz 3.6** Für den modifizierten Dekonvolutions–Kernschätzer mit Fourierintegral–Kernfunktion und Gewichtsfunktion  $G(\mathbf{t}, \Theta_\varphi) = e^{-(\mathbf{t}'\Theta_\varphi)^2 \frac{\sigma^2}{2}}$  ist die optimale Bandweite unter dem Mittleren Quadratischen Fehler gegeben durch

$$\lambda_n^{MSE} = C n^{-\frac{1}{2\alpha+1}}.$$

Eine obere Schranke für den Mittleren Quadratischen Fehler ist dann gegeben durch

$$MSE(\hat{f}_n(\mathbf{y})) \leq C n^{-\frac{2(\alpha-1)}{2\alpha+1}}.$$

**Beweis:** Es gilt gemäß Gleichung (3.2), daß

$$MSE(\hat{f}_n(\mathbf{y})) = Bias(\hat{f}_n(\mathbf{y}))^2 + var(\hat{f}_n(\mathbf{y})).$$

Mit Satz 3.3 und Satz 3.5 ergibt sich dann

$$MSE(\hat{f}_n(\mathbf{y})) \leq C \lambda_n^{2\alpha-2} + C \frac{1}{n\lambda_n^3}. \quad (3.14)$$

Um das Minimum der rechten Seite zu ermitteln, wird diese nach  $\lambda_n$  abgeleitet. Dann ergibt sich das optimale

$$\lambda_n^{MSE} = C n^{-\frac{1}{2\alpha+1}}.$$

und mit (3.14) liefert dies

$$MSE(\hat{f}_n(\mathbf{y})) \leq C n^{-\frac{2\alpha-2}{2\alpha+1}}.$$

**q.e.d.**

Bei einer Folge  $b_n = n^{-r}$ , nennen wir  $r$  die Konvergenzrate von  $b_n$ . Der Bias von  $\hat{f}_n(\mathbf{y})$  konvergiert mit zunehmender Konvergenzrate von  $\lambda_n$  schneller gegen 0, die Varianz von  $\hat{f}_n(\mathbf{y})$  konvergiert mit abnehmender Konvergenzrate von  $\lambda_n$  schneller gegen 0. Die Bestimmung des Minimums von  $MSE(\hat{f}_n(\mathbf{y}))$  in  $\lambda_n$  entspricht dem Gleichsetzen der Konvergenzraten von Bias und Varianz, um optimale Konvergenzgeschwindigkeit von  $MSE(\hat{f}_n(\mathbf{y}))$  gegen 0 zu erreichen. Da das Ergebnis für den Mittleren Quadratischen Fehler in Satz 3.6 unabhängig von  $\mathbf{y}$  ist, läßt sich das Ergebnis für den Mittleren Integrierten Quadratischen Fehler einfach aus Satz 3.6 ableiten.

**Satz 3.7** *Für den modifizierten Dekonvolutions-Kernschätzer mit Fourierintegral-Kernfunktion und mit Gewichtsfunktion  $G(\mathbf{t}, \Theta_\varphi) = e^{-(\mathbf{t}'\Theta_\varphi)^2 \frac{\sigma^2}{2}}$  ist die optimale Bandweite hinsichtlich Mittlerem Integrierten Quadratischen Fehler gegeben durch*

$$\lambda_n^{MISE} = C n^{-\frac{1}{2\alpha+1}}.$$

*Der Mittlere Integrierte Quadratische Fehler ist dann nach oben beschränkt durch*

$$MISE(\hat{f}_n) \leq C n^{-\frac{(2\alpha-2)}{2\alpha+1}}.$$

**Beweis:** Folgt direkt aus Satz 3.6.

### 3.2.3 Das Verhalten des Schätzers in der Supremum–Norm

In diesem Abschnitt wird das Konvergenzverhalten des Rekonstruktionsfehlers  $\hat{f}_n(\mathbf{y}) - f(\mathbf{y})$  in der Supremum–Norm untersucht. Eine obere Schranke wird hergeleitet. Mit Hilfe der Dreiecksungleichung erhält man zunächst

$$\sup_{\mathcal{B}} | \hat{f}_n(\mathbf{y}) - f(\mathbf{y}) | \leq \sup_{\mathcal{B}} | \hat{f}_n(\mathbf{y}) - E[\hat{f}_n(\mathbf{y})] | + \sup_{\mathcal{B}} | E[\hat{f}_n(\mathbf{y})] - f(\mathbf{y}) | . \quad (3.15)$$

Der zweite Summand in (3.15) ist dann die maximale Verzerrung

$$\sup_{\mathcal{B}} | E[\hat{f}_n(\mathbf{y})] - f(\mathbf{y}) | = \sup_{\mathcal{B}} | Bias(\hat{f}_n(\mathbf{y})) | .$$

Mit Satz 3.3 folgt

$$\sup_{\mathcal{B}} | \hat{f}_n(\mathbf{y}) - f(\mathbf{y}) | \leq \sup_{\mathcal{B}} | \hat{f}_n(\mathbf{y}) - E[\hat{f}_n(\mathbf{y})] | + C \lambda_n^{\alpha-1} . \quad (3.16)$$

Der erste Summand läßt sich umformen zu

$$\begin{aligned} \sup_{\mathcal{B}} | \hat{f}_n(\mathbf{y}) - E[\hat{f}_n(\mathbf{y})] | &= \sup_{\mathcal{B}} \left| \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \frac{\frac{1}{n} \sum_{j=1}^n e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) d\mathbf{t} \right. \\ &\quad \left. - \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \Phi_K(\lambda_n \mathbf{t}) \Phi_f(\mathbf{t}) d\mathbf{t} \right| . \end{aligned}$$

Es sei die Funktion  $D_j(\mathbf{y})$  definiert als

$$D_j(\mathbf{y}) := \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \left( \frac{e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} - \Phi_f(\mathbf{t}) \right) \Phi_K(\lambda_n \mathbf{t}) d\mathbf{t} .$$

Weiterhin sei

$$\Delta_n(\mathbf{y}) := \frac{1}{n} \sum_{j=1}^n D_j(\mathbf{y}) .$$

Dann ist

$$\sup_{\mathcal{B}} | \hat{f}_n(\mathbf{y}) - E[\hat{f}_n(\mathbf{y})] | = \sup_{\mathcal{B}} | \Delta_n(\mathbf{y}) | . \quad (3.17)$$

**Definition 3.8** Es sei  $(m_n)_{n \in \mathbb{N}}$  eine Zahlenfolge in  $\mathbb{N}$  mit  $m_n \rightarrow \infty$  für  $n \rightarrow \infty$ . Es sei  $(i, j)$  ein Indexpaar mit  $i = 0, \dots, m_n - 1$  und  $j = 0, \dots, m_n - 1$ . Ein **äquidistantes Gitter**  $\mathcal{G}$  mit Feinheit  $m_n$  über dem Bereich  $[-1, 1] \times [-1, 1]$  ist definiert über die Mittelpunkte

$$\mathbf{y}_{M_{ij}} = (y_{M_{ij1}}, y_{M_{ij2}}) := \left( -1 + \frac{(2i+1)}{m_n}, 1 - \frac{(2j+1)}{m_n} \right)$$

und über die Segmentlänge  $\frac{2}{m_n}$ . Ein Gittersegment ist definiert als

$$G_{ij} := \{(y_1, y_2) \in [-1, +1] \times [-1, +1] \text{ wobei}$$

$$y_1 \in [y_{M_{ij1}} - \frac{1}{m_n}, y_{M_{ij1}} + \frac{1}{m_n}] \text{ und } y_2 \in [y_{M_{ij2}} - \frac{1}{m_n}, y_{M_{ij2}} + \frac{1}{m_n}]\}.$$

**Bezeichnung 3.1** Es bezeichne  $k$  einen skalaren Index mit  $k = j m_n + i + 1$ .

Der Index  $k$  wird nun zur Numerierung der Gittermittelpunkte und Gittersegmente verwendet. Man erhält somit die Menge der Gittermittelpunkte

$$\mathcal{Y}_M := \{\mathbf{y}_{M_k} = \mathbf{y}_{M_{ij}} \mid k = j m_n + i + 1, i, j = 0, \dots, m_n - 1\}$$

und die Menge aller Gittersegmente

$$\mathcal{G} := \{G_k = G_{ij}, k = j m_n + i + 1 \mid i, j = 0, \dots, m_n - 1\}.$$

wobei  $y_{M_{ij}}$  und  $G_{ij}$  wie in Definition 3.8 definiert sind.

**Definition 3.9** Ein äquidistantes Gitter  $\tilde{\mathcal{G}}$  über der Menge  $\mathcal{B}$  ist definiert über die Teilmenge der Gittermittelpunkte

$$\tilde{\mathcal{Y}}_M := \{\mathbf{y}_{M_k} \in \mathcal{Y}_M \mid G_k \subset \mathcal{B}\}.$$

**Bezeichnung 3.2** Ein Gittersegment  $G_k \in \tilde{\mathcal{G}}$  wird mit  $\tilde{G}_k$  bezeichnet.

Es gibt  $K_n$  Gittersegmente  $\tilde{G}_k$ ,  $K_n < m_n^2$ .

**Definition 3.10** *Es werde die folgende Teilmenge von  $\mathcal{B}$  definiert.*

$$\tilde{\mathcal{B}} := \bigcup_{k=1}^{K_n} \tilde{G}_k.$$

$\tilde{\mathcal{B}}$  nähert für  $n \rightarrow \infty$  die Menge  $\mathcal{B}$  beliebig genau an. Das Supremum in (3.17) kann angenähert werden durch

$$\begin{aligned} \sup_{\mathcal{B}} |\hat{f}_n(\mathbf{y}) - E[\hat{f}_n(\mathbf{y})]| &\leq \sup_{\tilde{\mathcal{B}}} |\hat{f}_n(\mathbf{y}) - E[\hat{f}_n(\mathbf{y})]| \\ &= \max_k \sup_{\tilde{G}_k} |\Delta_n(\mathbf{y})|. \end{aligned}$$

Dies kann in zwei Anteile zerlegt werden.

$$\begin{aligned} \max_k \sup_{\tilde{G}_k} |\Delta_n(\mathbf{y})| &= \max_k \sup_{\tilde{G}_k} |\Delta_n(\mathbf{y}_{M_k}) - \Delta_n(\mathbf{y}_{M_k}) + \Delta_n(\mathbf{y})| \\ &\leq \max_k |\Delta_n(\mathbf{y}_{M_k})| + \max_k \sup_{\tilde{G}_k} |\Delta_n(\mathbf{y}) - \Delta_n(\mathbf{y}_{M_k})|. \end{aligned} \tag{3.18}$$

Eine Abschätzung für den zweiten Summanden in (3.18) enthält der folgende

**Satz 3.8** *Gegeben sei ein äquidistantes Gitter  $\tilde{\mathcal{G}}$  über  $\mathcal{B}$  mit Segmentlänge  $\frac{2}{m_n}$ .  $(m_n)_{n \in \mathbb{N}}$  sei eine Folge mit  $m_n = O(n^\gamma)$ ,  $\gamma > 0$ . Es bezeichne  $K_n$  die Anzahl der Gittersegmente  $\tilde{G}_k$  in  $\tilde{\mathcal{G}}$ . Es sei  $k \in \{1, \dots, K_n\}$  und  $K_n < m_n^2$ . Mit der Definition*

$$\begin{aligned} D_j(\mathbf{y}) &:= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \left( \frac{e^{-it' \mathbf{x}_j G(\mathbf{t}, \Theta_{\varphi_j})}}{M(\mathbf{t})} - \Phi_f(\mathbf{t}) \right) \Phi_K(\lambda_n \mathbf{t}) d\mathbf{t}, \\ \Delta_n(\mathbf{y}) &:= \frac{1}{n} \sum_{j=1}^n D_j(\mathbf{y}) \end{aligned}$$

*gilt*

$$\max_k \sup_{\tilde{G}_k} |\Delta_n(\mathbf{y}) - \Delta_n(\mathbf{y}_{M_k})| = C n^{-\gamma} \lambda_n^{-4}.$$

**Beweis:** Einsetzen für  $\Delta_n$  ergibt zunächst die folgende Abschätzung:



$$\begin{aligned}
& \max_k \sup_{\tilde{G}_k} |\Delta_n(\mathbf{y}) - \Delta_n(\mathbf{y}_{M_k})| \\
&= \max_k \sup_{\tilde{G}_k} \left| \frac{1}{n} \sum_{j=1}^n \left( \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}} \left( \frac{e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} - \Phi_f(\mathbf{t}) \right) \Phi_K(\lambda_n \mathbf{t}) dt \right. \right. \\
&\quad \left. \left. - \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}_{M_k}} \left( \frac{e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} - \Phi_f(\mathbf{t}) \right) \Phi_K(\lambda_n \mathbf{t}) dt \right) \right| \\
&\leq \max_k \sup_{\tilde{G}_k} \frac{1}{n} \sum_{j=1}^n \left( \frac{1}{(2\pi)^2} \left| \int_{\mathbb{R}^2} (e^{-it' \mathbf{y}} - e^{-it' \mathbf{y}_{M_k}}) \right. \right. \\
&\quad \left. \left. (e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j}) - \Phi_f(\mathbf{t}) M(\mathbf{t})) \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})} dt \right| \right) \\
&\leq \max_k \sup_{\tilde{G}_k} \frac{1}{n} \sum_{j=1}^n \left( \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} |(e^{-it' \mathbf{y}} - e^{-it' \mathbf{y}_{M_k}})| \right. \\
&\quad \left. |(e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j}) - \Phi_f(\mathbf{t}) M(\mathbf{t})) \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})}| dt \right).
\end{aligned}$$

Mit  $\|\mathbf{t}\| \leq \frac{\sqrt{2}}{\lambda_n}$  folgt

$$|e^{-it' \mathbf{y}_{M_k}} - e^{-it' \mathbf{y}}| \leq |\mathbf{y} - \mathbf{y}_{M_k}| \frac{\sqrt{2}}{\lambda_n} \leq \frac{2}{m_n \lambda_n}$$

und damit

$$\begin{aligned}
& \max_k \sup_{\tilde{G}_k} |\Delta_n(\mathbf{y}) - \Delta_n(\mathbf{y}_{M_k})| \\
&\leq \frac{1}{n} \frac{1}{(2\pi)^2} \frac{2}{m_n \lambda_n} \sum_{j=1}^n \int_{\mathbb{R}^2} |e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j}) - \Phi_f(\mathbf{t}) M(\mathbf{t})| \cdot \left| \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})} \right| dt.
\end{aligned}$$

Mit der Abschätzung

$$|e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j}) - \Phi_f(\mathbf{t}) M(\mathbf{t})| \leq |G(\mathbf{t}, \Theta_{\varphi_j})| + \left| \int_{\mathcal{B}} e^{it' \mathbf{y}} f(\mathbf{y}) d\mathbf{y} \right| |M(\mathbf{t})| \leq 2$$

erhält man

$$\max_k \sup_{\tilde{G}_k} |\Delta_n(\mathbf{y}) - \Delta_n(\mathbf{y}_{M_k})| \leq \frac{1}{n} \frac{1}{(2\pi)^2} \frac{4}{m_n \lambda_n} \sum_{j=1}^n \int_{\mathbb{R}^2} \left| \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})} \right| d\mathbf{t}.$$

Im Beweis zu Satz 3.5. wurde gezeigt, daß  $M(\mathbf{t}) \geq C \|\mathbf{t}\|^{-1}$  ist. Damit ergibt sich

$$\int_{\mathbb{R}^2} \left| \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})} \right| d\mathbf{t} = C \int_0^{\frac{1}{\lambda_n}} \int_0^{\frac{1}{\lambda_n}} \left| \frac{1}{M(\mathbf{t})} \right| dt_1 dt_2 \leq \frac{1}{\lambda_n^3}.$$

Wenn nun die Gitterfeinheit groß genug gewählt wird, daß heißt  $m_n = n^\gamma$  für ein genügend großes  $\gamma$ , dann folgt

$$\max_k \sup_{\tilde{G}_k} |\Delta_n(\mathbf{y}) - \Delta_n(\mathbf{y}_{M_k})| = C n^{-\gamma} \lambda_n^{-4}.$$

**q.e.d.**

Als nächstes wird der Ausdruck  $\max_k |\Delta_n(\mathbf{y}_{M_k})|$  in (3.18) abgeschätzt. Zur Erinnerung,

$$\begin{aligned} & \max_k |\Delta_n(\mathbf{y}_{M_k})| \\ &= \max_k \left| \frac{1}{n} \sum_{j=1}^n \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}_{M_k}} \left( \frac{e^{it \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j})}{M(\mathbf{t})} - \Phi_f(\mathbf{t}) \right) \Phi_K(\lambda_n \mathbf{t}) dt \right|. \end{aligned}$$

Gemäß dem Beweis zu Satz 2.1 gilt folgende Beziehung

$$E[e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j}) \frac{1}{M(\mathbf{t})}] = \Phi_f(\mathbf{t}).$$

Daraus folgt, daß für alle  $\mathbf{y}$

$$E[D_j(\mathbf{y})] = 0,$$

und damit

$$E[\Delta_n(\mathbf{y}_{M_k})] = 0.$$

Der folgende Satz enthält eine Aussage über das Konvergenzverhalten der Zufallsvariablen  $\max_k |\Delta_n(\mathbf{y}_{M_k})|$ .

**Satz 3.9** Gegeben sei ein äquidistantes Gitter  $\tilde{\mathcal{G}}$  über  $\mathcal{B}$  mit Segmentlänge  $\frac{2}{m_n}$ .  $(m_n)_{n \in \mathbb{N}}$  und  $(\epsilon_n)_{n \in \mathbb{N}}$  seien Folgen mit  $m_n \rightarrow \infty$ ,  $\epsilon_n \rightarrow 0$  für  $n \rightarrow \infty$ . Es bezeichne  $K_n$  die Anzahl der Gittersegmente  $\tilde{G}_k$  in  $\tilde{\mathcal{G}}$ . Es sei  $k \in \{1, \dots, K_n\}$  und  $K_n < m_n^2$ . Es seien  $D_j(\mathbf{y}_{M_k})$  und  $\Delta_n(\mathbf{y}_{M_k})$  definiert über

$$D_j(\mathbf{y}_{M_k}) := \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}_{M_k}} \left( \frac{e^{-it' \mathbf{x}_j G(\mathbf{t}, \Theta_{\varphi_j})}}{M(\mathbf{t})} - \Phi_f(\mathbf{t}) \right) \Phi_K(\lambda_n \mathbf{t}) dt,$$

$$\Delta_n(\mathbf{y}_{M_k}) := \frac{1}{n} \sum_{j=1}^n D_j(\mathbf{y}_{M_k}).$$

Unter der Bedingung

$$\sum_{n=1}^{\infty} K_n e^{-\frac{1}{2}(n\epsilon_n^2 B^2 \lambda_n^3)} < \infty,$$

gilt, daß

$$\max_k |\Delta_n(\mathbf{y}_{M_k})| = O(\epsilon_n) \quad \text{mit Wahrscheinlichkeit 1.}$$

Der Beweis zu Satz 3.9 erfolgt unter Verwendung des Lemma's von Borel-Cantelli und der Bernstein Ungleichung. Diese sind in Lemma 3.1 und Lemma 3.2 angegeben.

**Lemma 3.1 (Lemma von Borel–Cantelli)**

Es sei  $A_1, A_2, \dots$  eine unendliche Folge von Ereignissen. Das Ereignis  $\{A_i \text{ unendlich oft}\}$  ist definiert als

$$\{A_i \text{ unendlich oft}\} := A^\infty = \bigcap_{j=1}^{\infty} \bigcup_{i=j}^{\infty} A_i.$$

Wenn

$$\sum_{i=1}^{\infty} \Pr(A_i) < \infty,$$

dann gilt

$$\Pr(A^\infty) = 0.$$

**Lemma 3.2 (Bernstein Ungleichung)**

Es seien  $X_1, \dots, X_n$  unabhängig identisch verteilte Zufallsvariablen mit  $E[X_j] = 0$  und  $\text{var}(X_j) = \sigma^2 < \infty$ . Es sei weiterhin  $\|X_j\| \leq D$  und  $n\sigma^2 \leq V$ ,  $j = 1, \dots, n$ . Dann gilt  $\forall \epsilon > 0$ :

$$\Pr\left(\left|\sum_{j=1}^n X_j\right| > \epsilon\right) \leq 2 \exp\left(-\frac{1}{2} \frac{\epsilon^2}{V + \frac{1}{3} D\epsilon}\right).$$

Es wird im folgenden der Beweis von Satz 3.9 durchgeführt.

**Beweis** zu Satz 3.9: Nach Lemma 3.1 (Borel-Cantelli) folgt aus der Bedingung

$$\sum_{n=1}^{\infty} \Pr(\max_k |\Delta_n(\mathbf{y}_{M_k})| \geq B \epsilon_n) < \infty$$

daß

$$\max_k |\Delta_n(\mathbf{y}_{M_k})| = O(\epsilon_n) \quad \text{mit Wahrscheinlichkeit 1.}$$

Hierbei ist  $k = 1, \dots, K_n$  und  $K_n < m_n^2$ . Es gilt weiterhin, daß

$$\begin{aligned} \Pr(\max_k |\Delta_n(\mathbf{y}_{M_k})| \geq B \epsilon_n) &\leq \sum_{k=1}^{K_n} \Pr(|\Delta_n(\mathbf{y}_{M_k})| \geq B \epsilon_n) \\ &= \sum_{k=1}^{K_n} \Pr\left(\left|\frac{1}{n} \sum_{j=1}^n D_j(\mathbf{y}_{M_k})\right| \geq B \epsilon_n\right). \end{aligned}$$

Unter Anwendung der Bernstein-Ungleichung erhält man folgende Aussage:

$$\Pr\left(\left|\sum_{j=1}^n D_j(\mathbf{y}_{M_k})\right| \geq n B \epsilon_n\right) \leq 2 \exp\left(-\frac{1}{2} \frac{n^2 B^2 \epsilon_n^2}{V + \frac{1}{3} D \epsilon_n B n}\right). \quad (3.19)$$

Hierbei bezeichnet  $D$  den Betrag von  $D_j(\mathbf{y}_{M_k})$  und kann wie folgt abgeschätzt werden.

$$\begin{aligned} |D_j(\mathbf{y}_{M_k})| &= \left| \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-it' \mathbf{y}_{M_k}} \left( \frac{e^{it' \mathbf{x}_j G(\mathbf{t}, \Theta_{\varphi_j})}}{M(\mathbf{t})} - \Phi_f(\mathbf{t}) \right) \Phi_K(\lambda_n \mathbf{t}) d\mathbf{t} \right| \\ &\leq \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \left| \frac{e^{it' \mathbf{x}_j G(\mathbf{t}, \Theta_{\varphi_j})}}{M(\mathbf{t})} - \Phi_f(\mathbf{t}) \right| |\Phi_K(\lambda_n \mathbf{t})| d\mathbf{t} \\ &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \left| e^{it' \mathbf{x}_j G(\mathbf{t}, \Theta_{\varphi_j})} - \Phi_f(\mathbf{t}) M(\mathbf{t}) \right| \left| \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})} \right| d\mathbf{t} \end{aligned}$$

Es wird folgende Abschätzung verwendet

$$|e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j}) - \Phi_f(\mathbf{t}) M(\mathbf{t})| \leq |G(\mathbf{t}, \Theta_{\varphi_j})| + \left| \int_{\mathcal{B}} e^{it' \mathbf{y}} f(\mathbf{y}) d\mathbf{y} \right| |M(\mathbf{t})|.$$

Aus Definition 2.5 und Definition 2.6. folgt, daß  $|G(\mathbf{t}, \Theta_{\varphi_j})| \leq 1$  und  $|M(\mathbf{t})| \leq 1$ , und damit

$$|e^{it' \mathbf{x}_j} G(\mathbf{t}, \Theta_{\varphi_j}) - \Phi_f(\mathbf{t}) M(\mathbf{t})| \leq 2.$$

Im Beweis zu Satz 3.5 wurde gezeigt, daß  $|M(\mathbf{t})| \geq \|\mathbf{t}\|^{-1}$ . Es ergibt sich folgende Abschätzung für  $|D_j(\mathbf{y}_{M_k})|$  für alle  $j = 1, \dots, n$  und alle  $\mathbf{y}_{M_k}$ ,  $k = 1, \dots, K_n$ .

$$\begin{aligned} |D_j(\mathbf{y}_{M_k})| &\leq C \int_{\mathbb{R}^2} \left| \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})} \right| dt \\ &= C \int_0^{\frac{1}{\lambda_n}} \int_0^{\frac{1}{\lambda_n}} \left| \frac{1}{M(\mathbf{t})} \right| dt_1 dt_2 \\ &\leq C \frac{1}{\lambda_n^3}. \end{aligned}$$

Als nächstes soll eine obere Schranke für die Varianz von  $D_j(\mathbf{y}_{M_k})$  gefunden werden. Für  $j = 1, \dots, n$  gilt:

$$\begin{aligned} \text{var}(D_j(\mathbf{y}_{M_k})) &\leq E[D_j^2(\mathbf{y}_{M_k})] \\ &= \int_{\mathbb{R}^2} \int_0^\pi \left( \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{it' \mathbf{x}} e^{-it' \mathbf{y}_{M_k}} G(\mathbf{t}, \Theta_\varphi) \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})} dt \right)^2 w(\mathbf{x}|\varphi) \frac{1}{\pi} d\mathbf{x} d\varphi. \end{aligned}$$

Wie auch im Beweis zu Satz 3.2 sei angenommen, daß die Dichtefunktion  $w(\mathbf{x}|\varphi)$  nach oben für alle  $\mathbf{x}$ ,  $\varphi$  beschränkt ist. Weiterhin ist

$$\int_{\mathbb{R}^2} e^{it' \mathbf{x}} e^{-it' \mathbf{y}_{M_k}} G(\mathbf{t}, \Theta_\varphi) \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})} dt$$

die Fouriertransformation der Funktion

$$e^{-it' \mathbf{y}_{M_k}} G(\mathbf{t}, \Theta_\varphi) \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})}$$

betrachtet an der Stelle  $\mathbf{x}$ . Daher folgt mit Parseval's Gleichung aus Bemerkung 3.1

$$\int_{\mathbb{R}^2} \left( \int_{\mathbb{R}^2} e^{it' \mathbf{x}} e^{-it' \mathbf{y}_{M_k}} G(\mathbf{t}, \Theta_\varphi) \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})} dt \right)^2 d\mathbf{x} = \int_{\mathbb{R}^2} (e^{-it' \mathbf{y}_{M_k}} G(\mathbf{t}, \Theta_\varphi) \frac{\Phi_K(\lambda_n \mathbf{t})}{M(\mathbf{t})})^2 dt.$$

Eine obere Schranke für die Varianz von  $D_j(\mathbf{y}_{M_k})$  ist dann gegeben über

$$\begin{aligned} \text{var}(D_j(\mathbf{y}_{M_k})) &\leq C \int_0^\pi \int_{\mathbb{R}^2} \frac{1}{\pi} G(\mathbf{t}, \Theta_\varphi)^2 \frac{\Phi_K^2(\lambda_n \mathbf{t})}{M(\mathbf{t})^2} dt d\varphi \\ &= C \int_{\mathbb{R}^2} \frac{\int_0^\pi G(\mathbf{t}, \Theta_\varphi)^2 \frac{1}{\pi} d\varphi}{|M(\mathbf{t})|} \frac{\Phi_K^2(\lambda_n \mathbf{t})}{M(\mathbf{t})} dt. \end{aligned}$$

Wie im Beweis zu Satz 3.5 gezeigt, ist  $E[G(\mathbf{t}, \Theta_\varphi)^2] = |M(\mathbf{t})|$ . Es folgt für alle  $j = 1, \dots, n$

$$\text{var}(D_j(\mathbf{y}_{M_k})) \leq C \int_{\mathbb{R}^2} \frac{\Phi_K^2(\lambda_n \mathbf{t})}{M(\mathbf{t})} dt \leq C \frac{1}{\lambda_n^3}.$$

Damit gilt nun

$$D \leq C \frac{1}{\lambda_n^3},$$

$$V = C \frac{n}{\lambda_n^3}.$$

Unter Verwendung von (3.19) ergibt sich

$$\Pr\left(\left|\sum_{j=1}^n D_j(\mathbf{y}_{M_k})\right| \geq n B \epsilon_n\right) \leq 2 \exp\left(-\frac{1}{2} \left(\frac{n^2 B^2 \epsilon_n^2}{\frac{n}{\lambda_n^3} + \frac{1}{3} \frac{1}{\lambda_n^3} n \epsilon_n B}\right)\right).$$

Wenn  $B \epsilon_n$  kleiner ist als 3, wird der Nenner von  $\frac{n}{\lambda_n^3}$  dominiert. Es gilt dann

$$\Pr\left(\left|\sum_{j=1}^n D_j(\mathbf{y}_{M_k})\right| \geq n B \epsilon_n\right) \leq 2 \exp\left(-\left(\frac{1}{2} n \epsilon_n^2 B^2 \lambda_n^3\right)\right).$$

Daraus folgt

$$\begin{aligned}
\Pr(\max_k |\Delta_n(\mathbf{y}_{M_k})| \geq B \epsilon_n) &\leq \sum_{k=1}^{K_n} \Pr(|\frac{1}{n} \sum_{j=1}^n D_j(\mathbf{y}_{M_k})| \geq B \epsilon_n) \\
&\leq \sum_{k=1}^{K_n} 2 \exp\left(-\left(\frac{1}{2} n \epsilon_n^2 B^2 \lambda_n^3\right)\right) \\
&= K_n 2 \exp\left(-\left(\frac{1}{2} n \epsilon_n^2 B^2 \lambda_n^3\right)\right).
\end{aligned}$$

Für  $n \rightarrow \infty$  konvergiert die Anzahl der Gittersegmente  $K_n$  gegen  $\infty$ . Wenn  $B$ ,  $\epsilon_n$  und  $\lambda_n$  passend gewählt werden, sodaß

$$\sum_{n=1}^{\infty} K_n \exp\left(-\left(\frac{1}{2} n \epsilon_n^2 B^2 \lambda_n^3\right)\right) < \infty, \quad (3.20)$$

dann gilt auch

$$\sum_{n=1}^{\infty} \Pr(\max_k |\Delta_n(\mathbf{y}_{M_k})| \geq B \epsilon_n) < \infty. \quad (3.21)$$

Es folgt dann mit Borel–Cantelli, daß

$$\max_k |\Delta_n(\mathbf{y}_{M_k})| = O(\epsilon_n) \quad \text{mit Wahrscheinlichkeit 1,}$$

womit der Satz bewiesen ist.

**q.e.d.**

Satz 3.10 gibt die Wahl des Bandweiteparameters für die optimale Konvergenzrate unter der Supremum–Norm an.

**Satz 3.10** *Für den modifizierten Dekonvolutions–Kernschätzer  $\hat{f}_n(\mathbf{y})$  ist die optimale Bandweite unter der Supremum–Norm gegeben durch*

$$\lambda_n^{sup} = C \left( \frac{\ln(n)}{n} \right)^{\frac{1}{2\alpha+1}}.$$

*Damit ergibt sich als obere Schranke für  $\sup_{\mathcal{B}} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})|$*

$$\sup_{\mathcal{B}} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})| = O\left( n^{-\frac{\alpha-1}{2\alpha+1}} \ln(n)^{\frac{\alpha-1}{2\alpha+1}} \right) \quad \text{mit Wahrscheinlichkeit 1.}$$

**Beweis:** Zunächst wird unter Verwendung von Satz 3.8 und Satz 3.9. eine fast sichere Abschätzung für (3.18) hergeleitet. Die Voraussetzung von Satz 3.9,

$$\sum_{n=1}^{\infty} K_n \exp\left(-\frac{1}{2} n \epsilon_n^2 B^2 \lambda_n^3\right) < \infty, \quad (3.22)$$

ist erfüllt, wenn  $K_n$ ,  $\epsilon_n$ ,  $B$  und  $\lambda_n$  folgender Bedingung genügen

$$\lim_{n \rightarrow \infty} K_n \exp\left(-\frac{1}{2} n \epsilon_n^2 B^2 \lambda_n^3\right) = 0.$$

Daraus ergibt sich zunächst die folgende Beziehung zwischen  $\lambda_n$  und  $\epsilon_n$ .

$$\epsilon_n^2 \lambda_n^3 > O\left(\frac{1}{n}\right).$$

Man wähle

$$\epsilon_n^2 \lambda_n^3 = C \ln(n)/n.$$

Damit ist  $K_n \exp\left(-\frac{1}{2} n \epsilon_n^2 B^2 \lambda_n^3\right) = K_n n^{-C B^2}$ . Es gilt  $K_n = O(m_n^2)$ . Wenn das Gitter fein genug ist, d. h. wenn  $m_n$  schnell genug mit  $n$  anwächst, dann gilt  $K_n = O(n^{2\gamma})$ .  $\gamma$  werde so groß gewählt wie man es für ein hinreichend feines Gitter braucht. Für die getroffene Wahl von  $\gamma$  wird  $B$  groß genug gewählt, sodaß die Summe in (3.22) über  $n$  konvergiert. Damit ist die Voraussetzung von Satz 3.9 erfüllt und mit

$$\epsilon_n = \left(\frac{C \ln(n)}{n} \lambda_n^{-3}\right)^{\frac{1}{2}},$$

ergibt sich unter Hinzunahme von Satz 3.8 folgende Abschätzung für das Supremum in (3.18):

$$\sup_{\mathcal{B}} |\hat{f}_n(\mathbf{y}) - E[\hat{f}_n(\mathbf{y})]| = O\left(\frac{\lambda_n^{-4}}{n^\gamma}\right) + O\left(\frac{\sqrt{\ln(n)}}{\sqrt{n}} \lambda_n^{-3/2}\right) \quad \text{f. s.}$$

Für hinreichend feines Gitter, und damit hinreichend großes  $\gamma$ , ist der erste Term von geringerer Größenordnung in  $n$  als der zweite, und es gilt die folgende Annäherung für die Supremum-Norm in (3.16):

$$\sup_{\mathbf{y} \in \mathcal{B}} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})| \leq O(\lambda_n^{\alpha-1}) + O\left(\frac{\sqrt{\ln(n)}}{\sqrt{n}} \lambda_n^{-3/2}\right) \quad \text{f. s.}$$



Mit der Wahl

$$\lambda_n = C \left( \frac{\ln(n)}{n} \right)^{\frac{1}{2\alpha+1}}$$

ergibt sich die optimale obere Schranke

$$\sup_{\mathbf{y} \in \mathcal{B}} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})| = O \left( n^{-\frac{\alpha-1}{2\alpha+1}} \ln(n)^{\frac{\alpha-1}{2\alpha+1}} \right) \quad \text{mit Wahrscheinlichkeit 1.}$$

### 3.3 Anmerkungen zu Kapitel 3

In diesem Kapitel wurde die Konvergenzgeschwindigkeit des Rekonstruktionsfehlers  $|\hat{f}_n(\mathbf{y}) - f(\mathbf{y})|$  untersucht. Es wurden obere Schranken für die Konvergenzgeschwindigkeit unter dem Mittleren Quadratischen Fehler, dem Mittleren Integrierten Quadratischen Fehler und der Supremum-Norm hergeleitet. Bei einem Vergleich der Raten für die Supremum-Norm und den Mittleren Quadratischen Fehler ist zu bedenken, daß die Supremum-Norm den Betrag und der Mittlere Quadratische Fehler das Quadrat des Rekonstruktionsfehlers betrachtet. Außerdem bildet der Mittlere Quadratische Fehler den Erwartungswert über alle Stichproben während die Supremum-Norm von der jeweiligen Stichprobe abhängig ist. Bei Berechnung des Mittleren Quadratischen Fehler und des Mittleren Integrierten Quadratischen Fehler fallen daher Stichproben mit großem Rekonstruktionsfehler weniger ins Gewicht als bei der Berechnung der Supremum-Norm. Die Ergebnisse in Kapitel 3 zeigen, daß die Konvergenzrate von  $|\hat{f}_n(\mathbf{y}) - f(\mathbf{y})|$  unter der Supremum-Norm,  $r_1 = \frac{\alpha-1}{2\alpha+1}$ , der Hälfte der Konvergenzrate unter dem Mittleren Quadratischen Fehler,  $r_2 = \frac{2(\alpha-1)}{2\alpha+1}$ , entspricht, bei der Supremum-Norm fällt allerdings zusätzlich noch ein logarithmischer Faktor an.

Die hier erzielten oberen Schranken für die Konvergenzrate von  $\hat{f}_n(\mathbf{y})$  gegen  $f(\mathbf{y})$  werden in Kapitel 4 durch Minimax-Konvergenzraten ergänzt. Dies ermöglicht es, eine Aussage über die Optimalität des modifizierten Dekonvolutions-Kernschätzers im Hinblick auf die untersuchten Konvergenzraten zu treffen. Am Ende von Kapitel 4 werden dann die hier erzielten Ergebnisse in die Literatur eingeordnet.

# Kapitel 4

## Optimalitätseigenschaften

In Kapitel 3 wurden obere Schranken für die Konvergenzrate des modifizierten Dekonvolutions–Kernschätzers unter verschiedenen Fehlermaßen hergeleitet. In diesem Kapitel wird gezeigt, daß die vom modifizierten Dekonvolutions–Kernschätzer für das gegebene Schätzproblem erzielte Rate im Sinne des Minimax–Kriteriums optimal ist.

**Bezeichnung 4.1** *Die Klasse aller Schätzfunktionen für  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$  basierend auf  $n$  u.i.v. Beobachtungen wird mit  $\mathcal{T}_n$  bezeichnet. Ein Element aus  $\mathcal{T}_n$  wird mit  $\hat{T}_n$  bezeichnet.*

Ferner bezeichne  $L(\hat{T}_n, f)$  eines der Fehlermaße aus Definition 3.1. Die Konvergenzrate von

$$\inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} L(\hat{T}_n, f) \quad (4.1)$$

heißt Minimax–Konvergenzrate für die Schätzung von  $f$  aus  $\mathcal{H}^{\alpha,2}(\mathcal{B})$ . Es gibt keinen Schätzer, der gleichmäßig über  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  eine bessere Rate erzielt.

### 4.1 Ein Verfahren zur Bestimmung der Minimax Konvergenzrate

In diesem Abschnitt wird ein Verfahren zur Bestimmung der Minimax–Konvergenzrate eines Schätzproblem es eingeführt. Wir adaptieren das von Donoho und Liu (1991) [12] verwendete Verfahren für unsere Zwecke. Es sei  $t \in \mathbb{R}$ ,  $\mathbf{y}$  ein fest gewählter Punkt in  $\mathcal{B}$  und  $\Delta \in \mathbb{R}^+$ . Zwei Teilmengen von  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  werden betrachtet:

$$\mathcal{F}_{\leq t} := \{f \in \mathcal{H}^{\alpha,2}(\mathcal{B}), f(\mathbf{y}) \leq t\}, \quad (4.2)$$

$$\mathcal{F}_{\geq t+\Delta} := \{f \in \mathcal{H}^{\alpha,2}(\mathcal{B}), f(\mathbf{y}) \geq t + \Delta\}. \quad (4.3)$$

Daraus resultieren zwei Mengen von gemeinsamen Dichten

$$\mathcal{F}_{\leq t}^{(n)} := \{f^* : f^*(y_1, \dots, y_n) = f(y_1) \cdot \dots \cdot f(y_n), f \in \mathcal{F}_{\leq t}\},$$

$$\mathcal{F}_{\geq t+\Delta}^{(n)} := \{f^* : f^*(y_1, \dots, y_n) = f(y_1) \cdot \dots \cdot f(y_n), f \in \mathcal{F}_{\geq t+\Delta}\}.$$

Im folgenden bezeichnet  $\text{conv}(\mathcal{F}_{\leq t}^{(n)})$  die Menge der Linearkombinationen von Elementen aus  $\mathcal{F}_{\leq t}^{(n)}$  mit nicht-negativen, sich zu 1 addierenden Koeffizienten. Analog wird die Menge  $\text{conv}(\mathcal{F}_{\geq t+\Delta}^{(n)})$  definiert. Es sei  $f_0$  aus  $\mathcal{F}_{\leq t}$  und  $f_1$  aus  $\mathcal{F}_{\geq t+\Delta}$ . Gegeben sei eine Stichprobe vom Umfang  $n$ :  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ . Man führt eine Teststatistik  $\Phi$  für den Test

$$H_0 : f_0 \quad \text{gegen} \quad H_1 : f_1$$

ein. Falls  $\Phi = 0$  wird die Hypothese  $H_0$  angenommen, falls  $\Phi = 1$  wird  $H_0$  abgelehnt und stattdessen die Alternative  $H_1$  angenommen. Falls  $\Phi = p \in ]0, 1[$  wird  $H_0$  mit Wahrscheinlichkeit  $p$  abgelehnt. Es bezeichne von nun an  $E_{f_i}[\cdot]$  die Erwartungswertbildung unter der Bedingung  $f = f_i$ . Dann bezeichnet

$$\pi(f_0, f_1) := \inf_{0 \leq \Phi \leq 1} E_{f_0}[\Phi] + E_{f_1}[1 - \Phi]$$

die kleinstmögliche Summe der Wahrscheinlichkeiten von Fehler erster und zweiter Art für das Testproblem  $H_0$  gegen  $H_1$ . Die Größe  $\pi(f_0, f_1)$  kann als Maß für die Schwierigkeit zwischen  $f_0$  und  $f_1$  zu unterscheiden interpretiert werden. Mit Hilfe von  $\pi(\cdot, \cdot)$  führen wir die Minimax-Schwierigkeit eines Testproblem es ein.

**Definition 4.1 (Minimax-Schwierigkeit)**

Gegeben seien zwei Dichteklassen  $\mathcal{F}_{\leq t}$  und  $\mathcal{F}_{\geq t+\Delta}$ . Die Größe

$$\pi(\text{conv}(\mathcal{F}_{\leq t}^{(n)}), \text{conv}(\mathcal{F}_{\geq t+\Delta}^{(n)})) := \inf_{0 \leq \Phi \leq 1} \sup_{\substack{f_0 \in \mathcal{F}_{\leq t} \\ f_1 \in \mathcal{F}_{\geq t+\Delta}}} E_{f_0}[\Phi] + E_{f_1}[1 - \Phi]$$

heißt Minimax-Schwierigkeit für das Testproblem  $H_0 : f \in \mathcal{F}_{\leq t}$  gegen  $H_1 : f \in \mathcal{F}_{\geq t+\Delta}$  bei einer Stichprobe vom Umfang  $n$ .

Ferner benötigen wir

**Definition 4.2 (Maximale Affinität)**

Für  $\mathcal{F}_{\leq t}$  und  $\mathcal{F}_{\geq t+\Delta}$  heißt

$$A(n, \Delta, \mathbf{y}) := \sup_t \pi(\text{conv}(\mathcal{F}_{\leq t}^{(n)}), \text{conv}(\mathcal{F}_{\geq t+\Delta}^{(n)}))$$

die maximale Affinität des Testproblems  $H_0 : f \in \mathcal{F}_{\leq t}$  gegen  $H_1 : f \in \mathcal{F}_{\geq t+\Delta}$  bei einer Stichprobe vom Umfang  $n$ .

**Definition 4.3** Es sei  $a \in ]0, 1[$ , dann ist

$$\Delta_A(n, a, \mathbf{y}) := \sup\{\Delta, A(n, \Delta, \mathbf{y}) \geq a\}.$$

$\Delta_A(n, \cdot, \mathbf{y})$  ist die Umkehrfunktion von  $A(n, \cdot, \mathbf{y})$ . Falls für ein  $\Delta > 0$  und ein  $a \in ]0, 1[$  die Bedingung  $A(n, \Delta, \mathbf{y}) \geq a$  für alle  $n \geq n_0$  erfüllt ist, so bedeutet dies, daß für mindestens ein  $t$  die Minimax-Schwierigkeit  $\pi(\text{conv}(\mathcal{F}_{\leq t}^{(n)}), \text{conv}(\mathcal{F}_{\geq t+\Delta}^{(n)}))$  nicht gegen Null konvergiert. Somit ist es nicht möglich, asymptotisch eine fast sichere Entscheidung zwischen den Funktionenklassen  $\mathcal{F}_{\leq t}$  und  $\mathcal{F}_{\geq t+\Delta}$  zu erreichen. Bei gegebenem  $a$  ist  $\Delta_A(n, a, \mathbf{y})$  durch die Eigenschaften des Schätzproblems bestimmt.

Es wird im folgenden gezeigt, daß über die Größe  $\Delta_A(n, a, \mathbf{y})$  eine untere Schranke für die Minimax-Konvergenzrate des Schätzfehlers  $|\hat{T}_n(\mathbf{y}) - f(\mathbf{y})|$  gegeben ist. Zu diesem Zwecke sei ohne Beschränkung der Allgemeinheit angenommen, daß für ein fest gewähltes  $a \in ]0, 1[$  das Supremum über  $\Delta$  in Definition 4.3 angenommen wird, und daß für  $\Delta = \Delta_A$ , das Supremum über  $t$  in Definition 4.2 an der Stelle  $t = t_0$  auftritt. Es sei  $\hat{T}_n \in \mathcal{T}_n$  beliebig gewählt. Man teste die Hypothese  $H_0 : \mathcal{F}_{\leq t_0}$  gegen  $H_1 : \mathcal{F}_{\geq t_0 + \Delta_A(n, a, \mathbf{y})}$  basierend auf folgender Regel: für  $\hat{T}_n(\mathbf{y}) \leq t_0 + \frac{\Delta_A(n, a, \mathbf{y})}{2}$  entscheide für  $H_0$ , für  $\hat{T}_n(\mathbf{y}) > t_0 + \frac{\Delta_A(n, a, \mathbf{y})}{2}$  entscheide für  $H_1$ . Wegen Definition 4.1 bis 4.3 ist

$$\sup_{\substack{f_0 \in \mathcal{F}_{\leq t_0} \\ f_1 \in \mathcal{F}_{\geq t_0 + \Delta_A(n, a, \mathbf{y})}}} \Pr_{f_0}(\text{Ablehnung von } H_0) + \Pr_{f_1}(\text{Annahme von } H_0) \geq a. \quad (4.4)$$

Hierbei und im folgenden bezeichnet  $\Pr_{f_i}$  das Bestimmen einer Wahrscheinlichkeit unter der Bedingung  $f = f_i$ .

Daraus folgt

$$\sup_{\substack{f_0 \in \mathcal{F}_{\leq t_0} \\ f_1 \in \mathcal{F}_{\geq t_0 + \Delta_A(n, a, \mathbf{y})}}} \max [\Pr_{f_0}(\text{Ablehnung von } H_0), \Pr_{f_1}(\text{Annahme von } H_0)] \geq \frac{a}{2}. \quad (4.5)$$

Weiterhin gilt

$$\begin{aligned} \Pr_{f_0}(|\hat{T}_n(\mathbf{y}) - f_0(\mathbf{y})| \geq \frac{\Delta_A(n, a, \mathbf{y})}{2}) &\geq \Pr_{f_0}(\text{Ablehnung von } H_0), \\ \Pr_{f_1}(|\hat{T}_n(\mathbf{y}) - f_1(\mathbf{y})| \geq \frac{\Delta_A(n, a, \mathbf{y})}{2}) &\geq \Pr_{f_1}(\text{Annahme von } H_0). \end{aligned} \quad (4.6)$$

Aus (4.4) bis (4.6) wird die folgende Aussage abgeleitet.

$$\sup_{\substack{f_0 \in \mathcal{F}_{\leq t_0} \\ f_1 \in \mathcal{F}_{\geq t_0 + \Delta_A(n, a, \mathbf{y})}}} \max_{f_0, f_1} (\Pr_{f_i}(|\hat{T}_n(\mathbf{y}) - f_i(\mathbf{y})| \geq \frac{\Delta_A(n, a, \mathbf{y})}{2})) \geq \frac{a}{2}. \quad (4.7)$$

Hierbei ist  $a \in ]0, 1[$  fest gewählt. Da  $\hat{T}_n$  beliebig gewählt wurde und  $f_0, f_1 \in \mathcal{H}^{\alpha, 2}(\mathcal{B})$  ergibt sich schließlich folgendes Theorem (Theorem 2.1 in Donoho und Liu (1991) [12]).

**Theorem 4.1** *Es sei  $a \in ]0, 1[$  beliebig gewählt aber fest. Für*

$$\Delta_A(n, a, \mathbf{y}) := \sup\{\Delta, A(n, \Delta, \mathbf{y}) \geq a\},$$

*ist*

$$\inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha, 2}(\mathcal{B})} \Pr_f(|\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq \frac{\Delta_A(n, a, \mathbf{y})}{2}) \geq \frac{a}{2}.$$

**Beweis:** Siehe oben, (4.4) bis (4.7).

Aus Theorem 4.1 erhält man folgendes

**Korollar 4.1** *Es sei  $a \in ]0, 1[$  beliebig gewählt aber fest. Für*

$$\Delta_A(n, a, \mathbf{y}) := \sup\{\Delta, A(n, a, \mathbf{y}) \geq a\}$$

*ist*

$$\inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha, 2}(\mathcal{B})} E_f[|\hat{T}_n(\mathbf{y}) - f(\mathbf{y})|^2] \geq \frac{\Delta_A^2(n, a, \mathbf{y})}{4} \frac{a}{2}.$$

**Beweis:** Wir definieren die Zufallsvariable  $S := |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})|$ . Wir führen den Beweis von Korollar 4.1. für den Fall, daß  $S$  eine Dichtefunktion  $w_S$  besitzt. Andernfalls argumentiert man analog. Dann ist

$$\begin{aligned}
\sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} E[|\hat{T}_n(\mathbf{y}) - f(\mathbf{y})|^2] &= \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \int_0^\infty s^2 w_S(s) ds \\
&\geq \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \int_{\frac{\Delta_A(n,a,\mathbf{y})}{2}}^\infty s^2 w_S(s) ds \\
&\geq \frac{(\Delta_A(n,a,\mathbf{y}))^2}{4} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr(S \geq \frac{\Delta_A(n,a,\mathbf{y})}{2}) \\
&\geq \frac{(\Delta_A(n,a,\mathbf{y}))^2}{4} \frac{a}{2}.
\end{aligned}$$

Die letzte Ungleichung folgt aus Theorem 4.1. Dies gilt für alle  $\hat{T}_n \in \mathcal{T}_n$ . Daraus folgt die Behauptung.

**q.e.d.**

Aus Theorem 4.1 und Korollar 4.1 wird ersichtlich, wie aus  $\Delta_A(n, a, \mathbf{y})$  eine untere Schranke für die Minimax-Konvergenzrate hergeleitet werden kann. Zu diesem Zweck muß zunächst eine Abschätzung für  $\Delta_A(n, a, \mathbf{y})$  gefunden werden. Es sei

$$A_2(n, \Delta, \mathbf{y}) := \sup_t \pi(\mathcal{F}_{\leq t}^{(n)}, \mathcal{F}_{\geq t+\Delta}^{(n)})$$

wobei

$$\pi(\mathcal{F}_{\leq t}^{(n)}, \mathcal{F}_{\geq t+\Delta}^{(n)}) := \sup_{\substack{f_0 \in \mathcal{F}_{\leq t} \\ f_1 \in \mathcal{F}_{\geq t+\Delta}}} \inf_{\Phi} E_{f_0}[\Phi] + E_{f_1}[1 - \Phi].$$

Aus der Beziehung

$$\sup_{\substack{f_0 \in \mathcal{F}_{\leq t} \\ f_1 \in \mathcal{F}_{\geq t+\Delta}}} E_{f_0}[\Phi] + E_{f_1}[1 - \Phi] \geq \sup_{\substack{f_0 \in \mathcal{F}_{\leq t} \\ f_1 \in \mathcal{F}_{\geq t+\Delta}}} \inf_{\Phi} E_{f_0}[\Phi] + E_{f_1}[1 - \Phi]$$

folgt

$$\inf_{\Phi} \sup_{\substack{f_0 \in \mathcal{F}_{\leq t} \\ f_1 \in \mathcal{F}_{\geq t+\Delta}}} E_{f_0}[\Phi] + E_{f_1}[1 - \Phi] \geq \sup_{\substack{f_0 \in \mathcal{F}_{\leq t} \\ f_1 \in \mathcal{F}_{\geq t+\Delta}}} \inf_{\Phi} E_{f_0}[\Phi] + E_{f_1}[1 - \Phi]$$

und damit

$$\pi(\text{conv}(\mathcal{F}_{\leq t}^{(n)}), \text{conv}(\mathcal{F}_{\geq t+\Delta}^{(n)})) \geq \pi(\mathcal{F}_{\leq t}^{(n)}, \mathcal{F}_{\geq t+\Delta}^{(n)}).$$

Daraus ergibt sich die Ungleichung

$$A(n, \Delta, \mathbf{y}) \geq A_2(n, \Delta, \mathbf{y}).$$

Auch zu  $A_2(n, \cdot, \mathbf{y})$  führen wir die Umkehrfunktion ein:

$$\Delta_2(n, a, \mathbf{y}) := \sup\{\Delta, A_2(n, \Delta, \mathbf{y}) \geq a\}.$$

$\Delta_2(n, a, \mathbf{y})$  stellt eine untere Schranke für  $\Delta_A(n, a, \mathbf{y})$  dar und kann in folgender Weise dargestellt werden:

$$\Delta_2(n, a, \mathbf{y}) := \sup_{\mathcal{H}^{\alpha, 2}(\mathcal{B})} \{|f_1(\mathbf{y}) - f_0(\mathbf{y})|, \pi(f_1^{(n)}, f_0^{(n)}) \geq a\}.$$

Um nun eine untere Schranke für  $\Delta_2(n, a, \mathbf{y})$  zu bestimmen, werden folgende Begriffe eingeführt.

**Definition 4.4** Für die Dichten  $f_0, f_1$  heißt

$$H(f_0, f_1) := \left( \int_{\mathcal{B}} \left( \sqrt{f_0(\mathbf{y})} - \sqrt{f_1(\mathbf{y})} \right)^2 d\mathbf{y} \right)^{\frac{1}{2}}$$

Hellinger-Abstand zwischen  $f_0$  und  $f_1$  und

$$\chi^2(f_0, f_1) = \int_{\mathcal{B}} \frac{(f_0(\mathbf{y}) - f_1(\mathbf{y}))^2}{f_0(\mathbf{y})} d\mathbf{y}$$

heißt  $\chi^2$ -Abstand zwischen  $f_0$  und  $f_1$ .

Nach LeCam (1985) [33] besteht zwischen diesen Abstands Begriffen die Beziehung

$$\chi^2(f_0, f_1) \geq 2H^2(f_0, f_1). \quad (4.8)$$

**Definition 4.5 (Hellinger–Affinität)**

Es seien  $w$  und  $v$  Dichtefunktionen bezüglich eines Maßes  $\mu$ . Die Größe

$$\rho(w, v) = \int \sqrt{w} \sqrt{v} d\mu$$

heißt Hellinger–Affinität von  $w$  und  $v$ .

**Lemma 4.1** Die Hellinger–Affinität genügt folgenden Beziehungen:

1.  $\rho^2(w, v) \leq \pi(w, v) (2 - \pi(w, v))$ .
2.  $\rho(w^{(n)}, v^{(n)}) = \rho(w, v)^n$ .
3.  $\rho(w, v) = \frac{1}{2}(2 - H^2(w, v))$ .

**Beweis:** Siehe Donoho und Liu (1991), [12].

**Definition 4.6 (Stetigkeitsmodul)**

Sei  $T : \mathcal{H}^{\alpha,2}(\mathcal{B}) \rightarrow \mathbb{R}$  ein Funktional. Die Funktion

$$\omega(\epsilon) := \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \{|T(f_1) - T(f_0)| : H(f_1, f_0) \leq \epsilon\}$$

heißt Stetigkeitsmodul von  $T$  über  $\mathcal{H}^{\alpha,2}(\mathcal{B})$ .

Es bezeichne

$$\mathcal{M}_1 := \{f_0, f_1 \in \mathcal{H}^{\alpha,2}(\mathcal{B}), \pi(f_0^{(n)}, f_1^{(n)}) \geq a\}.$$

Mit  $\pi(f_0^{(n)}, f_1^{(n)}) \geq a$  gilt auch

$$2 - \pi(f_0^{(n)}, f_1^{(n)}) \leq 2 - a, \tag{4.9}$$

und

$$\pi(f_0^{(n)}, f_1^{(n)})(2 - \pi(f_0^{(n)}, f_1^{(n)})) \geq a(2 - \pi(f_0^{(n)}, f_1^{(n)})). \tag{4.10}$$

Daher kann  $\mathcal{M}_1$  alternativ dargestellt werden als

$$\mathcal{M}_1 = \{f_0, f_1 \in \mathcal{H}^{\alpha,2}(\mathcal{B}), \pi(f_0^{(n)}, f_1^{(n)})(2 - \pi(f_0^{(n)}, f_1^{(n)})) \geq a(2 - \pi(f_0^{(n)}, f_1^{(n)}))\}.$$



Aus (4.9) und (4.10) ist ersichtlich, daß die Menge

$$\mathcal{M}_2 := \{f_0, f_1 \in \mathcal{H}^{\alpha,2}(\mathcal{B}), \pi(f_0^{(n)}, f_1^{(n)})(2 - \pi(f_0^{(n)}, f_1^{(n)})) \geq a(2 - a)\}$$

eine Teilmenge von  $\mathcal{M}_1$  ist. Weiterhin gilt wegen

$$\rho^2(f_0^{(n)}, f_1^{(n)}) \leq \pi(f_0^{(n)}, f_1^{(n)}) (2 - \pi(f_0^{(n)}, f_1^{(n)}))$$

daß auch  $\mathcal{M}_2$  eine Obermenge ist von

$$\mathcal{M}_3 := \{f_0, f_1 \in \mathcal{H}^{\alpha,2}(\mathcal{B}), \rho^2(f_0^{(n)}, f_1^{(n)}) \geq a(2 - a)\}.$$

Da  $\Delta_2$  definiert ist über

$$\Delta_2(n, a, \mathbf{y}) := \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \{|f_1(\mathbf{y}) - f_0(\mathbf{y})|, \pi(f_1^{(n)}, f_0^{(n)}) \geq a\}$$

folgt somit, daß die Größe

$$\sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \{|f_0(\mathbf{y}) - f_1(\mathbf{y})|, \rho^2(f_0^{(n)}, f_1^{(n)}) \geq a(2 - a)\}$$

eine untere Schranke für  $\Delta_2(n, a, \mathbf{y})$  darstellt. Unter Verwendung von Lemma 4.1, 2. und 3., folgt aus

$$\rho^2(f_0^{(n)}, f_1^{(n)}) \geq a(2 - a)$$

die Bedingung

$$\left(\frac{1}{2}(2 - H^2(f_0, f_1))\right)^{2n} \geq a(2 - a),$$

und dies ist äquivalent zu

$$H^2(f_0, f_1) \leq 2(1 - (a(2 - a))^{\frac{1}{2n}}).$$

Unter Zuhilfenahme des sich anschließenden Lemmas ist eine weitere Umformung möglich.

**Lemma 4.2** Gegeben sei  $\epsilon_0 > 0$ . Es sei  $a \in ]0, 1[$  beliebig gewählt aber fest und es gebe ein  $n_0$  sodas  $\frac{|\log(a)|}{n} < \epsilon_0$  für alle  $n \geq n_0$ . Dann gilt für eine endliche, positive Konstante  $c$  und für  $n \geq n_0$

$$(1 - (a(2-a))^{\frac{1}{2n}}) > \frac{c^2}{2} \frac{|\log a|}{n}.$$

**Beweis:** Siehe Donoho und Liu (1991), [12].

Unter den Voraussetzungen von Lemma 4.2 gilt, daß die Menge

$$\mathcal{M}_4 := \{f_0, f_1 \in \mathcal{H}^{\alpha,2}(\mathcal{B}), H^2(f_0, f_1) \leq C/n\}$$

eine Teilmenge von  $\mathcal{M}_3$  ist, wobei  $C$  von der Form  $c|\log(a)|$  für  $0 < c < \infty$  ist. Dies gilt bei gegebenem  $a$  für  $n$  groß genug. Somit stellt das Stetigkeitsmodul

$$\omega(C/n) := \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \{|f_0(\mathbf{y}) - f_1(\mathbf{y})|, H^2(f_0, f_1) \leq C/n\}$$

eine untere Schranke für  $\Delta_2(n, a, \mathbf{y})$  dar. Wegen Gleichung (4.8) gilt die Implikation

$$\chi^2(f_0, f_1) \leq C/n \quad \Rightarrow \quad 2H^2(f_0, f_1) \leq C/n.$$

Damit ist die Menge

$$\mathcal{M}_5 := \{f \in \mathcal{H}^{\alpha,2}(\mathcal{B}), \chi^2(f_0, f_1) \leq C/n\}$$

eine Teilmenge von  $\mathcal{M}_4$  und die Größe

$$\Delta_2^*(n, a, \mathbf{y}) := \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \{|f_0(\mathbf{y}) - f_1(\mathbf{y})|, \chi^2(f_0, f_1) \leq C/n\} \quad (4.11)$$

stellt eine untere Schranke für  $\omega(C/n)$  und damit für  $\Delta_2(n, a, \mathbf{y})$  dar. Mit Theorem 4.1 folgt dann schließlich

**Theorem 4.2** Gegeben sei  $\epsilon_0 > 0$ . Es sei  $a \in ]0, 1[$  beliebig gewählt aber fest und es gebe ein  $n_0$  sodaß  $\frac{|\log(a)|}{n} < \epsilon_0$  für alle  $n \geq n_0$ . Mit

$$\Delta_2^*(n, a, \mathbf{y}) := \sup_{\mathcal{H}^{\alpha, 2}(\mathcal{B})} \{|f_0(\mathbf{y}) - f_1(\mathbf{y})|, \chi^2(f_0, f_1) \leq C/n\}$$

gilt für  $n \geq n_0$  die Bedingung

$$\inf_{T_n} \sup_{\mathcal{H}^{\alpha, 2}(\mathcal{B})} \Pr_f(|\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq \frac{\Delta_2^*(n, a, \mathbf{y})}{2}) \geq \frac{a}{2}.$$

Hierbei ist  $C$  von der Form  $c |\log(a)|$  mit  $0 < c < \infty$ .

**Beweis:** Siehe die obige Herleitung des Theorems.

Die untere Schranke  $\Delta_2^*(n, a, \mathbf{y})$  wurde für festes  $\mathbf{y}$  bestimmt und gilt bei gewähltem  $a$  für  $n$  groß genug.

## 4.2 Anwendung auf ToFPET Dichteschätzung

Die in Abschnitt 4.1 vorgestellte Methodik wird nun verwendet, um untere Schranken für den Mittleren Quadratischen Fehler (MSE), den Mittleren Integrierten Quadratischen Fehler (MISE) und die Supremum-Norm zu berechnen. Um Theorem 4.2 auf das hier vorliegende ToFPET-Schätzproblem anwenden zu können, ist zunächst eine Adaption auf den Fall indirekter Beobachtungen mit Flugzeitinformation nötig. Seien  $f_0 \in \mathcal{F}_{\leq t}$  und  $f_1 \in \mathcal{F}_{\geq t+\Delta}$  zwei Emissionsdichten. Die zugehörigen Detektionsdichten  $g_0$  und  $g_1$  ergeben sich aus

$$\begin{aligned} g_0(\mathbf{x}, \varphi) &:= \int_{-\infty}^{\infty} f_0(\mathbf{x} - r\Theta_\varphi)w(r)dr, \\ g_1(\mathbf{x}, \varphi) &:= \int_{-\infty}^{\infty} f_1(\mathbf{x} - r\Theta_\varphi)w(r)dr. \end{aligned} \tag{4.12}$$

Die Minimax-Schwierigkeit für das Testproblem  $H_0 : f \in \mathcal{F}_{\leq t}$  gegen  $H_1 : f \in \mathcal{F}_{\geq t+\Delta}$  bei einer Stichprobe  $(\mathbf{X}_1, \varphi_1), \dots, (\mathbf{X}_n, \varphi_n)$  ist wie zuvor definiert über

$$\pi(\text{conv}(\mathcal{F}_{\leq t}^{(n)}), \text{conv}(\mathcal{F}_{\geq t+\Delta}^{(n)})) := \inf_{0 \leq \Phi \leq 1} \sup_{\substack{f_0 \in \mathcal{F}_{\leq t} \\ f_1 \in \mathcal{F}_{\geq t+\Delta}}} E_{f_0}[\Phi] + E_{f_1}[1 - \Phi],$$

wobei die Teststatistik  $\Phi$  aus  $(\mathbf{X}_j, \varphi_j)$ ,  $j = 1, \dots, n$  erzeugt wird. Analog zu Definition 4.2 und 4.3 definiert man die maximale Affinität

$$A(n, \Delta, \mathbf{y}) := \sup_t \pi(\text{conv}(\mathcal{F}_{\leq t}^{(n)}), \text{conv}(\mathcal{F}_{\geq t+\Delta}^{(n)})).$$

Wenn wie im vorausgehenden Abschnitt für ein fest gewähltes  $a \in ]0, 1[$

$$\Delta_A(n, a, \mathbf{y}) := \sup\{\Delta, A(n, \Delta, \mathbf{y}) \geq a\}$$

so ist

$$\inf_{T_n} \sup_{\mathcal{H}^{\alpha, 2}(\mathcal{B})} \Pr_f(|\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq \frac{\Delta_A(n, a, \mathbf{y})}{2}) \geq \frac{a}{2}. \quad (4.13)$$

Hierbei bezeichnet  $\hat{T}_n(\mathbf{y})$  eine Schätzfunktion für  $f(\mathbf{y})$  basierend auf einer Stichprobe  $(\mathbf{X}_1, \varphi_1), \dots, (\mathbf{X}_n, \varphi_n)$  aus der Detektionsdichte  $g(\mathbf{x}, \varphi)$ . Aus Abschnitt 4.1 ergibt sich eine untere Schranke für  $\Delta_A$ :

$$\begin{aligned} \Delta_A(n, a, \mathbf{y}) &\geq \Delta_2(n, a, \mathbf{y}) \\ &:= \sup_{\mathcal{H}^{\alpha, 2}(\mathcal{B})} \{|f_0(\mathbf{y}) - f_1(\mathbf{y})|, \inf_{0 \leq \Phi \leq 1} E_{f_0}[\Phi] + E_{f_1}[1 - \Phi] \geq a\}. \end{aligned}$$

Es gelten die Voraussetzungen von Theorem 4.2. Bei gegebenem  $a$  wird für  $n$  groß genug  $\Delta_2(n, a, \mathbf{y})$  analog zur Vorgehensweise in Abschnitt 4.1 nach unten abgeschätzt durch

$$\Delta_2(n, a, \mathbf{y}) \geq \Delta_2^*(n, a, \mathbf{y}) := \sup_{\mathcal{H}^{\alpha, 2}(\mathcal{B})} \{|f_0(\mathbf{y}) - f_1(\mathbf{y})|, \chi^2(g_0, g_1) \leq C/n\}. \quad (4.14)$$

Hierbei ist  $C$  eine Konstante der Form  $c |\log(a)|$  für ein positives, endliches  $c$ . Im folgenden wird (4.14) verwendet, um eine untere Schranke des Mittleren Quadratischen Fehlers, des Mittleren Integrierten Quadratischen Fehlers und der Supremum-Norm für das ToFPET-Schätzproblem zu ermitteln.

### 4.3 Untere Schranke für den MSE

Wegen (4.14) und Korollar 4.1 ist zunächst

$$\inf_{T_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} E_f[|\hat{T}_n(\mathbf{y}) - f(\mathbf{y})|^2] \geq C \frac{(\Delta_2^*(n,a,\mathbf{y}))^2}{4}.$$

Im folgenden wird die Größe  $\Delta_2^*(n, a, \mathbf{y})$  bestimmt. Dazu werden zwei Repräsentanten aus  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  gewählt. Es sei  $\mathbf{y}_M \in \mathcal{B}$  fest aber beliebig.

$$f_2(\mathbf{y}) = \begin{cases} \frac{1}{\pi} & \text{für } \|\mathbf{y}\| \leq 1 \\ 0 & \text{sonst.} \end{cases}$$

$$f_3(\mathbf{y}) = f_2(\mathbf{y}) + \delta^2 H(m_n(\mathbf{y} - \mathbf{y}_M)).$$

Hierbei sei  $\delta$  eine nicht-negative Größe und  $(m_n)_{n \in \mathbb{N}}$  eine Folge positiver Zahlen mit  $m_n \rightarrow \infty$  für  $n \rightarrow \infty$ . Die Funktion  $H(\mathbf{y})$  erfülle folgende Bedingungen:

- i)  $H(\mathbf{y}) \in \mathcal{H}^{\alpha,2}(\mathcal{B})$ ,
- ii)  $H(0,0) \neq 0$ ,
- iii)  $|H(\mathbf{y})| \leq C$ ,
- iv)  $H(\mathbf{y}) \equiv 0$  wenn  $\|\mathbf{y}\| > 1$ ,
- v)  $\int_{\mathcal{B}} H(\mathbf{y}) d\mathbf{y} = 0$ .

Die Funktion  $f_2$  ist in  $\mathcal{H}^{\alpha,2}(\mathcal{B})$ . Damit auch  $f_3(\mathbf{y})$  in  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  liegt, müssen  $\delta$  und  $m_n$  so gewählt werden, daß

$$\|\delta^2 H(m_n(\mathbf{y} - \mathbf{y}_M))\|_{\mathcal{H}^{\alpha,2}(\mathcal{B})} < \infty.$$

**Lemma 4.3** Wenn  $\delta, m_n > 0$  der Bedingung

$$\delta \leq C m_n^{-\frac{(\alpha-1)}{2}}$$

genügen, dann ist  $\|\delta^2 H(m_n(\mathbf{y} - \mathbf{y}_M))\|_{\mathcal{H}^{\alpha,2}(\mathcal{B})} < \infty$ .

**Beweis:** Die Bedingung

$$\|\delta^2 H(m_n(\mathbf{y} - \mathbf{y}_M))\|_{\mathcal{H}^{\alpha,2}(\mathcal{B})} < \infty$$

ist äquivalent mit

$$\int_{\mathbb{R}^2} (1 + \|\mathbf{t}\|^2)^\alpha \delta^4 \frac{1}{m_n^4} |\Phi_H(\frac{\mathbf{t}}{m_n})|^2 d\mathbf{t} < \infty. \quad (4.15)$$

Die Substitution  $\mathbf{u} = \frac{\mathbf{t}}{m_n}$  führt zu

$$\frac{\delta^4}{m_n^4} \int_{\mathbb{R}^2} |\Phi_H(\mathbf{u})|^2 (1 + m_n^2 \|\mathbf{u}\|^2)^\alpha m_n^2 d\mathbf{u} < \infty$$

und wegen  $H(\mathbf{y}) \in \mathcal{H}^{\alpha,2}(\mathcal{B})$  gilt letzteres genau dann, wenn

$$\delta^4 m_n^{2\alpha-2} \leq C < \infty.$$

Daraus folgt die Bedingung

$$\delta \leq C m_n^{-\frac{(\alpha-1)}{2}}.$$

**q.e.d.**

Die mit  $f_2$  und  $f_3$  assoziierten Detektionssdichten sind

$$\begin{aligned} g_2(\mathbf{x}, \varphi) &:= \int_{\mathbb{R}} f_2(\mathbf{x} - r \Theta_\varphi) w(r) dr. \\ g_3(\mathbf{x}, \varphi) &:= \int_{\mathbb{R}} (f_2(\mathbf{x} - r \Theta_\varphi) + \delta^2 H(m_n(\mathbf{x} - \mathbf{y}_M - r \Theta_\varphi)) w(r) dr. \end{aligned}$$

Es folgt dann aus (4.14), daß an der Stelle  $\mathbf{y} = \mathbf{y}_M \in \mathcal{B}$

$$\begin{aligned} \Delta_2^*(n, a, \mathbf{y}_M) &:= \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \{|f_0(\mathbf{y}_M) - f_1(\mathbf{y}_M)|, \chi^2(g_0, g_1) \leq C/n\} \\ &\geq \sup_{\delta} \{|f_2(\mathbf{y}_M) - f_3(\mathbf{y}_M)|, \chi^2(g_2, g_3) \leq C/n\}. \end{aligned}$$

Also,

$$\Delta_2^*(n, a, \mathbf{y}_M) \geq \sup_{\delta} \{|\delta^2 H(\mathbf{0})|, \chi^2(g_2, g_3) \leq C/n\}. \quad (4.16)$$

Dabei ist  $C$  eine Konstante, welche von  $a$  abhängt. Eine Abschätzung für  $\Delta_2^*(n, a, \mathbf{y}_M)$  enthält der folgende

**Satz 4.1** Es sei  $\Delta_2^*(n, a, \mathbf{y}_M) := \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \{|f_0(\mathbf{y}_M) - f_1(\mathbf{y}_M)|, \chi^2(g_0, g_1) \leq C/n\}$ .

Für alle  $\mathbf{y}_M \in \mathcal{B}$  gilt

$$\Delta_2^*(n, a, \mathbf{y}_M) \geq C n^{-\frac{(\alpha-1)}{2\alpha+1}}.$$

**Beweis:**

$$\begin{aligned} \Delta_2^*(n, a, \mathbf{y}_M) &\geq \sup_{\delta} \{|f_2(\mathbf{y}_M) - f_3(\mathbf{y}_M)|, \chi^2(g_2, g_3) \leq C/n\} \\ &= \sup_{\delta} \{|\delta^2 H(\mathbf{0})|, \chi^2(g_2, g_3) \leq C/n\}. \end{aligned}$$

**1. Fall:** Es sei  $\mathbf{y}_M = \mathbf{0} := (0, 0)$ . Dann ist

$$f_3(\mathbf{y}) = f_2(\mathbf{y}) + \delta^2 H(m_n \mathbf{y}).$$

Es muß gelten

$$\chi^2(g_2, g_3) \leq C/n. \quad (4.17)$$

Wir machen, wie in Kapitel 2 motiviert, die Annahme, daß bei der Vorbereitung des Datensatzes alle Datenpunkte ausgesondert werden, welche nicht in  $\mathcal{B}$  liegen. Wir beschränken uns also auf den Fall  $\mathbf{X} \in \mathcal{B}$ . Der  $\chi^2$ -Abstand zwischen  $g_2$  und  $g_3$  ist dann

$$\begin{aligned} \chi^2(g_2, g_3) &= \int_{\mathcal{B}} \int_0^{\pi} \frac{\left( \int_{\mathbb{R}} (f_2(\mathbf{x}-r\Theta_\varphi) - f_3(\mathbf{x}-r\Theta_\varphi)) w(r) dr \right)^2}{\int_{\mathbb{R}} f_2(\mathbf{x}-r\Theta_\varphi) w(r) dr} d\varphi d\mathbf{x} \\ &= \int_{\mathcal{B}} \int_0^{\pi} \frac{\left( \int_{\mathbb{R}} \delta^2 H(m_n(\mathbf{x}-r\Theta_\varphi)) w(r) dr \right)^2}{\int_{\mathbb{R}} f_2(\mathbf{x}-r\Theta_\varphi) w(r) dr} d\varphi d\mathbf{x}. \end{aligned} \quad (4.18)$$

Wir betrachten zunächst das Integral  $\left| \int_{\mathbb{R}} H(m_n(\mathbf{x}-r\Theta_\varphi)) w(r) dr \right|$ . Aus den Eigenschaften der Funktion  $H$  folgt, daß  $H(m_n(\mathbf{x}-r\Theta_\varphi)) = 0$  falls

$$\|m_n(\mathbf{x}-r\Theta_\varphi)\| > 1.$$

Wir bezeichnen mit  $U_\gamma(\mathbf{x})$ ,  $\gamma > 0$  alle Punkte  $\mathbf{x}^* \in \mathcal{B}$  mit  $\|\mathbf{x}^* - \mathbf{x}\| \leq \gamma$ . Obige Bedingung gilt, wenn

$$r \Theta_\varphi \notin U_{\frac{1}{m_n}}(\mathbf{x}).$$

Damit  $r \Theta_\varphi \in U_{\frac{1}{m_n}}(\mathbf{x})$  ist, muß erfüllt sein:

a) eine Bedingung an  $\varphi$  für gegebenes  $\mathbf{x}$ :

$$\|\mathbf{x}\|^2 - (\mathbf{x}' \Theta_\varphi)^2 \leq \frac{1}{m_n^2}, \quad (4.19)$$

b) eine Bedingung an  $r$  für gegebenes  $\mathbf{x}$  und  $\varphi$ :

$$r \in [r_0, r_1] :=$$

$$\left[ \mathbf{x}' \Theta_\varphi - \sqrt{\frac{1}{m_n^2} - (\|\mathbf{x}\|^2 - (\mathbf{x}' \Theta_\varphi)^2)}, \mathbf{x}' \Theta_\varphi + \sqrt{\frac{1}{m_n^2} - (\|\mathbf{x}\|^2 - (\mathbf{x}' \Theta_\varphi)^2)} \right]. \quad (4.20)$$

Es werden zwei Fälle unterschieden.

**Fall 1.1:**

$$\|\mathbf{x}\| \leq \frac{1}{m_n}.$$

In diesem Fall ist Bedingung (4.19) für alle  $\varphi \in [0, \pi]$  erfüllt.

**Fall 1.2:**

$$\|\mathbf{x}\| > \frac{1}{m_n}.$$

Um Bedingung (4.19) zu erfüllen, kann der Winkel  $\varphi$  Werte aus dem Intervall  $[\varphi_0, \varphi_0 + 2 \arcsin(\frac{1}{m_n \|\mathbf{x}\|})] \subset [0, \pi]$  annehmen.

Sowohl für Fall 1.1 als auch für Fall 1.2 ist die an  $r$  zu stellende Bedingung:

$$r \in [r_0, r_1] \subset \left[ \mathbf{x}' \Theta_\varphi - \frac{1}{m_n}; \mathbf{x}' \Theta_\varphi + \frac{1}{m_n} \right].$$

Aus den Eigenschaften der Funktion  $H$  folgt, daß

$$\sup_{[r_0, r_1]} H(m_n(\mathbf{x} - r \Theta_\varphi)) \leq C.$$



Da  $w(r)$  die Dichte der Normalverteilung  $\mathcal{N}(0, \sigma^2)$  ist, gilt für gegebenes  $\mathbf{x}$  und  $\varphi$ :

$$\begin{aligned} \left| \int_{\mathbb{R}} H(m_n(\mathbf{x} - r\Theta_\varphi))w(r)dr \right| &= \left| \int_{r_0}^{r_1} H(m_n(\mathbf{x} - r\Theta_\varphi))w(r)dr \right| \\ &\leq C \frac{1}{m_n}. \end{aligned}$$

Es wird nun das Integral

$$\int_{\mathbb{R}} f_2(\mathbf{x} - r\Theta_\varphi)w(r)dr$$

in (4.18) abgeschätzt.  $f_2(\mathbf{x} - r\Theta_\varphi)$  ist ungleich 0 genau dann wenn  $r\Theta_\varphi \in U_1(\mathbf{x})$ , bei gegebenem  $\mathbf{x}$  und  $\varphi$ . Integration erfolgt also über das Intervall

$$[r_2, r_3] := [\mathbf{x}'\Theta_\varphi - \sqrt{1 - (\|\mathbf{x}\|^2 - (\mathbf{x}'\Theta_\varphi)^2)}, \mathbf{x}'\Theta_\varphi + \sqrt{1 - (\|\mathbf{x}\|^2 - (\mathbf{x}'\Theta_\varphi)^2)}].$$

Wegen (4.19) ist

$$[r_2, r_3] \supset [\mathbf{x}'\Theta_\varphi - (1 - \frac{1}{m_n}), \mathbf{x}'\Theta_\varphi + (1 - \frac{1}{m_n})].$$

Es gibt nun eine Konstante  $C > 0$ , sodaß

$$\int_{r_2}^{r_3} f_2(\mathbf{x} - r\Theta_\varphi)w(r)dr \geq C (2 - \frac{2}{m_n}). \quad (4.21)$$

Dann gilt für den  $\chi^2$ -Abstand zwischen  $g_2$  und  $g_3$

$$\begin{aligned} \chi^2(g_2, g_3) &\leq C \int_{U_{\frac{1}{m_n}}(0)} \int_0^\pi \delta^4 \frac{1}{m_n^2} \frac{1}{2 - \frac{2}{m_n}} d\mathbf{x} d\varphi \\ &\quad + C \int_{B \setminus U_{\frac{1}{m_n}}(0)} \int_{\varphi_0}^{\varphi_0 + 2 \arcsin(\frac{1}{m_n \|\mathbf{x}\|})} \delta^4 \frac{1}{m_n^2} \frac{1}{2 - \frac{2}{m_n}} d\mathbf{x} d\varphi. \end{aligned}$$

Mit der Substitution  $\mathbf{x} = s(\cos(\theta), \sin(\theta))$  ergibt sich

$$\chi^2(g_2, g_3) \leq \delta^4 \frac{1}{2 - \frac{2}{m_n}} \left( C \frac{1}{m_n^4} + C \int_{\frac{1}{m_n}}^1 \int_0^{2\pi} \frac{1}{m_n^2} \arcsin\left(\frac{1}{m_n s}\right) s ds d\theta \right).$$

Da  $m_n \rightarrow \infty$  für  $n \rightarrow \infty$ , kann  $\arcsin(\frac{1}{m_n s})$  für große  $n$  durch  $\frac{1}{m_n s}$  approximiert werden. Also ist

$$\begin{aligned} \chi^2(g_2, g_3) &\leq \delta^4 \left( C \frac{1}{m_n^4} + C \int \frac{1}{m_n^3} ds \right) \\ &\leq C \delta^4 \frac{1}{m_n^4} + C \delta^4 \frac{1}{m_n^3}, \end{aligned} \quad (4.22)$$

für  $n$  hinreichend groß. Die rechte Seite von (4.22) wird vom zweiten Summanden dominiert. Es folgt also für hinreichend große  $n$ , daß

$$\chi^2(g_2, g_3) \leq C \delta^4 \frac{1}{m_n^3}. \quad (4.23)$$

Aus (4.17) und (4.23) kann eine zusätzliche Beziehung zwischen  $\delta$  und  $m_n$  abgeleitet werden:

$$\delta^4 \frac{1}{m_n^3} \leq C \frac{1}{n}.$$

Unter Verwendung von Lemma 4.3 erhält man daraus

$$\delta \leq C n^{-\frac{\alpha-1}{2(2\alpha+1)}}.$$

Gemeinsam mit (4.16) ergibt sich somit für  $\mathbf{y} = \mathbf{y}_M = \mathbf{0}$  als untere Schranke für  $\Delta_2^*(n, a, \mathbf{y}_M)$

$$\begin{aligned} \Delta_2^*(n, a, \mathbf{y}_M) &\geq \sup_{\delta} \{ |\delta^2 H(\mathbf{0})|, \chi^2(g_2, g_3) \leq C/n \} \\ &= C n^{-\frac{\alpha-1}{2\alpha+1}}. \end{aligned}$$

**2. Fall:** Seien  $\mathbf{y}_M, \mathbf{x} \in B$ ,  $\mathbf{y}_M \neq \mathbf{0}$ . Ferner sei  $\|\mathbf{y}\| \leq 1 - \frac{2}{m_n}$ . Dann ist

$$\begin{aligned} \chi^2(g_2, g_3) &= \\ &\int_B \int_0^\pi \delta^4 \frac{|\int_{\mathbb{R}} H(m_n(\mathbf{x} - \mathbf{y}_M - r\Theta_\varphi)) w(r) dr|}{\int_{\mathbb{R}} f_2(\mathbf{x} - r\Theta_\varphi) w(r) dr} \left| \int_{\mathbb{R}} H(m_n(\mathbf{x} - \mathbf{y}_M - r\Theta_\varphi)) w(r) dr \right| d\varphi d\mathbf{x}. \end{aligned}$$

Analog zum 1. Fall ist  $H(m_n(\mathbf{x} - \mathbf{y}_M - r\Theta_\varphi)) = 0$ , wenn

$$\mathbf{x} - r\Theta_\varphi \notin U_{\frac{1}{m_n}}(\mathbf{y}_M). \quad (4.24)$$

Damit  $\mathbf{x} - r\Theta_\varphi \in U_{\frac{1}{m_n}}(\mathbf{y}_M)$ , muß gelten:

a) für  $\varphi$  bei gegebenem  $\mathbf{x}$ :

$$\|\mathbf{x} - \mathbf{y}_M\|^2 - ((\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi)^2 \leq \frac{1}{m_n^2}, \quad (4.25)$$

b) für  $r$  bei gegebenem  $\mathbf{x}$  und  $\varphi$ :

$$\begin{aligned} r \in [r_0, r_1] = & \left[ (\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi - \sqrt{\frac{1}{m_n^2} - \|\mathbf{x} - \mathbf{y}_M\|^2 + ((\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi)^2}, \right. \\ & \left. (\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi + \sqrt{\frac{1}{m_n^2} - \|\mathbf{x} - \mathbf{y}_M\|^2 + ((\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi)^2} \right]. \end{aligned} \quad (4.26)$$

Für gegebenes  $\mathbf{x}$  und  $\varphi$  kann in

$$\int_{\mathbb{R}} H(m_n(\mathbf{x} - \mathbf{y}_M - r\Theta_\varphi)) w(r) dr$$

der Integrationsbereich auf  $[r_0, r_1]$  eingeschränkt werden. Aus (4.25) und (4.26) folgt, daß  $|r_1 - r_0| \leq \frac{2}{m_n}$ . Damit das Integral

$$\int_{\mathbb{R}} f_2(\mathbf{x} - r\Theta_\varphi) w(r) dr$$

nicht verschwindet, muß es für gegebenes  $\mathbf{x}$ ,  $\varphi$  einen Bereich  $[r_2, r_3]$  geben, sodaß für  $r \in [r_2, r_3]$  jeweils  $\mathbf{x} - r\Theta_\varphi \in U_1(0)$  ist. Da  $\|\mathbf{y}_M\| \leq 1 - \frac{2}{m_n}$  und  $\mathbf{x} - r\Theta_\varphi \in U_{\frac{1}{m_n}}(\mathbf{y}_M)$  folgt, daß  $\mathbf{x} - r\Theta_\varphi \in U_1(0)$  für  $r \in [r_0, r_1]$ . Bei gegebenem  $\mathbf{x}$  und  $\varphi$  gilt daher die Inklusion

$$[r_0, r_1] \subset [r_2, r_3]. \quad (4.27)$$

Hinsichtlich der Länge des Intervalls ist dann

$$|r_3 - r_2| = |r_1 - r_0| + C \quad (4.28)$$

für ein  $C > 0$ . Es wird im folgenden der Ausdruck

$$\frac{\left| \int_{r_0}^{r_1} H(m_n(\mathbf{x} - \mathbf{y}_M - r\Theta_\varphi)) w(r) dr \right|}{\int_{r_2}^{r_3} f_2((\mathbf{x} - r\Theta_\varphi)) w(r) dr} \leq C \frac{\int_{r_0}^{r_1} w(r) dr}{\int_{r_2}^{r_3} w(r) dr}$$

bestimmt. Zu diesem Zweck definiert man

$$p := \int_{r_0}^{r_1} w(r) dr.$$

Wegen (4.28) gilt dann für eine Konstante  $C > 0$

$$\frac{\int_{r_0}^{r_1} w(r) dr}{\int_{r_2}^{r_3} w(r) dr} = \frac{p}{p + C}.$$

Da  $|r_1 - r_0| \leq 2/m_n$  ist, folgt  $p \leq 2C/m_n$ . Für alle  $\mathbf{x}, \varphi$  gibt es eine Konstante  $C$ , sodaß

$$\frac{\int_{r_0}^{r_1} w(r) dr}{\int_{r_2}^{r_3} w(r) dr} \leq C/m_n.$$

Weiterhin gilt für gegebenes  $\mathbf{y}_M$  wie in Fall 1

$$\left| \int_{\mathbb{R}} H(m_n(\mathbf{x} - \mathbf{y}_M - r\Theta_\varphi)) w(r) dr \right| \leq C/m_n.$$

Man betrachtet nun getrennt die beiden Fälle  $\mathbf{x} \in U_{\frac{1}{m_n}}(\mathbf{y}_M)$  sowie  $\mathbf{x} \in \mathcal{B} \setminus U_{\frac{1}{m_n}}(\mathbf{y}_M)$ . Für  $\mathbf{x} \in U_{\frac{1}{m_n}}$  ergibt sich der folgende Beitrag zum  $\chi^2$ -Abstand:

$$C \int_{U_{\frac{1}{m_n}}(\mathbf{y}_M)} \int_0^\pi \frac{1}{m_n^2} d\mathbf{x} d\varphi = C/m_n^4.$$

Für  $\mathbf{x} \in \mathcal{B} \setminus U_{\frac{1}{m_n}}(\mathbf{y}_M)$  kann  $\varphi$  Werte aus dem Bereich  $[\varphi_0, \varphi_0 + 2 \arcsin(\frac{1}{\|\mathbf{y}_M - \mathbf{x}\| m_n})]$  annehmen. Damit ist der Beitrag zum  $\chi^2$ -Abstand für diesen Fall:

$$C \int_{\mathcal{B} \setminus U_{\frac{1}{m_n}}(\mathbf{y}_M)} \int_{\varphi_0}^{\varphi_0 + 2 \arcsin(\frac{1}{m_n \|\mathbf{y}_M - \mathbf{x}\|})} \frac{1}{m_n^2} d\varphi d\mathbf{x} \leq C \int_{\mathcal{B} \setminus U_{\frac{1}{m_n}}(\mathbf{y}_M)} \frac{1}{m_n \|\mathbf{y}_M - \mathbf{x}\|} \frac{1}{m_n^2} d\mathbf{x}.$$

Die Substitution  $\mathbf{u} := \mathbf{x} - \mathbf{y}_M$  führt für die rechte Seite zu

$$C \int_{U_1(-\mathbf{y}_M) \setminus U_{\frac{1}{m_n}}(0)} \frac{1}{m_n^3} \frac{1}{\|\mathbf{u}\|} d\mathbf{u},$$

und mit  $\mathbf{u} = (s \sin(\theta), s \cos(\theta))$  erhält man:

$$\int_0^{\frac{2\pi}{m_n}} \int_{\frac{1}{m_n}}^1 \frac{1}{m_n^3} \frac{1}{s} s ds d\theta \leq C/m_n^3.$$

Dann folgt für den gesamten  $\chi^2$ -Abstand:

$$\chi^2(g_2, g_3) \leq C\delta^4/m_n^3 + C\delta^4/m_n^4.$$

Diese Schranke wird vom ersten Summanden dominiert und daher ist für  $n \rightarrow \infty$

$$\chi^2(g_2, g_3) \leq C \delta^4/m_n^3.$$

Daraus folgt unter Verwendung von (4.17) und Lemma 4.3:

$$\delta \leq C n^{-\frac{(\alpha-1)}{2(2\alpha+1)}}.$$

**3. Fall:** Seien  $\mathbf{y}_M, \mathbf{x} \in \mathcal{B}$  mit  $\|\mathbf{y}_M\| > 1 - \frac{2}{m_n}$ . Für diesen Fall gilt wie in (4.25) und (4.26) bei gegebenem  $\mathbf{x}$  und  $\mathbf{y}_M$  die folgende Bedingung an  $\varphi$ :

$$\|\mathbf{x} - \mathbf{y}_M\|^2 - ((\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi)^2 \leq 1/m_n,$$

und an  $r$  bei gegebenem  $\mathbf{y}_M$ ,  $\mathbf{x}$  und  $\varphi$ :

$$r \in [r_0, r_1] = \left[ (\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi - \sqrt{\frac{1}{m_n^2} - \|\mathbf{x} - \mathbf{y}_M\|^2 + ((\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi)^2}, \right. \\ \left. (\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi + \sqrt{\frac{1}{m_n^2} - \|\mathbf{x} - \mathbf{y}_M\|^2 + ((\mathbf{x} - \mathbf{y}_M)' \Theta_\varphi)^2} \right].$$

Damit  $\int_{\mathbb{R}} f_2(\mathbf{x} - r\Theta_\varphi)w(r)dr \neq 0$  ist muß gelten:

$$r \in [r_2, r_3] = [\mathbf{x}\Theta_\varphi - \sqrt{1 - (\|\mathbf{x}\|^2 - (\mathbf{x}\Theta_\varphi)^2)}, \mathbf{x}\Theta_\varphi + \sqrt{1 - (\|\mathbf{x}\|^2 - (\mathbf{x}\Theta_\varphi)^2)}].$$

Im Gegensatz zum 2. Fall, ist es hier möglich, daß  $[r_0, r_1]$  keine Teilmenge von  $[r_2, r_3]$  ist. Es gilt wie zuvor  $|r_1 - r_0| \leq \frac{1}{m_n}$ . Für  $(\mathbf{x}, \varphi)$  mit  $\mathbf{x} \in U_{\frac{1}{m_n}}(\mathbf{y}_M)$ ,  $\|\mathbf{x}\| = 1$  und  $\mathbf{x}\Theta_\varphi = 0$ , ist der Ausdruck

$$\frac{|\int_{r_0}^{r_1} H(m_n(\mathbf{x} - \mathbf{y}_M - r\Theta_\varphi)) w(r) dr|}{\int_{r_2}^{r_3} f_2(\mathbf{x} - r\Theta_\varphi) w(r) dr}$$

nicht beschränkt. Für alle anderen Fälle gilt, daß es ein  $C > 0$  und ein  $n_0$  gibt so daß für  $n \geq n_0$

$$\frac{|\int_{r_0}^{r_1} H(m_n(\mathbf{x} - \mathbf{y}_M - r\Theta_\varphi)) w(r) dr|}{\int_{r_2}^{r_3} f_2(\mathbf{x} - r\Theta_\varphi) w(r) dr} \leq C \frac{\int_{r_0}^{r_1} w(r) dr}{\int_{r_2}^{r_3} w(r) dr} \leq C/m_n.$$

Die Berechnung des  $\chi^2$ -Abstandes für den 3. Fall finde daher unter Ausnahme der Menge  $M^{(0)} := \{(\mathbf{x}, \varphi) \text{ mit } \mathbf{x} \in U_{\frac{1}{m_n}}(\mathbf{y}_M), \|\mathbf{x}\| = 1, \mathbf{x}\Theta_\varphi = 0\}$  statt, wobei  $M^{(0)}$  eine Nullmenge bezüglich der Integration in (4.18) darstellt. Es ergibt sich analog zum 2. Fall:

$$\chi^2(g_2, g_3) \leq C\delta^4/m_n^3,$$

und damit

$$\delta \leq C n^{-\frac{(\alpha-1)}{2(2\alpha+1)}}.$$

Daher gilt, wenn  $n$  hinreichend groß ist, für alle  $\mathbf{y}_M \in \mathcal{B}$

$$\begin{aligned}\Delta_2^*(n, a, \mathbf{y}_M) &\geq \sup_{\delta} \{\delta^2 H(\mathbf{0}), \chi^2(g_2, g_3) \leq C/n\} \\ &= C n^{-\frac{\alpha-1}{2\alpha+1}},\end{aligned}$$

wobei  $C$  von  $\mathbf{y}_M$  und von  $a$  abhängt. Damit ist Satz 4.1 bewiesen.

**q.e.d.**

Gemeinsam mit Theorem 4.2 ergibt sich der folgende

**Satz 4.2** *Gegeben sei  $\epsilon_0 > 0$ . Es sei  $a \in ]0, 1[$  beliebig gewählt aber fest und es gebe ein  $n_0$  sodaß  $\frac{|\log(a)|}{n} < \epsilon_0$  für alle  $n \geq n_0$ . Es sei  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$ . Dann ist für  $n \geq n_0$  und für  $\mathbf{y}_M \in \mathcal{B}$*

$$\inf_{T_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} MSE(\hat{T}_n(\mathbf{y}_M)) \geq C n^{-\frac{2(\alpha-1)}{2\alpha+1}}.$$

*Hierbei ist  $C$  eine Konstante, welche von  $\mathbf{y}_M$  und von  $a$  abhängt.*

**Beweis:** Die Aussage erhält man analog zu Korollar 4.1, unter Anwendung von Theorem 4.2 und Satz 4.1.

**q.e.d.**

Somit ist die Rate  $r = \frac{2(\alpha-1)}{2\alpha+1}$  die Minimax-Konvergenzrate für das hier behandelte Schätzproblem unter dem Mittleren Quadratischen Fehler. Das Ergebnis zur oberen Schranke aus Kapitel 3

$$\sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} MSE(\hat{f}_n(\mathbf{y})) \leq C n^{-\frac{2(\alpha-1)}{2\alpha+1}}$$

zeigt, daß der modifizierte Dekonvolutions-Kernschätzer  $\hat{f}_n$  für den Mittleren Quadratischen Fehler die optimale Konvergenzrate erreicht und damit im Sinne des Minimax-Kriteriums optimal ist.

## 4.4 Untere Schranke für den MISE

Die Konstante  $C$  in Satz 4.2 hängt von der betrachteten Stelle  $\mathbf{y}_M$  ab. Um eine untere Schranke für den Mittleren Integrierten Quadratischen Fehler zu berechnen, muß ein Ansatz gewählt werden, der die Bestimmung einer globalen Konstanten

ermöglicht. In Fan (1993) [18] wird im Fall eindimensionaler Zufallsvariablen eine Methode dafür eingeführt, welche als „cubical lower bound approach“ bezeichnet wird. Wir adaptieren diese Methode für die Bestimmung der unteren Schranke des Mittleren Integrierten Quadratischen Fehlers für das hier behandelte ToF-PET Dekonvolutions-Schätzproblem.

Es wird ein äquidistantes Gitter über der Menge  $[-1, 1] \times [-1, 1]$  mit der Gitterfeinheit  $m_n$  analog zu Definition 3.8 eingeführt und mit  $\mathcal{G}$  bezeichnet. Die Mittelpunkte der Gittersegmente von  $\mathcal{G}$  sind wie zuvor

$$\mathbf{y}_{M_{ij}} := \left(-1 + \frac{(2i+1)}{m_n}, 1 - \frac{(2j+1)}{m_n}\right), \quad i, j = 0, \dots, m_n - 1.$$

Mit  $k$  bezeichnet man einen skalaren Index zur Numerierung der Gittersegmente. Es sei  $k = j m_n + i + 1$ . Die Mittelpunkte der Gittersegmente sind dann  $\mathbf{y}_{M_k} = (y_{M_{k1}}, y_{M_{k2}}) \in \mathcal{Y}_M$ ,  $k \in [1, m_n^2]$ . Die Gittersegmente selbst sind definiert über

$$G_k := \{(y_1, y_2) \in \mathcal{B}, |y_1 - y_{M_{k1}}| \leq \frac{1}{m_n}, |y_2 - y_{M_{k2}}| \leq \frac{1}{m_n}\}.$$

Die Menge  $\tilde{\mathcal{B}}$  ist definiert wie in Definition 3.10 als

$$\tilde{\mathcal{B}} := \bigcup_{\substack{k \in \{1, \dots, m_n^2\} \\ G_k \subset \mathcal{B}}} G_k.$$

Für  $n \rightarrow \infty$  nähert  $\tilde{\mathcal{B}}$  die Menge  $\mathcal{B}$  beliebig genau an. Es bezeichne  $\tilde{\mathcal{Y}}_M := \{\mathbf{y}_{M_k} \in \mathcal{Y}_M \mid G_k \subset \mathcal{B}\}$  die Menge der Mittelpunkte der Gittersegmente  $\tilde{G}_k$  in  $\tilde{\mathcal{B}}$ . Es bezeichne  $K_n$  die Anzahl der Gittersegmente  $\tilde{G}_k$  in  $\tilde{\mathcal{B}}$ . Es gilt  $K_n < m_n^2$ . Es werden Funktionen  $f_2$  und  $f_\tau$  aus  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  wie folgt gewählt. Zunächst setzen wir wie in Abschnitt 4.3

$$f_2(\mathbf{y}) = \begin{cases} \frac{1}{\pi} & \text{für } \|\mathbf{y}\| < 1 \\ 0 & \text{sonst.} \end{cases}$$

Weiterhin sei  $\tau = (\tau_1, \dots, \tau_{K_n})$  ein Vektor aus  $\{0, 1\}^{K_n}$  und

$$f_\tau(\mathbf{y}) := f_2(\mathbf{y}) + \delta^2 \sum_{j=1}^{K_n} \tau_j H(m_n(\mathbf{y} - \mathbf{y}_{M_j})). \quad (4.29)$$



Damit  $f_\tau$  in  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  liegt, muß gemäß Lemma 4.3  $\delta \leq C m_n^{-\frac{\alpha-1}{2}}$  sein. Man betrachte  $f_{\tau_{0k}}(\mathbf{y})$  und  $f_{\tau_{1k}}(\mathbf{y})$  mit

$$\tau_{0k} - \tau_{1k} = (0, \dots, 1, \dots, 0).$$

Die 1 steht hierbei an  $k$ -ter Stelle,  $k \in \{1, \dots, K_n\}$ . Dann ist

$$f_{\tau_{0k}}(\mathbf{y}) - f_{\tau_{1k}}(\mathbf{y}) = \delta^2 H(m_n(\mathbf{y} - \mathbf{y}_{M_k})).$$

Man kann  $\Delta_2^*(n, a, \mathbf{y})$  für ein beliebiges  $\mathbf{y} \in \mathcal{B}$  abschätzen wie folgt

$$\begin{aligned} \Delta_2^*(n, a, \mathbf{y}) &:= \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \{|f_0(\mathbf{y}) - f_1(\mathbf{y})|, \chi^2(g_0, g_1) \leq C/n\} \\ &\geq \max_k \max_{\tau_{0k}, \tau_{1k}} \{|f_{\tau_{0k}}(\mathbf{y}) - f_{\tau_{1k}}(\mathbf{y})|, \chi^2(g_{\tau_{0k}}, g_{\tau_{1k}}) \leq C/n\} \\ &= \max_k \max_{\tau_{0k}, \tau_{1k}} \{|\delta^2 H(m_n(\mathbf{y} - \mathbf{y}_{M_k}))|, \chi^2(g_{\tau_{0k}}, g_{\tau_{1k}}) \leq C/n\}. \end{aligned} \quad (4.30)$$

Die Detektionsdichten  $g_{\tau_{0k}}$  und  $g_{\tau_{1k}}$  sind analog zu (4.12) definiert. Um eine untere Schranke für  $\Delta_2^*(n, a, \mathbf{y})$  für ein  $\mathbf{y} \in \mathcal{B}$  zu bestimmen, werden zwei Fälle unterschieden:

a) es gibt kein  $\mathbf{y}_{M_k} \in \tilde{\mathcal{Y}}_M$  für welches  $\mathbf{y} \in U_{\frac{1}{m_n}}(\mathbf{y}_{M_k})$ . Dann folgt aus (4.30)

$$\Delta_2^*(n, a, \mathbf{y}) \geq 0,$$

b) es gibt genau ein  $\mathbf{y}_{M_k} \in \tilde{\mathcal{Y}}_M$  mit  $\mathbf{y} \in U_{\frac{1}{m_n}}(\mathbf{y}_{M_k})$ . Dann ist

$$\Delta_2^*(n, a, \mathbf{y}) \geq \max_{\tau_{0k}, \tau_{1k}} \{\delta^2 H(m_n(\mathbf{y} - \mathbf{y}_{M_k})), \chi^2(g_{\tau_{0k}}, g_{\tau_{1k}}) \leq C/n\}. \quad (4.31)$$

Dies gilt unter den Voraussetzungen von Satz 4.2 mit  $C = c |\log(a)|$ .

Es wird der nicht-triviale Fall b) betrachtet. Der  $\chi^2$ -Abstand für gegebenes  $g_{\tau_{0k}}$ ,  $g_{\tau_{1k}}$  und  $\mathbf{y}_{M_k}$  ist

$$\chi^2(g_{\tau_{0k}}, g_{\tau_{1k}}) = \int_{\mathcal{B}} \int_0^\pi \frac{\left( \int_{\mathbb{R}} \delta^2 H(m_n(\mathbf{x} - r\boldsymbol{\Theta}_\varphi - \mathbf{y}_{M_k})) w(r) dr \right)^2}{\int_{\mathbb{R}} (f_2(\mathbf{x} - r\boldsymbol{\Theta}_\varphi) + \delta^2 \sum_{j=1}^{K_n} \tau_{0kj} H(m_n(\mathbf{x} - r\boldsymbol{\Theta}_\varphi - \mathbf{y}_{M_j}))) w(r) dr} dx d\varphi. \quad (4.32)$$

Für gegebenes  $\mathbf{x}$ ,  $\varphi$  und  $r$  ist

$$\left| \sum_{j=1}^{K_n} \tau_{0kj} H(m_n(\mathbf{x} - r\Theta_\varphi - \mathbf{y}_{M_j})) \right| \leq C < \infty.$$

Für  $n \rightarrow \infty$  konvergiert  $\delta$  gegen 0. Daraus folgt

$$\lim_{n \rightarrow \infty} \delta^2 \left| \sum_{j=1}^{K_n} \tau_{0kj} H(m_n(\mathbf{x} - r\Theta_\varphi - y_{M_j})) \right| = 0.$$

Für große  $n$  ergibt sich daher für alle  $g_{\tau_{0k}}, g_{\tau_{1k}}$  die Annäherung:

$$\chi^2(g_{\tau_{0k}}, g_{\tau_{1k}}) = \int_{\mathcal{B}} \int_0^\pi \frac{\left( \int_{\mathbb{R}} \delta^2 H(m_n(\mathbf{x} - r\Theta_\varphi - \mathbf{y}_{M_k})) w(r) dr \right)^2}{\int_{\mathbb{R}} f_2(\mathbf{x} - r\Theta_\varphi) w(r) dr} d\mathbf{x} d\varphi + o(1). \quad (4.33)$$

Damit  $\|m_n(\mathbf{x} - r\Theta_\varphi - \mathbf{y}_{M_k})\| \leq 1$  ist, muß für gegebenes  $\mathbf{x}$  und  $\varphi$  die Größe  $r$  aus  $[r_0, r_1]$  sein, mit

$$\mathbf{x} - r\Theta_\varphi \in U_{\frac{1}{m_n}}(\mathbf{y}_{M_k}) \quad \text{für } r \in [r_0, r_1].$$

Damit  $\|\mathbf{x} - r\Theta_\varphi\| \leq 1$  ist, muß bei gegebenem  $\mathbf{x}$  und  $\varphi$  die Größe  $r$  aus  $[r_2, r_3]$  sein, mit

$$\mathbf{x} - r\Theta_\varphi \in U_1(0) \quad \text{für } r \in [r_2, r_3].$$

Für  $\mathbf{y}_{M_k} \in \tilde{\mathcal{Y}}_M$  und  $\mathbf{x} - r\Theta_\varphi \in U_{\frac{1}{m_n}}(\mathbf{y}_{M_k})$  folgt  $\mathbf{x} - r\Theta_\varphi \in U_1(0)$ . Daher ist

$$[r_0, r_1] \subset [r_2, r_3].$$

Analog zu den Berechnungen in Abschnitt 4.3 ermitteln wir, daß es für  $\mathbf{y}_{M_k}$  eine Konstante  $C_{M_k}$  gibt mit

$$\chi^2(g_{\tau_{0k}}, g_{\tau_{1k}}) \leq C_{M_k} \delta^2 \frac{1}{m_n^3}$$

für alle  $g_{\tau_{0k}}, g_{\tau_{1k}}$ . Dann folgt

$$\delta \leq C_{M_k} n^{-\frac{(\alpha-1)}{2(2\alpha+1)}}.$$

Dann gilt für alle  $\mathbf{y} \in U_{\frac{1}{m_n}}(\mathbf{y}_{M_k})$ :

$$\begin{aligned} & \max_{\tau_{0k}, \tau_{1k}} \{ \delta^2 |H(m_n(\mathbf{y} - \mathbf{y}_{M_k}))|, \chi^2(g_{\tau_{0k}}, g_{\tau_{1k}}) \leq C/n \} \\ &= C_{M_k} n^{-\frac{(\alpha-1)}{2\alpha+1}} |H(m_n(\mathbf{y} - \mathbf{y}_{M_k}))|. \end{aligned}$$

Für jedes  $\mathbf{y}_{M_k}$ ,  $k = 1, \dots, K_n$  kann auf diese Weise eine Konstante  $C_{M_k}$  bestimmt werden. Dann existiert eine globale Konstante  $C$  für alle  $\mathbf{y} \in \mathcal{B}$ , für die es ein  $\mathbf{y}_{M_k}$  aus  $\tilde{\mathcal{Y}}_M$  mit der Eigenschaft  $\mathbf{y} \in U_{\frac{1}{m_n}}(\mathbf{y}_{M_k})$  gibt, sodaß  $\Delta_2^*$  für diese  $\mathbf{y}$  folgendermaßen abgeschätzt werden kann.

$$\Delta_2^*(n, a, \mathbf{y}) \geq C n^{-\frac{(\alpha-1)}{2\alpha+1}} |H(m_n(\mathbf{y} - \mathbf{y}_{M_k}))|.$$

Für den Mittleren Integrierten Quadratischen Fehler ergibt sich dann unter den Voraussetzungen von Satz 4.2:

$$\begin{aligned} \inf_{\hat{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} MISE(\hat{T}_n) &\geq C \int_{\mathcal{B}} \frac{(\Delta_2^*(n, a, \mathbf{y}))^2}{4} \frac{a}{2} d\mathbf{y} \\ &\geq C n^{-\frac{2(\alpha-1)}{2\alpha+1}} \int_{\mathcal{B}} \sum_{k=1}^{K_n} |H(m_n(\mathbf{y} - \mathbf{y}_{M_k}))|^2 d\mathbf{y}. \end{aligned}$$

Mit den Eigenschaften der Funktion  $H$  folgt

$$\inf_{\hat{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} MISE(\hat{T}_n) \geq C n^{-\frac{2(\alpha-1)}{2\alpha+1}}.$$

Ein Vergleich mit der oberen Schranke des Mittleren Integrierten Quadratischen Fehlers aus Kapitel 3

$$\sup_{\mathcal{H}^{\alpha,2}} MISE(\hat{f}_n) \leq C n^{-\frac{2(\alpha-1)}{2\alpha+1}},$$

zeigt die Optimalität des modifizierten Dekonvolutions-Kernschätzers  $\hat{f}_n$  unter dem Minimax-Kriterium bezüglich des Mittleren Quadratischen Fehlers.

## 4.5 Untere Schranke für die Supremum–Norm

Die Bestimmung einer unteren Schranke für die Supremum–Norm erfolgt ähnlich zu den Berechnungen in Abschnitt 4.3 und 4.4. Jedoch im Gegensatz zum Mittleren Quadratischen Fehler und dem Mittleren Integrierten Quadratischen Fehler ist  $\sup_{\mathcal{B}} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})|$  eine Zufallsvariable. In Anlehnung an Stone (1980, 1983) [52], [54] und Donoho und Liu (1991) verwenden wir folgende Definition einer unteren Schranke der Konvergenzgeschwindigkeit.

**Definition 4.7** *Es sei  $(b_n)$  eine Folge von positiven Zahlen.*

- Die Folge  $(b_n)$  heißt untere Schranke der Konvergenzgeschwindigkeit für die Schätzung von  $f$  in der Supremum–Norm, wenn für alle  $C > 0$

$$\liminf_{n \rightarrow \infty} \inf_{T_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(\sup_{\mathcal{B}} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| > C b_n) > 0.$$

- $(b_n)$  heißt erreichbare Konvergenzgeschwindigkeit, falls es einen Schätzer  $\hat{T}_n^*$  gibt, sodaß

$$\lim_{C \rightarrow \infty} \limsup_{n \rightarrow \infty} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(\sup_{\mathcal{B}} |\hat{T}_n^*(\mathbf{y}) - f(\mathbf{y})| > C b_n) = 0.$$

- Eine Folge  $(b_n)$  heißt optimale Konvergenzgeschwindigkeit, falls sie eine untere Schranke der Konvergenzgeschwindigkeit und erreichbare Konvergenzgeschwindigkeit ist.
- Falls  $(b_n)$  untere Schranke ist und  $(b'_n)$  erreichbar ist, dann existieren positive Konstanten  $C$  und  $n_0$ , sodaß  $b'_n \geq C b_n$  für  $n \geq n_0$ .

Man definiert die Menge  $\tilde{\mathcal{B}}$  analog zu Abschnitt 4.4 als die Menge der Gittersegmente eines äquidistanten Gitters  $\mathcal{G}$  der Feinheit  $m_n$ , welche ganz in  $\mathcal{B}$  enthalten sind. Das  $k$ -te Gittersegment sei mit  $\tilde{G}_k$  bezeichnet, hierbei ist  $k$  aus  $\{1, \dots, K_n\}$ , mit  $K_n < m_n^2$ . Der Mittelpunkt von  $\tilde{G}_k$  wird mit  $\mathbf{y}_{M_k}$  bezeichnet. Es sei  $f \in \mathcal{H}^{\alpha,2}(\mathcal{B})$  und  $g$  sei die mit  $f$  assoziierte Detektionsdichte. Dann ist für alle  $C > 0$

$$\begin{aligned} \Pr_f(\sup_{\mathcal{B}} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq C b_n) &= \Pr_f(\max_k \sup_{\tilde{G}_k} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq C b_n) \\ &\geq \Pr_f(\max_k |\hat{T}_n(\mathbf{y}_{M_k}) - f(\mathbf{y}_{M_k})| \geq C b_n) \\ &\geq \max_k \Pr_f(|\hat{T}_n(\mathbf{y}_{M_k}) - f(\mathbf{y}_{M_k})| \geq C b_n) \end{aligned} \tag{4.34}$$

und damit

$$\begin{aligned} & \inf_{T_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(\sup_{\mathcal{B}} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq C b_n) \geq \\ & \max_{1, \dots, K_n} \inf_{T_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(|\hat{T}_n(\mathbf{y}_{M_k}) - f(\mathbf{y}_{M_k})| \geq C b_n). \end{aligned} \quad (4.35)$$

Die Bestimmung der unteren Schranke der Supremum-Norm geht also über in die Lösung  $K_n$  lokaler Probleme, nämlich die Bestimmung der unteren Schranke des Schätzfehlers  $|\hat{T}_n(\mathbf{y}_{M_k}) - f(\mathbf{y}_{M_k})|$  für  $k = 1, \dots, K_n$ . Aus Abschnitt 4.3 ist bekannt, daß

$$\inf_{T_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(|\hat{T}_n(\mathbf{y}_{M_k}) - f(\mathbf{y}_{M_k})| \geq C \frac{\Delta_2^*(n, a, \mathbf{y}_{M_k})}{2}) \geq \frac{a}{2}.$$

Es sei wie in Abschnitt 4.4

$$f_\tau(\mathbf{y}) := f_2(\mathbf{y}) + \delta^2 \sum_{j=1}^{K_n} \tau_j H(m_n(\mathbf{y} - \mathbf{y}_{M_j}))$$

und

$$f_{\tau 0k}(\mathbf{y}) - f_{\tau 1k}(\mathbf{y}) = \delta^2 H(m_n(\mathbf{y} - \mathbf{y}_{M_k})).$$

Es sei  $a \in ]0, 1[$  beliebig gewählt aber fest und  $\epsilon_0 > 0$ . Wenn  $\frac{\log(a)}{n} < \epsilon_0$  für  $n \geq n_0$ , dann gilt für  $n \geq n_0$ :

$$\Delta_2^*(n, a, \mathbf{y}) \geq \max_k \max_{\tau 0k, \tau 1k} \{|f_{\tau 0k}(\mathbf{y}) - f_{\tau 1k}(\mathbf{y})|, \chi^2(g_{\tau 0k}, g_{\tau 1k}) \leq C/n\},$$

wobei  $C$  von  $a$  abhängt. Für  $\mathbf{y} = \mathbf{y}_{M_k}$ ,  $k = 1, \dots, K_n$  gilt dann

$$\Delta_2^*(n, a, \mathbf{y}_{M_k}) \geq \delta^2 |H(\mathbf{0})|,$$

wobei  $\delta$  so gewählt ist, daß

$$\chi^2(g_{\tau 0k}, g_{\tau 1k}) \leq C/n.$$

Wegen Theorem 4.2 und Satz 4.1 gilt die folgende Aussage, (4.36), bei gegebenem  $a \in ]0, 1[$  und  $\epsilon_0 > 0$  unter der Voraussetzung  $\frac{|\log(a)|}{n} < \epsilon_0$  für  $n \geq n_0$ .

Für  $n \geq n_0$  gilt

$$\inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(|\hat{T}_n(\mathbf{y}_{M_k}) - f(\mathbf{y}_{M_k})| \geq c |\log(a)| n^{-\frac{\alpha-1}{2\alpha+1}}) \geq \frac{a}{2}, \quad (4.36)$$

wobei  $c$  eine endliche, positive Konstante ist.

Für ein  $C > 0$  wählt man ein  $a \in ]0, 1[$ , sodaß  $C = c |\log(a)|$  für ein positives, endliches  $c$ , und bei gegebenem  $\epsilon_0 > 0$  wählt man dann  $n_0$ , sodaß  $\frac{|\log(a)|}{n} < \epsilon_0$  für  $n \geq n_0$ . (4.36) gilt dann für alle  $C > 0$  mit passend gewähltem  $a$  und bei gegebenem  $\epsilon_0$  für  $n$  groß genug. Da (4.36) auch für alle  $\mathbf{y}_{M_k}$ ,  $k = 1, \dots, K_n$  gilt, folgt für alle  $C > 0$  und  $n$  groß genug:

$$\max_{1, \dots, K_n} \inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(|\hat{T}_n(\mathbf{y}_{M_k}) - f(\mathbf{y}_{M_k})| \geq C n^{-\frac{\alpha-1}{2\alpha+1}}) \geq \frac{a}{2},$$

und daher folgt mit (4.35)

$$\inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(\sup_{\mathcal{B}} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq C n^{-\frac{\alpha-1}{2\alpha+1}}) \geq \frac{a}{2},$$

für alle  $C > 0$  und für  $n$  groß genug. Daraus folgt

$$\liminf_{n \rightarrow \infty} \inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(\sup_{\mathcal{B}} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq C n^{-\frac{\alpha-1}{2\alpha+1}}) > 0,$$

für alle  $C > 0$ . Daraus folgt, daß  $b_n := n^{-\frac{\alpha-1}{2\alpha+1}}$  die Bedingung an die untere Schranke der Konvergenzgeschwindigkeit von  $\sup_{\mathcal{B}} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})|$  gemäß Definition 4.7 erfüllt. Es ergibt sich der folgende

**Satz 4.3** *Für die Klasse  $\mathcal{T}_n$  aller Schätzer einer Dichtefunktion aus der Menge  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  gilt für alle  $C > 0$ :*

$$\liminf_{n \rightarrow \infty} \inf_{\mathcal{T}_n} \sup_{\mathcal{H}^{\alpha,2}(\mathcal{B})} \Pr_f(\sup_{\mathcal{B}} |\hat{T}_n(\mathbf{y}) - f(\mathbf{y})| \geq C n^{-\frac{\alpha-1}{2\alpha+1}}) > 0.$$

In Kapitel 3 wurde folgende obere Schranke für die Supremum-Norm des modifizierten Dekonvolutions-Kernschätzers hergeleitet:

$$\sup_{\mathcal{B}} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})| = O\left(n^{-\frac{\alpha-1}{2\alpha+1}} \ln(n)^{\frac{\alpha-1}{2\alpha+1}}\right) \quad \text{mit Wahrscheinlichkeit 1.}$$

Diese fast sichere obere Schranke weist bis auf einen logarithmischen Faktor eine Konvergenzordnung vergleichbar mit der unteren Schranke der Konvergenzgeschwindigkeit aus Satz 4.3 auf. Daher zeigt der modifizierte Dekonvolutions-Kernschätzer (bis auf einen logarithmischen Faktor) optimales Konvergenzhalten bezüglich der Supremum-Norm.

## 4.6 Einordnung der Ergebnisse in die Literatur

In Kapitel 3 haben wir die Sobolev-Räume  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  als Verallgemeinerung der Räume  $\mathcal{C}^{k,\mu}(\mathcal{B})$  eingeführt. Für  $\alpha \in \mathbb{R}^+$  mit  $\alpha > 1$  und  $\mu \in ]0, 1[$  sei  $\epsilon \in [0, 1[$  sodaß  $\alpha - \epsilon - \mu - 1 \in \mathbb{N}$ . Dann gilt, wie in Theorem 3.1 zu sehen war,  $\mathcal{H}^{\alpha,2}(\mathcal{B}) \subset \mathcal{C}^{\alpha-\epsilon-\mu-1,\mu}(\mathcal{B})$ . Im folgenden werden die Resultate dieser Arbeit Ergebnissen aus der Literatur über Dichteschätzung gegenübergestellt, bei denen Schätzungen über dem Raum  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  oder dem Hölder-Raum  $\mathcal{C}^{\alpha-\epsilon-\mu-1,\mu}(\mathcal{B})$  untersucht werden.

In Zhang (1992) [61] wird die Schätzung einer PET Emissionsdichte über dem Raum der Hölder-stetigen Funktionen studiert. Die der Schätzung zugrunde liegenden Meßdaten sind Beobachtungen  $(\mathbf{X}_j, \varphi_j)$ ,  $j = 1, \dots, n$ . Zhang erzielt eine Minimaxrate für den Mittleren Quadratischen Fehler der Ordnung  $O\left(n^{-\frac{2\alpha-2\epsilon-2}{2\alpha-2\epsilon+1}}\right)$  gleichmäßig über dem Funktionenraum  $\mathcal{C}^{\alpha-\epsilon-\mu-1,\mu}(\mathcal{B})$ . Dieses Ergebnis ist vergleichbar (aber schwächer) mit der optimalen Rate, die in der vorliegenden Arbeit gleichmäßig über  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  bestimmt wird.

Johnstone und Silverman (1990) [31] behandeln die Schätzung einer PET Emissionsdichte auf dem Raum  $\mathcal{H}^{\alpha,2}(\mathcal{B})$ . Zwei Fälle werden einander gegenübergestellt: die Schätzung basierend auf direkten Emissions-Beobachtungen (ein Fall welcher in der Praxis nie auftritt) und die Schätzung basierend auf indirekten Beobachtungen ohne Flugzeitinformation. Im zweiten Fall bestehen die Beobachtungen aus den Auftreffpunkten von Photonenpaaren auf dem Detektorring. Bekannt sind damit nur die Flugbahnen der Photonen. Unter der Annahme direkter Beobachtungen, ist die Minimax-Konvergenzrate des Mittleren Integrierten Quadratischen Fehlers in Johnstone und Silverman (1990) von der Ordnung  $O\left(\left(\frac{n}{\ln(n)}\right)^{-\frac{(2\alpha-2)}{2\alpha}}\right)$  und im Falle indirekter Beobachtungen von der Ordnung  $O\left(n^{-\frac{(2\alpha-2)}{(2\alpha+2)}}\right)$ .

Das Ergebnis der vorliegenden Arbeit interpoliert zwischen den Ergebnissen von Johnstone und Silverman für den Fall direkter Beobachtungen und für den Fall indirekter Beobachtungen ohne Flugzeitinformation. Für  $\alpha > 1$  gilt

$$O(MISE^{direkt}) \leq O(MISE^{ToFPET}) \leq O(MISE^{indirekt}).$$

Damit liefert diese Arbeit eine Quantifizierung der Wirkung zusätzlicher Flugzeitinformation hinsichtlich der Beschleunigung der Konvergenzrate gegenüber der Situation ohne Flugzeitinformation. Wenn direkte Meßdaten für die Schätzung einer Dichte vorliegen, sollte der Meßfehler schneller konvergieren als bei Schätzung unter Verwendung indirekter Meßdaten mit oder ohne Flugzeitinformation. Die Verbesserung der Konvergenzrate bei Verwendung von Flugzeitinformation trotz der auch in diesem Fall bestehenden Überlagerung mit einem nicht vernachlässigbaren Fehler, hängt damit zusammen, daß die Realisierungen der unbekanntesten Fehlergröße jeweils auf bekannten eindimensionalen Unterräumen von  $\mathcal{B}$  liegen (den durch die Flugbahnen definierten Strecken), und damit geringere effektive Dimension haben als die Beobachtungspunkte.



# Kapitel 5

## Test des Schätzverfahrens

In Kapitel 3 und 4 wurden theoretische Ergebnisse zur Konvergenz des modifizierten Dekonvolutions–Kernschätzers vorgestellt. In diesem Kapitel werden die theoretischen Ergebnisse durch eine praktische Anwendung des Schätzverfahrens auf simulierte Daten ergänzt. Hierbei wird insbesondere das Konvergenzverhalten des Rekonstruktionsergebnisses unter der Supremum–Norm beobachtet. Die Ermittlung des Schätzers aus Gleichung (2.10) wird von einem **C**–Programm durchgeführt. Die Erzeugung von ToFPET–Beobachtungspaaren  $(\mathbf{X}_i, \varphi_i), i = 1, \dots, n$ , findet gemäß (2.1)–(2.2) und im Einklang mit den Modellannahmen (M1) statt. Dabei wurde auf das Programmpaket *Splus* zurückgegriffen. Zunächst wird im nächsten Abschnitt die Simulation eines geeigneten Datensatzes beschrieben.

### 5.1 Simulation von ToFPET–Beobachtungen

Um  $n$  u.i.v. ToFPET–Beobachtungspaare zu erhalten, erzeugen wir zunächst  $n$  u.i.v. Positronen–Emissionsorte  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  mittels einer von uns gewählten Beispieldichte. Gemäß (2.1)–(2.2) werden die  $\mathbf{Y}_i$  dann mit Beobachtungsfehlern  $\mathbf{Z}_i = r_i \Theta_{\varphi_i}$  überlagert.  $r_i$  und  $\varphi_i$  unterliegen den Bedingungen von (M1). Als Beispieldichte zur Erzeugung von Emissionsorten verwenden wir die Dichtefunktion einer Pearson–Typ–II Verteilung. Diese Wahl ist allerdings nicht als Modell eines tatsächlichen Emissionsprozesses zu verstehen. Die hier gewählte Dichtefunktion dient lediglich dazu, das implementierte Rekonstruktionsverfahren zu testen.

### 5.1.1 Die Pearson–Typ–II Dichtefunktion

Eine Dichte über  $\mathbb{R}^p$  heißt Pearson–Typ–II Dichte mit Parametern  $m$ ,  $\mu$  und  $\Sigma$  falls sie die Darstellung

$$f_{m,\mu,\Sigma}(\mathbf{y}) := \frac{\Gamma(p/2+m+1)}{\Gamma(m+1)\pi^{p/2}} |\Sigma|^{-1/2} (1 - (\mathbf{y} - \mu)' \Sigma^{-1} (\mathbf{y} - \mu))^m,$$

für  $\mathbf{y}$  mit  $(\mathbf{y} - \mu)' \Sigma^{-1} (\mathbf{y} - \mu) \leq 1$  besitzt und sonst den Wert 0 annimmt. Hierbei ist  $\Sigma$  eine  $p \times p$  positiv–definite Matrix,  $\mu$  ist ein Vektor aus  $\mathbb{R}^p$  und  $\Gamma$  bezeichnet die Gamma–Funktion:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt, \quad x > 0.$$

Es gilt

$$E[\mathbf{y}] = \mu, \quad Cov(\mathbf{y}) = \frac{1}{2m + p + 2} \Sigma.$$

Wir betrachten den Fall  $p = 2$  und den Spezialfall

$$\Sigma = \begin{pmatrix} \sigma_P^2 & 0 \\ 0 & \sigma_P^2 \end{pmatrix}, \quad \mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

den wir in folgender Weise bezeichnen.

$$f_{m,\sigma_P^2}(\mathbf{y}) = \begin{cases} \frac{\Gamma(m+2)}{\Gamma(m+1)} \frac{1}{\pi} \frac{1}{\sigma_P^2} \left(1 - \frac{1}{\sigma_P^2} (y_1^2 + y_2^2)\right)^m, & \text{für } \|\mathbf{y}\| \leq \sigma_P \\ 0, & \text{sonst.} \end{cases}$$

Wir wählen nun die Parameter  $m$  und  $\sigma_P^2$  wie folgt:

$$m = 2, \quad \sigma_P^2 = 8.$$

Abbildung 5.1 enthält eine Darstellung der gewählten Dichtefunktion  $f_{2,8}(\mathbf{y})$ . Diese Dichtefunktion dient uns als einfaches Beispiel einer Intensitätsverteilung mit kompaktem Träger.

### 5.1.2 Erzeugen von ToFPET–Beobachtungspaaren

Um Daten aus der Dichte  $f_{2,8}$  zu erzeugen, wird gemäß Johnson (1987) [35] der folgende Algorithmus verwendet, welchen wir in Splus implementieren:

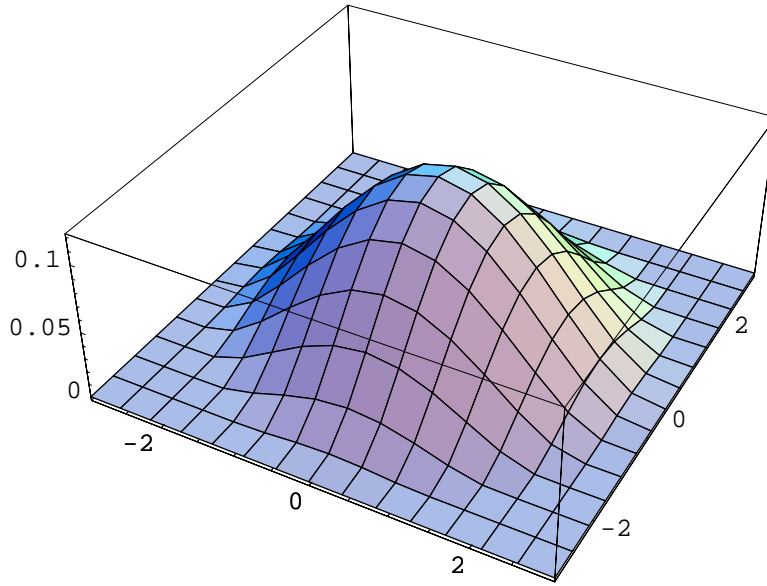


Abbildung 5.1: Darstellung der Pearson–Typ–II Dichte  $f_{2,8}(\mathbf{y})$ .

1. Erzeuge  $Y_1$  mittels einer Beta–Verteilung mit Parametern  $p = \frac{1}{2}, q = m + \frac{3}{2}$ . Eine Beta–Verteilung mit Parametern  $p > 0$  und  $q > 0$  besitzt die Dichtefunktion (siehe zum Beispiel Johnson und Kotz (1970) [38]):

$$w(y) := \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1} (1-y)^{q-1}, \quad 0 \leq y \leq 1.$$

2. Setze  $Y_1^* = \pm Y_1^{\frac{1}{2}}$ . Hierbei wird das Vorzeichen zufällig gewählt, positives und negatives Vorzeichen treten jeweils mit Wahrscheinlichkeit  $1/2$  auf.
3. Erzeuge eine Zufallsvariable  $V_2$  aus einer Beta–Verteilung mit Parametern  $p = \frac{1}{2}, q = m + 1$ .
4. Wähle  $Y_2^* = \pm (V_2(1 - (Y_1^*)^2))^{\frac{1}{2}}$ .
5. Setze  $\mathbf{Y} = \sigma_P \begin{pmatrix} Y_1^* \\ Y_2^* \end{pmatrix} + \mu$ .

Die resultierenden Daten  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  müssen nun zusätzlich mit einem Beobachtungsfehler der Form  $r_i(\cos(\varphi_i), \sin(\varphi_i))$ ,  $i = 1, \dots, n$  überlagert werden. Daraus ergeben sich die Beobachtungsdaten  $\mathbf{X}_1, \dots, \mathbf{X}_n$ . Gemäß Kapitel 2, sind  $r_i$ ,  $i = 1, \dots, n$ , unabhängige  $\mathcal{N}(0, \sigma^2)$ –verteilte Zufallsvariablen, und  $\varphi_i$ ,  $i = 1, \dots, n$ ,

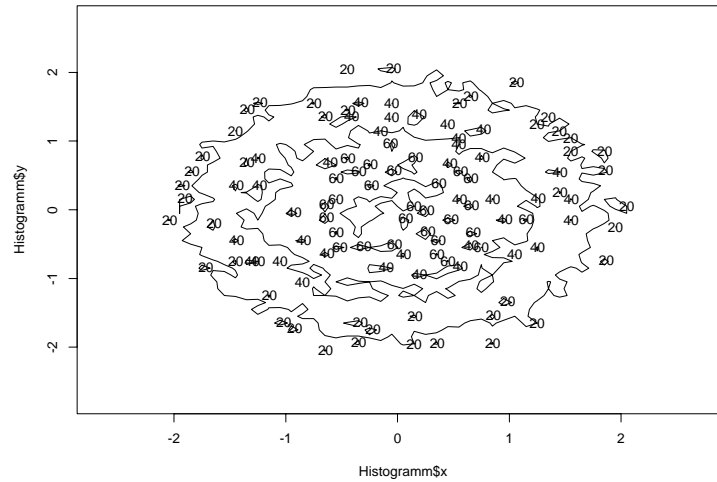
unabhängige  $\mathcal{U}([0, \pi])$ -verteilte Zufallsgrößen. Für die hier vorgestellten Rechnungen wurde  $\sigma = 0.02$  gewählt. Im folgenden ist das Splus Programm „Datengen.fkt“ aufgeführt, welches zur Erzeugung von  $n$  unabhängig identisch verteilten Beobachtungspaaren  $(\mathbf{X}_1, \varphi_1), \dots, (\mathbf{X}_n, \varphi_n)$  verwendet wurde.

### Funktion Datengen.fkt:

```
function(n, m, sigma)
{
X1data <- c(rbeta(n, 0.5, 1.5 + m))
hilf1 <- c(rbinom(n, 1, 0.5))
sign1 <- 2 * (hilf1 - 0.5)
X1 <- sign1 * sqrt(X1data)
V2 <- c(rbeta(n, 0.5, m + 1))
hilf2 <- c(rbinom(n, 1, 0.5))
sign2 <- 2 * (hilf2 - 0.5)
X2 <- sign2 * sqrt((V2 * (1 - X1^2)))
xdat <- sigma * X1
ydat <- sigma * X2
rad <- c(rnorm(n, mean = 0, sd = 0.02))
phi <- c(runif(n, 0, 3.1415))
arcx <- rad * cos(phi)
arcy <- rad * sin(phi)
xdat2 <- xdat + arcx
ydat2 <- ydat + arcy
cat(xdat2, file = „Xkoord “, fill = T)
cat(ydat2, file = „Ykoord “, fill = T)
cat(phi, file = „Winkel “, fill = T)
Histogramm <- hist2d(xdat2, ydat2, nxbins = 64, nybins = 64)
postscript(„ergebnis.ps “)
contour(Histogramm)
persp(Histogramm)
}
```

Abbildung 5.2 zeigt die Konturdarstellung und die perspektivische Darstellung eines mit der Splus-Funktion Datengen.fkt generierten Datensatzes vom Umfang  $n = 50.000$ .

a)



b)

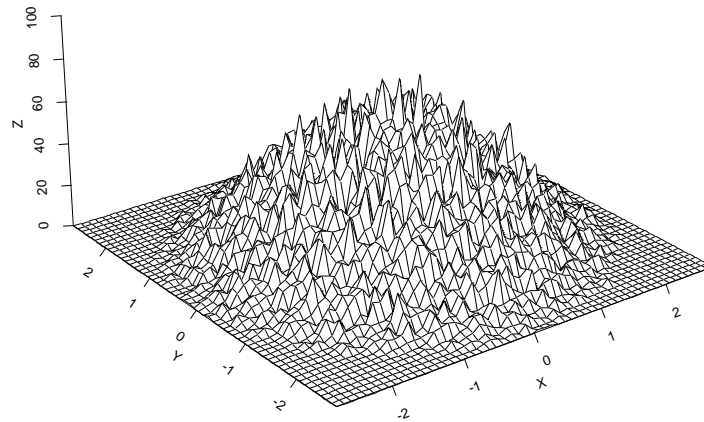


Abbildung 5.2: a) Konturdarstellung eines Datensatzes  $\mathbf{X}_1, \dots, \mathbf{X}_n$  erzeugt mit `Datengen.fkt` für  $n = 50.000$ . b) Perspektivische Darstellung des gleichen Datensatzes.

## 5.2 Bestimmung der Funktionenklasse

Um die später in diesem Kapitel dargestellten experimentellen Ergebnisse im Rahmen von Kapitel 3 und 4 betrachten zu können, soll untersucht werden, in welchen Sobolev-Raum  $\mathcal{H}^{\alpha,2}(\mathbb{R}^2)$  wir die gewählte Emissionsdichte  $f_{2,8}(\mathbf{y})$  einordnen können. Man betrachtet die Norm:

$$\|f_{2,8}(\mathbf{y})\|_{\mathcal{H}^{\alpha,2}(\mathbb{R}^2)} := \int_{\mathbb{R}^2} (1 + \|\mathbf{t}\|^2)^\alpha |\Phi_f(\mathbf{t})|^2 dt_1 dt_2,$$

welche sich auch schreiben läßt als

$$\|f_{2,8}(\mathbf{y})\|_{\mathcal{H}^{\alpha,2}(\mathbb{R}^2)} = \int_0^{\infty} \int_0^{2\pi} (1 + r^2)^\alpha |\Phi_f(r, \varphi)|^2 r dr d\varphi. \quad (5.1)$$

Die Dichtefunktion  $f_{2,8}$  ist radialsymmetrisch, also gibt es eine Darstellung der Form  $f_{2,8}(y_1, y_2) = \zeta(z)$  mit  $z = \sqrt{y_1^2 + y_2^2}$ . Es gilt gemäß Arsac (1966) [3], daß die Fouriertransformierte einer radialsymmetrischen Funktion ebenfalls radialsymmetrisch ist. Nach [3] kann man  $\Phi_f(r, \varphi)$  darstellen als

$$\Phi_f(r, \varphi) = \Xi_\zeta(r) := \int_0^{\infty} z \zeta(z) J_0(rz) dz, \quad (5.2)$$

welches gleich der Hankeltransformation zur Ordnung 0 der Funktion  $\zeta(z)$  an der Stelle  $r > 0$  ist. Dabei ist  $J_0$  die Besselfunktion der Ordnung 0. Allgemein ist die Besselfunktion der Ordnung  $\nu \in \mathbb{R}$  definiert über:

$$J_\nu(z) = \left(\frac{1}{2} z\right)^\nu \sum_{j=0}^{\infty} \frac{(-1)^j \left(\frac{1}{2} z\right)^{2j}}{\Gamma(j+\nu+1)j!}, \quad z > 0. \quad (5.3)$$

Die Funktion  $\zeta(z)$  ist definiert als

$$\zeta(z) = \begin{cases} \frac{\Gamma(4)}{\Gamma(3)\pi} \frac{1}{8} (1 - \frac{1}{8} z^2)^2 & \text{für } z^2 \leq 8, \\ 0 & \text{sonst.} \end{cases}$$

Die Hankeltransformation in (5.2) nimmt für die Funktion  $\zeta(z)$  folgende Form an, [15]:

$$\Xi_{\zeta}(r) = \frac{\Gamma(4)}{2\pi\sqrt{8}} r^{-3} J_3(\sqrt{8}r). \quad (5.4)$$

Wenn (5.4) in (5.1) eingesetzt wird, folgt

$$\|f_{2,8}\|_{\mathcal{H}^{\alpha,2}(\mathbb{R}^2)} := C \int_0^{\infty} (1+r^2)^{\alpha} r^{-6} J_3^2(\sqrt{8}r) r dr. \quad (5.5)$$

Die Besselfunktion  $J_3(\sqrt{8}r)$  kann für  $r \rightarrow 0$  angenähert werden über

$$J_3(\sqrt{8}r) \sim Cr^3.$$

Der Integrand in (5.5) kann daher für  $r = 0$  stetig durch 0 fortgesetzt werden. Um die Existenz des Integrals zu untersuchen, muß das Verhalten von  $J_3(\sqrt{8}r)$  für  $r \rightarrow \infty$  betrachtet werden. Für  $r \rightarrow \infty$  gilt gemäß [15]

$$J_3(\sqrt{8}r) = C \frac{1}{\sqrt{r}} (\cos(\sqrt{8}r - 7/4\pi) + O(1/r)).$$

Für große  $r$  wird  $J_3^2(\sqrt{8}r)$  dominiert von  $C \frac{1}{r} \cos^2(\sqrt{8}r - 7/4\pi)$ . Für  $r \rightarrow \infty$  wird der Integrand in (5.5) daher dominiert von

$$C (1+r^2)^{\alpha} r^{-6} r \frac{1}{r} \cos^2(\sqrt{8}r - 7/4\pi). \quad (5.6)$$

Existenz der Sobolev-Norm in 5.1 ist gewährleistet, wenn  $\alpha < 2.5$  ist.  $f_{2,8}$  ist daher Element eines Sobolev-Raumes  $\mathcal{H}^{\alpha,2}$  über  $\mathbb{R}^2$  mit Parameter  $\alpha \in ]1, 2.5[$ .

## 5.3 Algorithmische Berechnung der Rekonstruktionsdichte

### 5.3.1 Einführende Bemerkungen

Ein C-Programm wird verwendet, um aus den in Abschnitt 5.1 beschriebenen Daten den modifizierten Dekonvolutions-Kernschätzer  $\hat{f}_n(\mathbf{y})$  der Funktion  $f_{2,8}(\mathbf{y})$

zu bestimmen. Berechnung von  $\hat{f}_n(\mathbf{y})$  erfolgt über einem äquidistanten Gitter  $\mathcal{G}$  der Feinheit  $\kappa$  über dem Bereich  $\mathcal{M}' := [-3, 3] \times [-3, 3]$ , wobei  $\kappa$  fest gewählt wird. Die Gittermittelpunkte sind

$$\mathbf{y}_{M_{ij}} = \left(-3 + \frac{3 \cdot (2i + 1)}{\kappa}, 3 - \frac{3 \cdot (2j + 1)}{\kappa}\right), \quad i, j = 0, \dots, \kappa - 1.$$

Für jedes  $\mathbf{y}_{M_{ij}}$  wird der rekonstruierte Wert mit dem Wert der Pearson–Typ–II Dichte verglichen und der Schätzfehler  $|\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  berechnet. Die Größe

$$\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$$

wird anschließend bestimmt, sie dient hier als Approximation der Supremum–Norm  $\sup_{\mathcal{M}'} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})|$ . In Abschnitt 5.4 wird das Maximum  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  für selektive  $n$  beobachtet und eine experimentelle Konvergenzrate in  $n$  wird ermittelt. Diese wird mit den theoretischen Ergebnissen zur Konvergenzrate der Supremum–Norm aus Kapitel 3 und 4 verglichen.

Die Implementierung verwendet die folgenden Eingabedateien:

- Jeweils eine Datei mit den x-Koordinaten der generierten Datenpunkte, den y-Koordinaten sowie den Auftreffwinkeln der Photonenpaare auf dem Detektorring.
- Eine Definitions–Datei, in welcher die Anzahl der vorliegenden Datenpunkte  $n$  festgelegt wird. Hier werden auch die Parameter  $m$  und  $\sigma_P^2$  der Pearson–Typ–II Dichte, sowie  $\sigma$ , die Standardabweichung der Fehlergröße  $r$  spezifiziert.

Im Programm wird die Gitterfeinheit als  $\kappa = 20$  festgelegt. Eine Implementierung des Schätzalgorithmus (2.10) für die Punkte  $\mathbf{y}_{M_{ij}}$

$$\hat{f}_n(\mathbf{y}_{M_{ij}}) := \frac{1}{(2\pi)^{2n}} \sum_{k=1}^n \int_{\mathbf{t} \in \mathbb{R}^2} e^{-i\mathbf{t}'(\mathbf{y}_{M_{ij}} - \mathbf{x}_k)} \frac{G(\mathbf{t}, \Theta_{\varphi_k})}{M(\mathbf{t})} \Phi_K(\lambda_n \mathbf{t}) d\mathbf{t}$$

erfordert folgende Diskretisierungen.

- Die Integration über den Bereich  $\mathbf{t} \in \left[\frac{-1}{\lambda_n}, \frac{1}{\lambda_n}\right] \times \left[\frac{-1}{\lambda_n}, \frac{1}{\lambda_n}\right]$  wird durch Summation über ein äquidistantes Gitter  $\mathcal{G}_{\mathbf{t}}$  der Segmentlänge  $\Delta t$  ersetzt. Tests der Implementierung zeigten einen geringen Einfluß von  $\Delta t$  auf das Rekonstruktionsergebnis und für die hier vorgestellten Ergebnisse wurde  $\Delta t = 0.1$  gewählt.



- Die Berechnung der Funktion  $M(\mathbf{t}) = E[e^{i\mathbf{t}'\Theta_\varphi} G(\mathbf{t}, \Theta_\varphi)]$  für  $\mathbf{t} \in \mathcal{G}_t$  wird diskretisiert.

Da  $\hat{f}_n(\mathbf{y}_{M_{ij}})$  nach Korollar 3.1 eine reellwertige Funktion ist, kann der Exponentialterm  $e^{i\mathbf{t}'(\mathbf{y}_{M_{ij}} - \mathbf{X}_k)}$  in der Definition von  $\hat{f}_n(\mathbf{y}_{M_{ij}})$  durch den Term  $\cos(\mathbf{t}'(\mathbf{y}_{M_{ij}} - \mathbf{X}_k))$  ersetzt werden.

### 5.3.2 Wahl der Bandweite

Satz 3.10 spezifiziert als optimale Bandweite für den modifizierten Dekonvolutions-Kernschätzer unter der Supremum-Norm

$$\lambda_n^{sup} = C \left( \frac{\ln(n)}{n} \right)^{\frac{1}{2\alpha+1}}. \quad (5.7)$$

Da jedoch die Konstante in (5.7) unbekannt ist, kann dieses Ergebnis nicht verwendet werden, um in der Praxis die optimale Bandweite zu bestimmen. Zur praktischen und effizienten Umsetzung des modifizierten Dekonvolutions-Kernschätzers müsste ein Verfahren zur automatischen Bandweitenselektion für das Schätzverfahren entwickelt werden. Unser Ziel in diesem Kapitel ist eine Veranschaulichung der Ergebnisse, welche in Kapitel 2 bis 4 erörtert wurden. Daher haben wir hier auf die Anwendung verfeinerter Bandweiten-Selektions Verfahren, so wie sie z. B. in Herrmann et al. (1995) [24], Werthenbach und Herrmann (1998) [57], Marron und Padgett (1987) [32], Hardle et al. (1990) [23] beschrieben sind, verzichtet und ein einfaches iteratives Verfahren zur Approximation von  $\lambda_n^{sup}$  implementiert. Hierbei führen wir für jeden Datensatz die Berechnung von  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  für ein Gitter von  $\lambda_n$ -Werten durch. Es wird das  $\lambda_n$  aus dem Gitter bestimmt, für welches  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  minimal ist. Dieses,  $\lambda_n^{opt}$ , wird als Approximation für  $\lambda_n^{sup}$  angesehen und als Bandweite des Schätzers verwendet. Dieses Verfahren wird für jeden neuen Datensatz wiederholt. Man verfährt wie folgt:

- 1) Setze ein grobes Gitter von  $\lambda_n$ -Werten fest, z. B.  $\lambda_n = 1, \dots, 10$  mit Schrittweite 1.
- 2) Berechne ein Rekonstruktionsergebnis für alle  $\lambda_n$  und bestimme das  $\lambda_n^*$  aus dem Gitter, für welches der zugehörige Schätzer den kleinsten Wert für  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  liefert.
- 3) Definiere ein zweites, feineres Gitter um  $\lambda_n^*$  herum z. B.  $\lambda_n = \lambda_n^* - 1, \dots, \lambda_n^* + 1$  mit Schrittweite  $\Delta\lambda$ .

- 4) Bestimme das  $\lambda_n^{**}$  aus dem Gitter, für welches der zugehörige Schätzer den kleinsten Wert für  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  liefert.
- 5) Wenn  $\lambda_n^{**}$  am Rand des Gitterbereiches liegt, oder wenn höhere Genauigkeit gewünscht ist, wiederhole Schritt 4) mit einem neuen Gitter. Ansonsten setze  $\lambda_n^{opt} = \lambda_n^{**}$ .

Diese Vorgehensweise liefert eine Annäherung des bezüglich der Supremum-Norm optimalen Rekonstruktionsergebnisses. Wenn man - zum Beispiel basierend auf bereits durchgeführten Rechnungen mit anderen Datensätzen - Vorwissen über die Lage von  $\lambda_n^{sup}$  besitzt, so kann man bereits mit einem feineren Gitter über dem entsprechenden Bereich beginnen. Dies ist insbesondere dann günstig, wenn der zu Grunde liegende Datensatz groß ist und daher die Rechenzeit pro Rekonstruktions-Berechnung ansteigt. Wir beobachteten bei Durchführung der Schritte 1) und 2), daß für  $\lambda_n > 1$  die Funktion überglättet wird und demzufolge  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  sehr groß wird. In den hier dargestellten Beispielen ergaben sich jeweils  $\lambda_n^{opt}$ -Werte zwischen 0 und 1.

Wir stellen in Abschnitt 5.4 die Rekonstruktionsergebnisse, sowie die Schätzfehler, welche sich bei Verwendung von  $\lambda_n^{opt}$  ergaben mit Hilfe von Splus graphisch dar. Hierbei wird zwischen den Ergebnissen an den Gittermittelpunkten interpoliert und eine stetige, perspektivische Darstellung von  $\hat{f}_n(\mathbf{y})$  und  $|\hat{f}_n(\mathbf{y}) - f(\mathbf{y})|$  erzeugt.

In Tabelle 5.1 und Abbildung 5.13 führen wir die Ergebnisse für  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  auf, welche für verschiedene  $n$  mit den jeweils ermittelten  $\lambda_n^{opt}$  erzielt wurden.

## 5.4 Darstellung der Rekonstruktionsergebnisse

In diesem Abschnitt stellen wir die Ergebnisse der Implementierung des modifizierten Dekonvolutions-Kernschätzers für die Werte  $n = 5000$ ,  $n = 10.000$ ,  $n = 50.000$  und  $n = 75.000$  dar.

Abbildung 5.3 enthält das Ergebnis für  $n = 5000$ . Für die Bandweite  $\lambda_n^{opt}$  wurde der Wert 0.43 ermittelt, das Maximum des Fehlerbetrages  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  ergab sich zu 0.005859. Es wurde hierbei zunächst ein Gitter von  $\lambda_n$ -Werten über  $[0.34, 0.46]$  mit Gitterfeinheit 0.02 verwendet, das kleinste Ergebnis für  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f_{2,8}(\mathbf{y}_{M_{ij}})|$  ergab sich für  $\lambda_n = 0.42$ . Es wurde dann ein

Rekonstruktion, N=5000,Bandweite=0.43

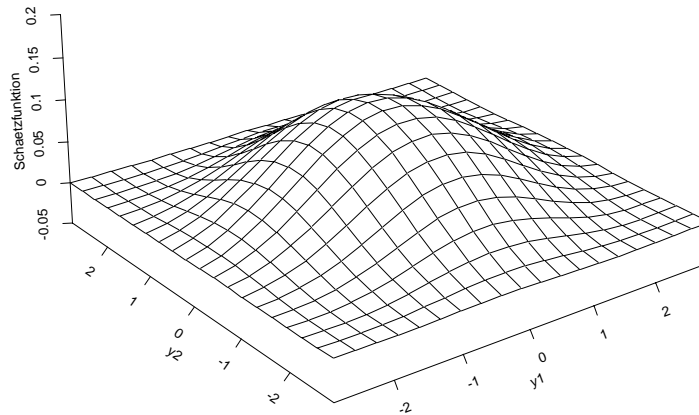


Abbildung 5.3: Rekonstruktion der Pearson–Typ–II Dichte  $f_{2,8}(\mathbf{y})$  aus 5000 Datenpunkten.

Gitter über  $[0.41, 0.43]$  mit Schrittweite 0.002 eingesetzt und das beste Resultat bezüglich des maximalen Fehlerbetrages bei  $\lambda_n = 0.43$  erzielt. Ein weiteres Gitter über  $[0.425, 0.435]$  mit Schrittweite 0.001 bestätigte  $\lambda_n^{opt} = 0.43$  als optimale Bandweite. Der Fehlerbetrag ist um ein bis zwei Größenordnungen kleiner als die Werte der Rekonstruktionsdichte, daher stellen wir den Fehlerbetrag in Abbildung 5.4 auf einer Skala von 0 bis 0.006 dar, um die Struktur des Rekonstruktionsfehlers hervorzuheben. In Abbildung 5.5 stellen wir den Fehlerbetrag dann zusätzlich auf der Skala  $[0, 0.1]$  dar, um einen direkten optischen Vergleich des Fehlers mit der Rekonstruktionsdichte zu ermöglichen.

Abbildung 5.6 zeigt das Ergebnis der Rekonstruktion für  $n = 10.000$ . Als optimale Bandweite wurde  $\lambda_n^{opt} = 0.38$  ermittelt, das damit erreichte Maximum des Fehlerbetrages liegt bei 0.004605. Es wurden hier zwei Gitter verwendet, ein Gitter über  $[0.22, 0.34]$  und ein zweites Gitter über  $[0.34, 0.46]$ , jeweils mit Schrittweite 0.02.

Abbildung 5.8 zeigt das Rekonstruktionsergebnis für  $n = 50.000$ . Als optimale Bandweite wurde  $\lambda_n^{opt} = 0.32$  bestimmt. Das Maximum des Fehlerbetrages liegt dann bei 0.002807. Es wurde ein Gitter über  $[0.28, 0.35]$  mit Schrittweite 0.01 eingesetzt. Für  $n = 50.000$  wurde die Bestimmung von  $\lambda_n^{opt}$  auf ein einziges Gitter mit kleiner Schrittweite beschränkt, da zum Zeitpunkt der Rechnung lediglich ein sequentieller Rechner zur Verfügung stand, auf dem die Rechenzeit für eine

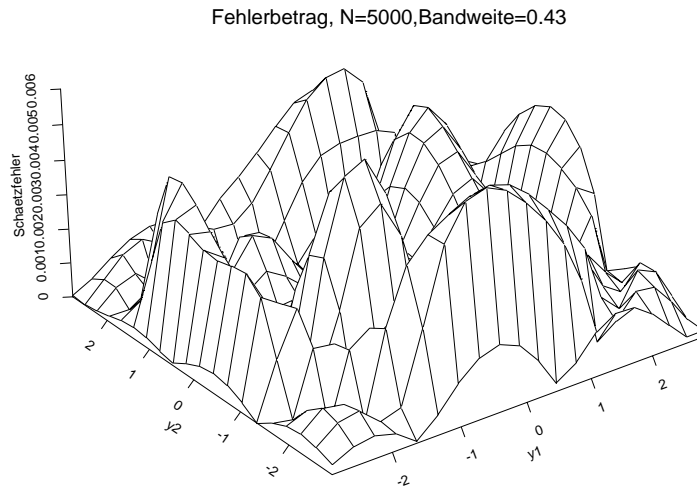


Abbildung 5.4: Rekonstruktionsfehler  $|\hat{f}_n(\mathbf{y}) - f_{2,8}(\mathbf{y})|$  bei  $n = 5000$ .

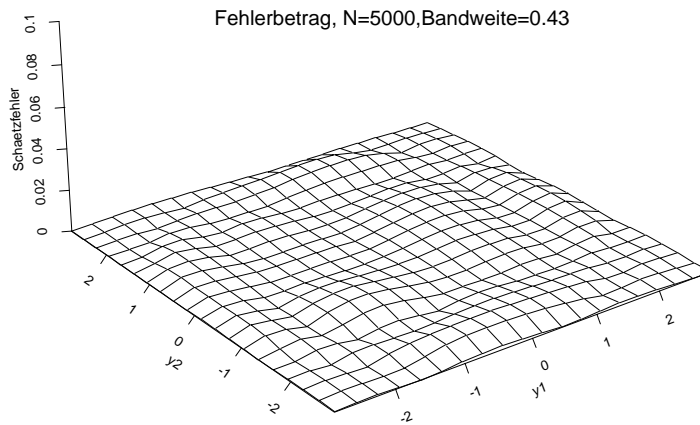


Abbildung 5.5: Rekonstruktionsfehler  $|\hat{f}_n(\mathbf{y}) - f_{2,8}(\mathbf{y})|$  bei  $n = 5000$ , große Skala.

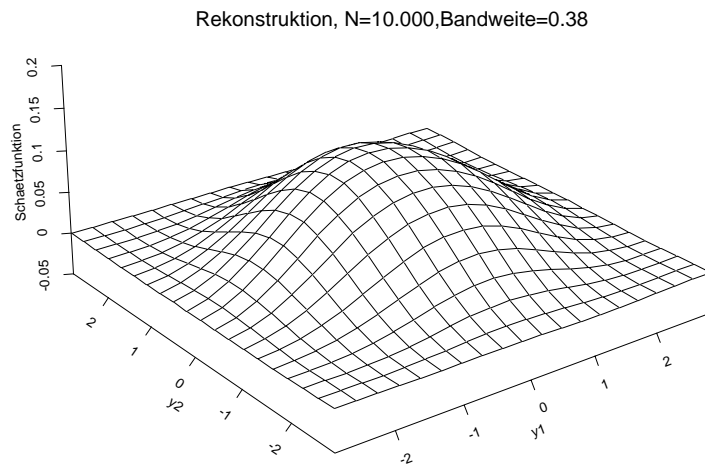


Abbildung 5.6: Rekonstruktion der Pearson-Typ-II Dichte  $f_{2,8}(\mathbf{y})$  aus 10.000 Datenpunkten.

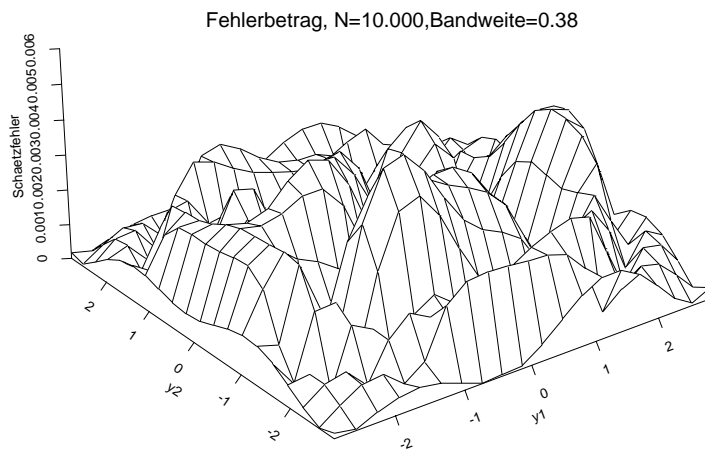


Abbildung 5.7: Rekonstruktionsfehler  $|\hat{f}_n(\mathbf{y}) - f_{2,8}(\mathbf{y})|$  bei  $n = 10.000$ .

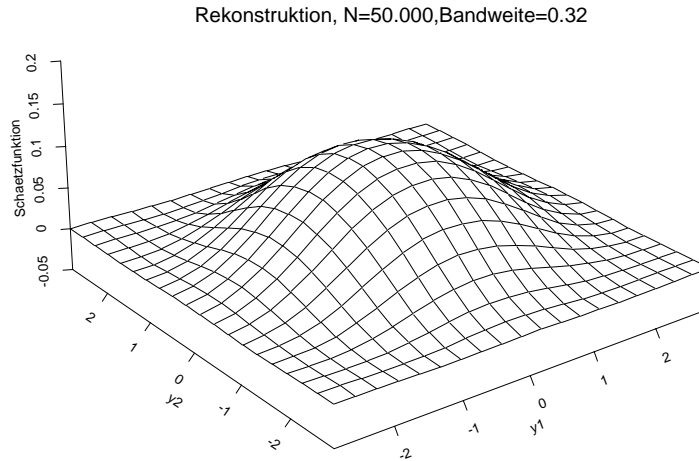


Abbildung 5.8: Rekonstruktion der Pearson–Typ–II Dichte  $f_{2,8}(\mathbf{y})$  aus 50.000 Datenpunkten.

Rekonstruktion bei dieser Größe des Datensatzes bereits etwa 24 Stunden betrug.

Abbildung 5.10 stellt das Rekonstruktionsergebnis für  $n = 75.000$  dar. Die ermittelte optimale Bandweite ist  $\lambda_n^{opt} = 0.30$ , das Maximum des Fehlerbetrages ist dann 0.002427. Es wurde dabei zunächst ein Gitter über  $[0.2, 0.3]$  mit Schrittweite 0.02 eingesetzt und dann, da das optimale  $\lambda_n$  am Rand des Gitters lag, ein zweites Gitter über  $[0.29, 0.31]$  mit Schrittweite 0.01. In Abbildung 5.11 und 5.12 ist der Fehlerbetrag für dieses Rekonstruktionsergebnis über der Skala  $[0, 0.006]$  beziehungsweise  $[0, 0.1]$  dargestellt, um einen direkten Vergleich mit den Ergebnissen für  $n = 5000$  zu erleichtern.

Die erzielten Rekonstruktionsergebnisse sind bereits für kleine  $n$  sehr gut, was dadurch begünstigt wird, daß die Standardabweichung der Fehlergröße  $r$  als  $\sigma = 0.02$ , und damit klein im Vergleich zur Segmentlänge 0.3 des Gitters über  $\mathcal{M}' = [-3, +3] \times [-3, +3]$  gewählt wurde.

Tabelle 5.1 enthält eine Auflistung des optimal erreichten  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}})|$  und des jeweils verwendeten  $\lambda_n^{opt}$  für die Fälle  $n = 5000, 10.000, 50.000, 75.000$ .

Abbildung 5.13 stellt  $\ln(\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}})|)$  über  $\ln(n)$  dar. An die vier

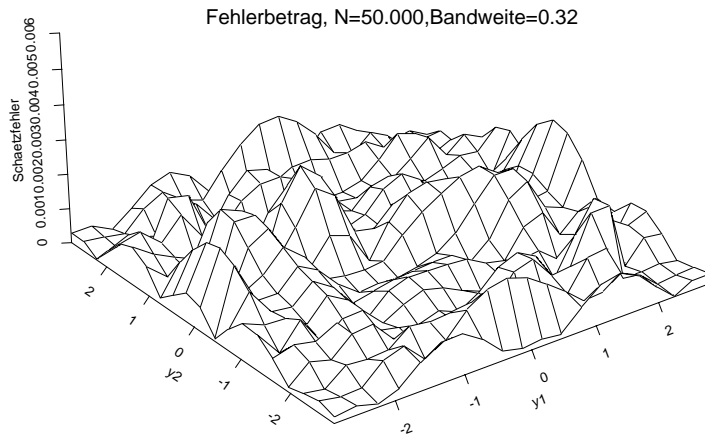


Abbildung 5.9: Rekonstruktionsfehler  $|\hat{f}_n(\mathbf{y}) - f_{2,8}(\mathbf{y})|$  bei  $n = 50.000$ .

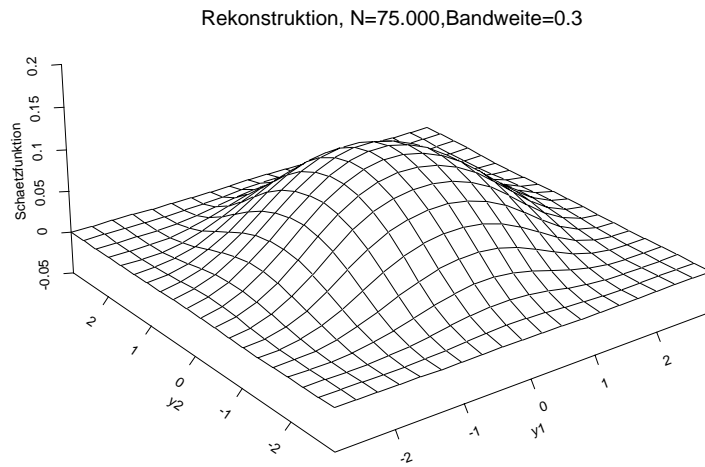


Abbildung 5.10: Rekonstruktion der Pearson-Typ-II Dichte  $f_{2,8}(\mathbf{y})$  aus 75.000 Datenpunkten.

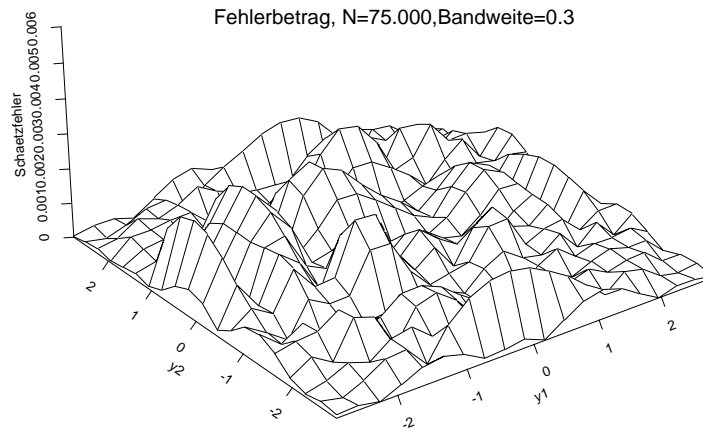


Abbildung 5.11: Rekonstruktionsfehler  $|\hat{f}_n(\mathbf{y}) - f_{2,8}(\mathbf{y})|$  bei  $n = 75.000$ .

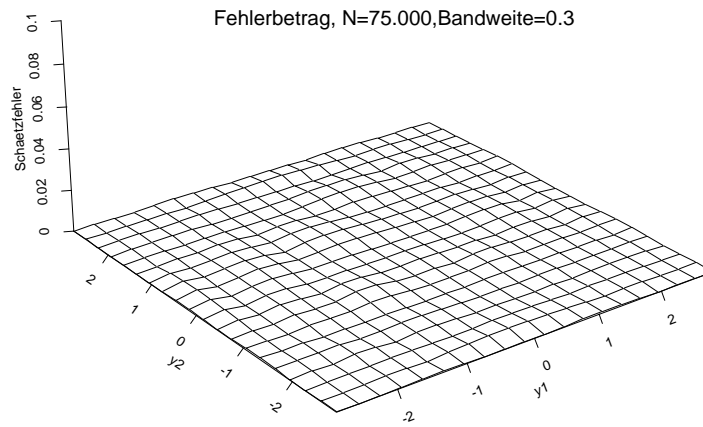


Abbildung 5.12: Rekonstruktionsfehler  $|\hat{f}_n(\mathbf{y}) - f_{2,8}(\mathbf{y})|$  bei  $n = 75.000$ , große Skala.



Tabelle 5.1: Darstellung von  $\lambda_n^{opt}$  und zugehörigem  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}})|$  für verschiedene  $n$ .

$n$	$\lambda_n^{opt}$	$\max_{i,j=0,\dots,\kappa-1}  \hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}}) $
5000	0.43	0.005859
10.000	0.38	0.004605
50.000	0.32	0.002807
75.000	0.3	0.002427

zur Verfügung stehenden Beobachtungspunkte wird eine Regressionsgerade angepaßt. Die Steigung dieser Geraden gibt eine Rate  $r_{exp}$  an, welche das Abklingen des maximalen Fehlerbetrages mit wachsendem  $n$  bei experimentell ermitteltem  $\lambda_n^{opt}$  beschreibt. Die Regression wurde mit Hilfe des Programm-Paketes *Mathematica* durchgeführt. Das Ergebnis der Regression ist in Abbildung 5.14 zu sehen. Es wird die Rate  $r_{exp} = 0.32$  bestimmt.

Die experimentell ermittelte Rate  $r_{exp}$  wird nun mit den theoretischen Ergebnissen zur oberen Schranke der Supremum-Norm für  $\hat{f}_n(\mathbf{y})$  aus Kapitel 3 verglichen. In Abschnitt 5.2 wurde gezeigt, daß  $f_{2,8}$  Element eines Sobolev-Raumes mit Parameter  $\alpha < 2.5$  ist. Dies zusammen mit den Ergebnissen aus Abschnitt 3.2.3 liefert die folgende theoretische Konvergenzordnung

$$\sup_{\mathcal{M}'} |\hat{f}_n(\mathbf{y}) - f_{2,8}(\mathbf{y})| = O\left(n^{-(0.25-\gamma)} (\ln(n))^{0.25-\gamma}\right) \quad \text{mit Wahrscheinlichkeit 1.}$$

Hierbei ist  $\gamma > 0$ . Die in den Simulationen ermittelte Rate für  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}})|$  ist im Einklang mit dieser Konvergenzordnung. In den Beispielrechnungen klingt die Supremum-Norm sogar etwas schneller ab.

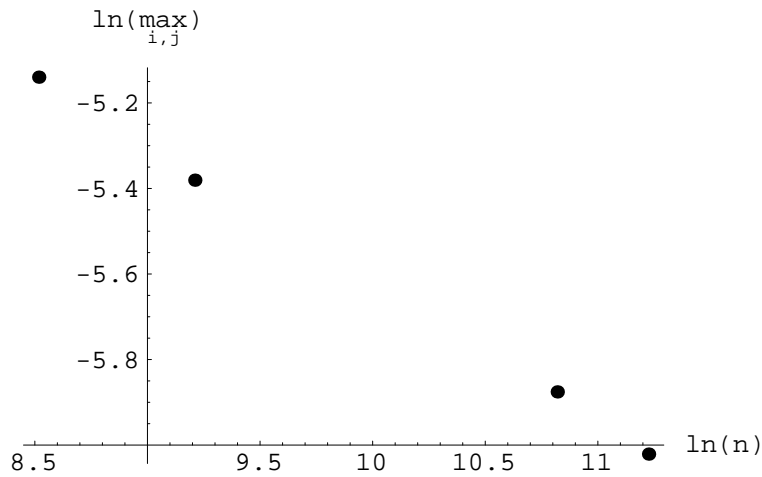


Abbildung 5.13: Darstellung von  $\ln(\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}})|)$  über  $\ln(n)$ .

	Estimate	SE	TStat	PValue
1	-2.41426	0.0556221	-43.4047	0.000530372
x	-0.320826	0.00555911	-57.7118	0.000300106

RSquared 0.9994,  
AdjustedRSquared 0.9991, EstimatedVariance 0.000153981, ANOVATable

	DF	SumOfSq	MeanSq	FRatio	PValue
Model	1	0.512857	0.512857	3330.66	0.000300106
Error	2	0.000307961	0.000153981		
Total	3	0.513165			

Abbildung 5.14: Anpassung einer Regressionsgeraden an  $(\ln(\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}})|), \ln(n))$  mit *Mathematica*.

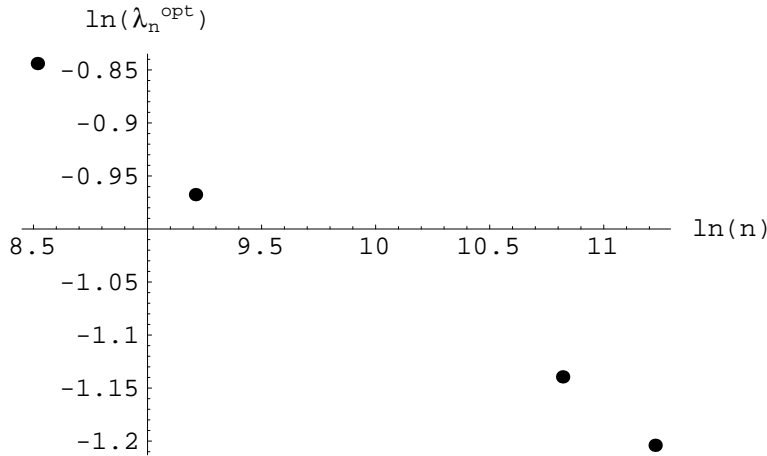


Abbildung 5.15: Darstellung von  $\ln(\lambda_n^{opt})$  über  $\ln(n)$ .

Ein Vergleich der experimentell ermittelten Rate des maximalen Fehlerbetrages mit der optimalen Konvergenzgeschwindigkeit unter der Supremum-Norm, welche in Abschnitt 4.5 gleichmäßig über  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  ermittelt wurde, ergibt, daß die experimentell ermittelte Konvergenzrate größer ist. Die untere Schranke für die Supremum-Norm wurde in Abschnitt 4.5 gleichmäßig über ganz  $\mathcal{H}^{\alpha,2}(\mathcal{B})$  ermittelt. Für ein festes  $f$  kann diese Schranke unterschritten werden.

In gleicher Weise kann man die Rate, mit welcher das experimentell bestimmte  $\lambda_n^{opt}$  abklingt mit der theoretischen optimalen Konvergenzrate von  $\lambda_n$  aus Satz 3.10 vergleichen. Dazu wird zunächst  $\ln(\lambda_n^{opt})$  über  $\ln(n)$  abgebildet. Dies ist in Abbildung 5.15 geschehen. Die Anpassung einer Regressionsgeraden ergibt die Abklingrate  $r_{exp,\lambda} = 0.13$  für  $\lambda_n^{opt}$ , siehe Abbildung 5.16. Aus Satz 3.10 mit  $\alpha < 2.5$  ergibt sich eine theoretische Rate  $r_{theo,\lambda} = \frac{1}{6} + \gamma$ ,  $\gamma > 0$ .

Die experimentell ermittelten Raten sind von der gleichen Größenordnung wie die theoretischen Raten. Gründe für die beobachteten Differenzen zwischen den theoretischen und den experimentell ermittelten Raten sind

1. die Ungenauigkeit bei der Approximation von  $\lambda_n^{opt}$ ;
2. die Approximation von  $\sup_{\mathcal{M}'} |\hat{f}_n(\mathbf{y}) - f(\mathbf{y})|$  durch  $\max_{i,j=0,\dots,\kappa-1} |\hat{f}_n(\mathbf{y}_{M_{ij}}) - f(\mathbf{y}_{M_{ij}})|$ ;
3. die Implementierung des modifizierten Dekonvolutions-Kernschätzers, bei welcher einige Diskretisierungen auftreten, was zu zusätzlichen Ungenauigkeiten führt.

	Estimate	SE	TStat	PValue	
1	0.218455	0.0887718	2.46086	0.132975	,
x	-0.126438	0.00887223	-14.251	0.00488781	
RSquared	0.990248	AdjustedRSquared	0.985372		
EstimatedVariance	0.000392213 , ANOVATable				
	DF	SumOfSq	MeanSq	FRatio	PValue
Model	1	0.0796553	0.0796553	203.092	0.004887
Error	2	0.000784426	0.000392213		
Total	3	0.0804397			

Abbildung 5.16: Regression von  $\ln(\lambda_n^{opt})$  über  $\ln(n)$  mit *Mathematica*.

## 5.5 Ausblick

Für die tatsächliche Anwendung des Schätzverfahrens in der Praxis muß ein theoretisches Verfahren entwickelt werden, um die optimale Bandweite  $\lambda_n^{opt}$  aus einem gegebenen Datensatz zu bestimmen. Dies könnte zum Beispiel ein Kreuzvalidierungsverfahren - angepaßt an das vorliegende Schätzproblem - sein. Ein Überblick über bestehende Verfahren zur automatischen Bandweitenwahl kann in [45] und [46], sowie in [57], [24], [32], [23] gefunden werden. Jede Wahl von  $\lambda_n$  hebt bestimmte Eigenschaften der Dichte  $f(\mathbf{y})$  hervor. Eine große Bandweite vermittelt einen Eindruck z. B. über das Vorhandensein von Clustern oder über die Schiefe der Funktion. Kleine Bandweiten betonen tatsächlich vorhandene aber auch künstliche Details in der Struktur von  $f(\mathbf{y})$ . Um Funktionen mit großen Intensitätsschwankungen zu rekonstruieren, wäre es daher von Vorteil, einen räumlich veränderlichen Parameter  $\lambda_n$  in Betracht zu ziehen.

Eine weitere Möglichkeit, das Schätzverfahren weiterzuentwickeln, ist die Verwendung von Vorwissen über die Struktur des Gehirnes. Da man bei PET-Untersuchungen nicht nach anatomischen Veränderungen wie etwa Tumoren sucht, kann man Annahmen treffen über die innere Struktur des Gehirnes. Man legt dabei die einzelnen Hauptregionen  $\mathcal{R}_i$  fest und nimmt an, daß sich innerhalb dieser Regionen die Emissionsdichte nicht extrem ändert. Man nimmt ferner an, daß die Regionen unabhängig voneinander betrachtet werden können, daß also Daten, die zu einer Region  $\mathcal{R}_i$  gehören, nicht für die Berechnung der Emissionsdichte innerhalb der anderen Regionen relevant sind. Da bei ToFPET die Beobachtungsdaten in Form von Punkten aus  $\mathbb{R}^2$  vorliegen, kann man den Datensatz aufteilen in mehrere zu den einzelnen Regionen gehörende Datensätze. Die Berechnung

der Emissionsdichte an der Stelle  $\mathbf{y} \in \mathcal{R}_i$  kann dann für alle Regionen  $\mathcal{R}_i$  parallel unter Verwendung des reduzierten Datensatzes geschehen. Dies wird zu einer erheblichen Reduzierung der Rechenzeit beitragen.

Bei dieser Vorgehensweise besteht auch die Möglichkeit, für jede Region eine separate Emissionsdichte anzunehmen und diese mit dem entsprechenden eingeschränkten Datensatz zu rekonstruieren. Dies führt dazu, daß man die Existenz starker Kontraste an den Gebietsrändern nicht mehr mitmodelliert. Man kann daher für jede Region  $\mathcal{R}_i$  eine Funktionenklasse wählen, die hohe Glattheitsanforderungen erfüllt und ein passendes  $\lambda_n$  einsetzen. Damit kann die Konvergenzrate des Schätzers gegen die tatsächliche Funktion schneller werden, was bei gegebener Genauigkeitsanforderung einen geringeren Datenbedarf bedeutet.

Bei der Aufteilung des Datensatzes vernachlässigt man die Existenz des Beobachtungsfehlers  $\mathbf{Z}$ . Dies führt zu Ungenauigkeiten hauptsächlich an den Gebietsgrenzen. Das Ausmaß dieses Effektes müsste quantitativ untersucht werden.

# Anhang: Parallele Implementierung des Schätzverfahrens

## A.1 Einleitung

Da der vorliegende Algorithmus die Berechnung einer Schätzfunktion für eine große Anzahl von Punkten  $\mathbf{y} \in \mathcal{M}'$  und für ein Gitter von  $\lambda_n$ -Werten erfordert, ist es günstig eine Parallelisierung der Berechnungen in Erwägung zu ziehen. Parallelisierung eines Algorithms bedeutet, daß ein Programm gleichzeitig von  $p$  Prozessoren bearbeitet wird, was zu einer Reduzierung der Rechenzeit führen soll. Idealerweise soll hierbei die Kommunikation zwischen den Prozessoren (Datenaustausch) gering gehalten werden.

Das hier entwickelte Schätzverfahren liefert einen guten Ausgangspunkt für eine Parallelisierung. Parallelisierung kann in folgender Weise stattfinden.

- a) Über der Menge der Gitterpunkte  $\mathbf{y}_{M_{ij}}$ ,  $i, j = 0, \dots, \kappa - 1$  aus  $\mathcal{M}'$ . Die Menge der Gitterpunkte wird dann aufgeteilt in  $p$  Teilbereiche, die parallel abgearbeitet werden. Für die hier vorgestellten Ergebnisse betrug die verwendete Gitterfeinheit 20 und es ergaben sich  $(20)^2$  Gitterpunkte  $\mathbf{y}_{M_{ij}}$ . Es sei  $(20)^2/p \in \mathbb{N}$ , dann besteht jeder Teilbereich aus  $(20)^2/p$  Datenpunkten.
- b) Über der Datenmenge. Die Berechnung der Schätzfunktion an einer Stelle  $\mathbf{y}_{M_{ij}}$  aus  $n$  Datenpunkten wird dabei in die Berechnung von  $p$  Teilsummen aufgeteilt.
- c) Über dem  $\lambda_n$ -Gitter. Jeder Prozessor übernimmt die Berechnung für Schätzfunktion und Rekonstruktionsfehler für einen Teilbereich der spezifizierten  $\lambda_n$ 's.

Im nächsten Abschnitt ist ein Beispiel für eine Parallelisierung der Berechnung von  $\hat{f}_n(\mathbf{y})$  dargestellt. In diesem Beispiel erfolgte die Parallelsierung über der

Menge der Gitterpunkte. Beim Testen des parallelen Programmcodes auf einem Parix-Parallelrechner mit  $p$  Prozessoren ergab sich eine Reduzierung der Programmlaufzeit um einen Faktor  $p$  im Vergleich zum entsprechenden sequentiellen Programmablauf. Damit ist die optimale Verkürzung der Rechenzeit erreicht, was zeigt, daß der vorliegende Schätzalgorithmus optimal vom Einsatz paralleler Programmierverfahren profitieren würde.

Die Struktur der Parallelisierung ist wie folgt:

- Masterprogramm (Prozessor 0):
  - Einlesen externer Daten:  $n$  Zufallsdaten  $(X_k, \varphi_k)$ , Speichern als globale Größen **Data.Xdat**, **Data.phifeld**.
  - Spezifikation globaler Parameter, z. B. der Feinheit des Gitters über dem Ortsraum  $\mathcal{M}' : 20 \times 20$ .
  - Für Prozessor 1: Berechnen des Schätzwertes  $\hat{f}_n(\mathbf{y}_{M_{ij}})$  für die Einträge **Yfeld** [0] bis **Yfeld**  $[\frac{(20)^2}{p} - 1]$ .
  - Exportieren der berechneten Werte in eine externe Datei „zkoord“.
- Prozessor  $i, i = 1 \cdots p - 1$ :
  - Empfangen der Daten von Prozessor  $i - 1$ .
  - Weiterleiten der globalen Daten an Prozessor  $i + 1$  (außer bei Prozessor  $p - 1$ ).
  - Berechnen des Schätzwertes für die Einträge **Yfeld**  $[i \cdot \frac{(20)^2}{p}]$  bis **Yfeld**  $[(i + 1) \cdot \frac{(20)^2}{p} - 1]$ .
  - Sukzessive Ausgabe der Ergebnisse, wobei die Prozessoren nacheinander die berechneten Ergebnisse in eine Datei schreiben. Dies erfolgt in einem Format, welches für die Visualisierung der Ergebnisse mit Splup geeignet ist.

Der vollständige Programmcode für eine parallelisierte Berechnung ist im nächsten Abschnitt aufgelistet. Der sequentielle Code ist nicht mehr extra aufgeführt, da er aus dem parallelen Code einfach abgeleitet werden kann. Folgende Prozeduren wurden verwendet.

- *main.c*, das Hauptprogramm, welches gleichzeitig von allen Prozessoren bearbeitet wird.

- *Tfeld.c*, eine Unterprozedur, welche ein Gitter fester Segmentlänge 0.1 im Fourierraum erzeugt.
- *Yfeld.c*, ein Unterprogramm, welches ein Gitter der Feinheit  $\kappa = 20$  über dem Gehirnraum erzeugt.
- *Mfeld.c*, eine Unterprozedur zur Berechnung der Funktion  $M(\mathbf{t})$  an den in *Tfeld.c* erzeugten Punkten.
- *FourIntEst.c*, die Prozedur zur Berechnung des Schätzwertes  $\hat{f}_n(\mathbf{y}_{M_{ij}})$  an einer Stelle  $\mathbf{y}_{M_{ij}} \in \mathcal{M}'$ .
- *init1d.c*, *init2d.c*, *init3d.c*, Prozeduren zur dynamischen Speicherplatzverwaltung. Sie erzeugen jeweils Zeiger auf ein-, zwei- und dreidimensionale Felder beliebiger Größe.

Bei Verwendung von 4 Prozessoren lautet der Programmaufruf mittels eines Parix-Rechners wie folgt:

```
RUN -A P4 2 2 MAINPARA
```

## A.2 C-Programmcode für einen Parix-Rechner

Um das Programm auf einem Parix-Rechner zu betreiben, müssen die folgenden „include-Dateien“ verwendet werden:

```
stdio.h
malloc.h
sys/memory.h
math.h
„defp.h“
„Funktp.h“
sys/root.h
sys/link.h
stdlib.h
string.h
sys/logerror.h
```

Dabei sind *defp.h* und *Funkt.h* „Headerfiles“, welche selbstdefinierte Datentypen und Funktionen enthalten. In *defp.h* wird auch  $n$ , die Größe des Datensatzes, festgelegt. *Funkt.h* enthält die Funktionendeklarationen aller Unterprozeduren.



### Definitionsdatei defp.h:

```
typedef double* zdou; typedef double** zzdou;
typedef double*** zzzdou;
typedef struct {zzdou Xdat;zdou phifeld;zdou t1dat, t2dat;zzdou Mdat;zzdou Y;
double Intk;} datpaket;
/* Fuer das Beispiel n=100000: */
/* (im Programm bezeichnet N die Größe des Datensatzes)*/
#define N 100000
#define sigma 0.02
#define Pearsig 2.828
#define Pearm 2
#define Gamma 3
```

### Definitionsdatei Funktp.h:

```
void Tfeld(zdou t1, zdou t2,double intber, int n1, int n2);
void Yfeld(zzzdou Yvec,int nrad,int nphi,double Durch);
void Mfeld(zzdou Mt,zdou t1,zdou t2,int n1,int n2);
double FourIntEst(zzdou Mt,zdou t1, zdou t2,int n1,int n2, zzdou Xdat,zdou phi,
double y[2]);
zzzdou init3d(int dim1,int dim2, int dim3);
zzdou init2d(int dim1,int dim2);
zdou init1d(int dim1);
```

### Variablenliste für mainpara.c:

```
const float pi=3.1415926;
zdou tf1,tf2;
zzdou Mt;
zzzdou Yvec;
zzdou fyfeld;
zzdou Xfeld,schaetzer,error;
zdou y;
zdou lambarr, intarr;
double yo[2],y1[2];
double Durch;
double feinhphi,feinhr;
double k, fra;
double test1, test, supremum;
int i,j,nx,ny,n1,n2;
```

```

int a,f,c1,c2,ein;
int v,t,s,p;
int va,v1,v2,u,e,g,h,b;
int fclose();
int end1,end2,end3,end4,end5,end6, end12, end22, end62, end63;
FILE *fopen(),*fpx,*fpy, *fpp;
FILE *feinx,*feiny,*feinz,*feinerr, *feinsup, *fcomm;
static char *xdat = „Pearsonx.1“;
static char *ydat = „Pearsony.1“;
static char *phidat = „Winkel.1“;
static char *xval = „/home/silke/local/xkoord“;
static char *yval = „/home/silke/local/ykoord“;
static char *zval = „/home/silke/local/zkoord“;
static char *errval=„/home/silke/local/abserror“;
static char *supval=„/home/silke/local/Supremum“;
static char *comm=„/home/silke/local/Communication“;
zdou xwert,ywert,phifeld;
double invsig, invsq, invpi, sigsq;
double lamb_low, lamb_hi;
double dlamb, deltat;
int lambtotal;
int err,Fl,teil,ok;
int MyProcID;
int nProcs;
datapaket Data;
char *FName = „main“;
LinkCB_t *NextLink, *PrevLink;

```

### **C-Programmcode für das Hauptprogramm „mainpara.c“:**

```

main() {
/*Variablenliste*/

/*****/
/* Das Main-Programm wird gleichzeitig von allen zur */
/* Verfügung stehenden Prozessoren bearbeitet. */
/* Die Prozessor-ID wird abgefragt, bei weniger als 2*/
/* Prozessoren, STOP */
/*****/

MyProcID = GET_ROOT() - > ProcRoot - > MyProcID;

```

```

printf(„%d“,MyProcID);
nProcs = GET_ROOT() -> ProcRoot -> nProcs;
if (nProcs <2) printf(„network too small \n“);
AbortServer (1);
if(MyProcID ==0)

/*****/
/* Das Masterprogramm, Verteilung der Daten */
/*****/
{
/*****/
/* 1.Schritt Einlesen der von S+ erzeugten Zufallsdaten*/
/*****/
xwert=init1d(N);
ywert=init1d(N);
phifeld=init1d(N);

s=0;
fpx=fopen(xdat,„r“);
while(c1=fscanf(fpx,„%lf“, &(xwert[s]))!=EOF)
s++;
end1=fclose(fpx);

t=0;
fpy=fopen(ydat,„r“);

while(c2=fscanf(fpy,„%lf“, &(ywert[t]))!=EOF)
t++;
end2=fclose(fpy);

p=0;
fpp=fopen(phidat, „r“);
while(c2=fscanf(fpp, „%lf“, &(phifeld[p]))!=EOF)
p++;
end3=fclose(fpp);

Data.Xdat=init2d(N,2);
for(va=0;va<N;va+=1) Data.Xdat[va][0]=xwert[va];
Data.Xdat[va][1]=ywert[va];

lamb_low=0.45;

```

```

lamb_hi=0.53;
dlamb=0.01;
lambtotal=9;
lambarr=init1d(lambtotal);
intarr=init1d(lambtotal);
for(f=0;f<lambtotal;f++)
{
lambarr[f]=lamb_hi-f*dlamb;
intarr[f]=(double) pow(lambarr[f],-1);
}

for(g=0;g<lambtotal;g++)
{
n1=0;
n2=0;
deltat=0.1;
n1=ceil(intarr[g]/deltat);
n2=2*n1;

fcomm = fopen(comm,,,"a");
fprintf(fcomm,,,"Integrationsbereich: t1 in [0, %f]\n",intarr[g]);
fprintf(fcomm,,,"t2 in [- %f,%f]\n",intarr[g],intarr[g]);
fprintf(fcomm,,,"Anzahl der Unterteilungen in t1 Richtung:%i\n", n1);
fprintf(fcomm,,,"Anzahl der Unterteilungen in t2 Richtung:%i\n\n", n2);
fclose(fcomm);

Data.t1dat=init1d(n1);
Data.t2dat=init1d(n2);
Tfeld(Data.t1dat,Data.t2dat,intber,n1,n2);

Data.Mdat=init2d(n1,n2);
Mfeld(Data.Mdat,Data.t1dat,Data.t2dat,n1,n2);
fra=(double) n1*n2*N*pi*pi;
Data.Intk=(double) pow(intarr[g],2) * pow(fra,-1);

/*****
/* 3.Schritt:Erzeugen eines Gitters im Gehirnraum */
*****/
Durch = 6;
nx=20;
ny=20;

```

```

Yvec=init3d(nx,ny,2);
Yfeld(Data.Y,nx,ny,Durch);
/*****/
/* Rausschreiben der Gitterpunkte fuer die S+-Visualisierung */
/*****/

feinx = fopen(xval,,w");
for(u=0;u<=(nx-1);u++)
{
fprintf(feinx,,"%f",Yvec[u][0][0]);
}
end4=fclose(feinx);

feiny = fopen(yval,,w");
for(u=0;u<=(ny-1);u++)
{
fprintf(feiny,,"%f",Yvec[0][u][1]);
}
end5=fclose(feiny);
printf(„Rausschreiben der Gitterpunkte in externes file beendet.“);

/*****/
/* Verschicken der Daten an die einzelnen Prozessoren*/
/*****/
NextLink=ConnectLink(MyProcID + 1,1234,&err);
if(NextLink == NULL)
{ LogError(EC_ERROR,FName,,Failed to connect link to 1,err“
„code &d“,err); exit(1); }
SendLink(NextLink,&n1,sizeof(int));
SendLink(NextLink,&n2,sizeof(int));
SendLink(NextLink,&*(Data.phifeld),N*sizeof(double));
for(j=0;j<N;j++)
{
SendLink(NextLink,Data.Xdat[j],2*sizeof(double));
}
SendLink(NextLink,&*(Data.t1dat),n1*sizeof(double));
SendLink(NextLink,&*(Data.t2dat),n2*sizeof(double));
for(i=0;i<20;i++)for(j=0;j<20;j++)
SendLink(NextLink,Data.Y[i][j],2*sizeof(double));
for(i=0;i<n1;i++)
SendLink(NextLink,Data.Mdat[i],n2*sizeof(double));

```

```

SendLink(NextLink,&Data.Intk,sizeof(double));
/*****/
/* Berechnung des Schaetzers auf einem Teilgitter ueber dem*/
/* Gehirnraum */
/*****/

teil=(2*nrad)*pow(nProcs,-1);

sigsq=(double) pow(Pearsig,2);
invsig=(double) pow(Pearsig,-1);
invsq=(double) pow(Pearsig,-2);
invpi=(double) pow(pi,-1);

schaetzer=init2d(teil,20);
error=init2d(teil,20);
supremum=0;

for(i=0;i<=(teil-1);i=i+1) for(j=0;j<20;j++)
{
y[0]=Data.Y[i][j][0]; y[1]=Data.Y[i][j][1];
schaetzer[i][j]=Data.Intk * FourIntEst(Data.Mdat,Data.t1dat,
Data.t2dat,n1,n2,Data.Xdat,Data.phifeld,y);

if((y[0]*y[0]+y[1]*y[1])<=sigsq)
{
error[i][j]=fabs( schaetzer[i][j] -
Gamma*invpi*invsq*pow((1-invsq*(pow(y[0],2)+pow(y[1],2))),Pearm));
}
else
{
error[i][j] = fabs(schaetzer[i][j]);
}

/*****/
/* Jeder Prozessor schreibt die Ergebnisse des jeweils */
/* bearbeiteten Teilbereiches in das file zkoord */
/*****/
feinz = fopen(zval,„w“);
feinerr = fopen(errval,„w“);
feinsup = fopen(supval, „w“);
for(u=0;u<=(teil-1);u++) for(g=0;g<=(2*nrad-1);g++)

```

```

{
fprintf(feinz,,,"%f",ergebnis[u][g]);
fprintf(feinerr,,,"%f",error[u][g]);
}

end6=fclose(feinz);
end62=fclose(feinerr);
end63=fclose(feinsup);
Fl=1;
SendLink(NextLink,&Fl,sizeof(int));
}
/*****/
/* Die uebrigen Prozessoren */
/*****/
elseif (MyProcID < nProcs-1)
PrevLink=ConnectLink(MyProcID-1,1234,&err);
NextLink=ConnectLink(MyProcID+1,1234,&err);
RecvLink(PrevLink,&n1,1024);
RecvLink(PrevLink,&n2,1024);
RecvLink(PrevLink,&nrad,1024);
/*****/
/*Initialisierung der Datenstruktur */
/*****/
Data.phifeld=init1d(N);
Data.Xdat=init2d(N,2);
Data.t1dat=init1d(n1);
Data.t2dat=init1d(n2);
Data.Y=init3d(20,20,2);
Data.Mdat=init2d(n1,n2);
/*****/
/* Belegen der Datenstruktur */
/*****/
RecvLink(PrevLink,&(*(Data.phifeld)),N*sizeof(double));
for(i=0;i<N;i++)
RecvLink(PrevLink,Data.Xdat[i],2*sizeof(double));
RecvLink(PrevLink,&(*(Data.t1dat)),n1*sizeof(double));
RecvLink(PrevLink,&(*(Data.t2dat)),n2*sizeof(double));
for(i=0;i<20;i++) for(j=0;j<20;j++)
RecvLink(PrevLink,Data.Y[i][j],2*sizeof(double));
for(i=0;i<n1;i++)
RecvLink(PrevLink,Data.Mdat[i],n2*sizeof(double));

```

```

RecvLink(PrevLink,&Data.Intk,sizeof(double));
SendLink(NextLink,&n1,sizeof(int));
SendLink(NextLink,&n2,sizeof(int));
SendLink(NextLink,&*(Data.phifeld),N*sizeof(double));
for(j=0;j<N;j++)
SendLink(NextLink,Data.Xdat[j],2*sizeof(double));
SendLink(NextLink,&*(Data.t1dat),n1*sizeof(double));
SendLink(NextLink,&*(Data.t2dat),n2*sizeof(double));
for(i=0;i<20;i++) for(j=0;j<20;j++)
SendLink(NextLink,Data.Y[i][j],2*sizeof(double));
for(i=0;i<n1;i++)
SendLink(NextLink,Data.Mdat[i],n2*sizeof(double));
SendLink(NextLink,&Data.Intk,sizeof(double));
teil=20*pow(nProcs,-1);
schaetzer=init2d(teil,20);
schaetzer2=init2d(teil,20);
error=init2d(teil,20);

sigsq=(double) pow(Pearsig,2);
invsig=(double) pow(Pearsig,-1);
invsq=(double) pow(Pearsig,-2);
invpi=(double) pow(pi,-1);

/*****
/* Berechnen der Schaetzwerte fuer den zugeordneten Teilbereich*/
*****/
for(i=MyProcID * teil;i<(MyProcID +1)*teil;i++)
for(j=0;j<20;j++)
{ y[0]=Data.Y[i][j][0];
y[1]=Data.Y[i][j][1]; e=i-MyProcID*teil;
schaetzer[e][j]=Data.Intk * FourIntEst(Data.Mdat,Data.t1dat,
Data.t2dat,n1,n2,Data.Xdat,Data.phifeld,y);

if((y[0]*y[0]+y[1]*y[1])<=sigsq)
{
error[e][j]=fabs( schaetzer[e][j] -
Gamma*invpi*invsq*pow((1-invsq*(
pow(y[0],2)+pow(y[1],2))),Pearn));
}
else
{

```



```

error[e][j] = fabs(schaetzer[e][j]);
}

}

RecvLink(PrevLink,&F1,1024);
if(F1==1)
{ feinz = fopen(zval,,a");
feinerr = fopen(errval,,a");
for(u=0;u<4=(teil-1);u++) for(g=0;g<=(20-1);g++)
{
fprintf(feinz,,"%f",schaetzer[u][g]);
fprintf(feinerr,,"%f",error[u][g]);
}
}

end7=fclose(feinz);
end72=fclose(feinerr);
end73=fclose(feinsup);

F1=1;
SendLink(NextLink,&F1,sizeof(int));
}

if(MyProcID==nProcs-1){
PrevLink=ConnectLink(MyProcID - 1,1234,&err);
if(PrevLink==NULL)printf(,"failed to connect to previous
processor, err %d",err); exit(1); }
RecvLink(PrevLink,&n1,1024);
RecvLink(PrevLink,&n2,1024);
Data.phifeld=init1d(N);
Data.Xdat=init2d(N,2);
Data.t1dat=init1d(n1);
Data.t2dat=init1d(n2);
Data.Y=init3d(20,20,2);
Data.Mdat=init2d(n1,n2);
RecvLink(PrevLink,&*(Data.phifeld),N*sizeof(double));
for(i=0;i<N;i++)
RecvLink(PrevLink,Data.Xdat[i],2*sizeof(double));
RecvLink(PrevLink,&*(Data.t1dat),n1*sizeof(double));
RecvLink(PrevLink,&*(Data.t2dat),n2*sizeof(double));

```

```

for(i=0;i<2*nrad;i++) for(j=0;j<20;j++)
{ RecvLink(PrevLink,Data.Y[i][j],2*sizeof(double)); }
for(i=0;i<n1;i++)
RecvLink(PrevLink,Data.Mdat[i],n2*sizeof(double));
RecvLink(PrevLink,&Data.Intk,sizeof(double));

teil=2*20*pow(nProcs,-1);
schaetzer=init2d(teil,20);
error=init2d(teil,20);

sigsq=(double) pow(Pearsig,2);
invsig=(double) pow(Pearsig,-1);
invsq=(double) pow(Pearsig,-2);
invpi=(double) pow(pi,-1);

/*****
/* Berechnen der Schaetzwerte fuer den zugeordneten Teilbereich*/
*****/
for(i=MyProcID * teil;i<(MyProcID +1)*teil;i++)
for(j=0;j<20;j++)
{ y[0]=Data.Y[i][j][0];
y[1]=Data.Y[i][j][1]; e=i-MyProcID*teil;
schaetzer[e][j]=Data.Intk * FourIntEst(Data.Mdat,Data.t1dat,
Data.t2dat,n1,n2,Data.Xdat,Data.phifeld,y);

if((y[0]*y[0]+y[1]*y[1])<=sigsq)
{
error[e][j]=fabs( schaetzer[e][j] -
Gamma*invpi*invsq*pow((1-invsq*(
pow(y[0],2)+pow(y[1],2))),Pearn));
}
else
{
error[e][j] = fabs(schaetzer[e][j]);
}

RecvLink(PrevLink,&Fl,1024);
if(Fl==1)
{ feinz = fopen(zval,,,"a");
feinerr = fopen(errval,,,"a");
for(u=0;u<=(teil-1);u++) for(g=0;g<4=(2*20-1);g++)

```

```

{
fprintf(feinz,,,"%f", ergebnis[u][g]);
fprintf(feinerr,,,"%f",error[u][g];
}
end8=fclose(feinz);
end82=fclose(feinerr);
end83=fclose(feinsup);
}
exit(0);
} }

```

Es folgt der Programmcode der Unterprozeduren.

### Unterprozedur Tfeld.c:

```

void Tfeld(zdou T1,zdou T2,double ib,int n1 ,int n2)
{ double ft1,ft2; int i,j;
ft1=ib/n1; ft2=(2*ib)/n2;
for(i=0;i<=n1-1;i=i+1)
{ T1[i]=(i+0.5)*ft1; }
for(j=0;j<=n2-1;j=j+1)
{ T2[j]=ib-(j+0.5)*ft2; } }

```

### Unterprozedur Yfeld.c:

```

void Yfeld(zzzdou Yv,int nr,int np, double d)
{
int i,j; double feinhx, feiny;
double R;
feinhx=d/(nr);
feiny=d/(np);
R=0.5*d;
for(i=0;i<=nr-1;i++)
{
for(j=0;j<=np-1;j++)
{
Yv[i][j][0]=-R+(i+0.5)*feinhx;
Yv[i][j][1]=R-(j+0.5)*feiny; } } }

```

### Unterprozedur Mfeld.c:

```
void Mfeld(zzdou M,zdou T1,zdou T2,int dim1,int dim2)
{
const float Pi=3.1415926;
int i,j,t;
double sum;
k=Pi/100.0;
for(i=0;i<=dim1-1;i++) {
for(j=0;j<=dim2-1;j++) {
sum=0;
for(t=0;t<=99;t++) {
sum += k* exp(-pow(sigma,2)*pow(T1[i]*cos(t*(Pi/100.0))+
T2[j]*sin(t*(Pi/100.0)),2)); }
M[i][j]=sum; } } }
```

### Unterprozedur FourIntEst.c:

```
double FourIntEst(zzdou Mf,zdou t1,zdou t2,int
n1,int n2,zzdou Xfe,zdou phi,double y[2])
{
const float pi=3.1415926;
int i,j,t,fd;
double sum;
double hilf,d1,d2;
int test;
sum=0;
for(t=0;t<N;t++) { hilf=0;
for(i=0;i<=n1-1;i++) {
for(j=0;j<=n2-1;j++) {
d1=y[0]-Xfe[t][0];
d2=y[1]-Xfe[t][1];
hilf = hilf+ (cos(t1[i]*d1 + t2[j]*d2)
exp(-pow(t1[i]*cos(phi[t])+t2[j]*sin(phi[t]),2)*pow(sigma,2)*0.5)
(pow(Mf[i][j],-1))); } }
sum=sum+hilf; }
return(sum); }
```

## Danksagung

Mein herzlichster Dank gilt

Herrn Prof. Dr. Christian Hesse für die interessante Auswahl des Themas, für die Anregungen, die Betreuung und die Unterstützung bei der Entstehung dieser Arbeit.

Herrn Prof. Dr. Harro Walk, der sich freundlicherweise zur Übernahme des Mitberichts bereiterklärt hat.

Herrn Dr. Ing. habil. Fritz Schmidt, der die Mitbetreuung im Rahmen des Graduiertenkolleges „Parallele und Verteilte Systeme“ (GKPVS) übernahm.

Frau Dipl. math. Christine Ackermann für die kritische Durchsicht des Manuskripts.

Mein inniger Dank gilt meiner Familie und meinen Freunden, besonders meinen Eltern und meinem Mann Martin für die Unterstützung, die zur Entstehung dieser Arbeit notwendig war.

# Literaturverzeichnis

- [1] R. A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.
- [2] R. A. Adams and L. I. Hedberg, *Function spaces and potential theory*, *Grundlehren der mathematischen Wissenschaften 314*, Springer Verlag, 1996.
- [3] J. Arsac, *Fourier transforms and the theory of distributions*, Prentice-Hall, New Jersey, 1966.
- [4] P. J. Bickel and K. A. Doksum, *Mathematical statistics: Basic ideas and selected topics*, Holden-Day, San Francisco, 1977.
- [5] D. Billingsley, *Probability and measure*, Wiley, New York, 1979.
- [6] R. J. Carroll and P. Hall, *Optimal rates of convergence for deconvolving a density*, *JASA, Theory and Methods* **83** (1988), no. 404, 1184–1186.
- [7] Y. S. Chow and H. Teicher, *Probability theory*, Springer-Verlag, New York, 1978.
- [8] N. A. C. Cressie, *Statistics for spatial data*, Wiley, New York, 1991.
- [9] K. Bullock Davis, *Mean square error properties of density estimates*, *Annals of Statistics* **3** (1975), no. 4, 1025–1030.
- [10] ———, *Mean integrated square error properties of density estimates*, *Annals of Statistics* **5** (1977), no. 3, 530–535.
- [11] D. L. Donoho and R. C. Liu, *Geometrizing rates of convergence i*, Technical Report 137 a), Dept. of Statistics, Univ. California, Berkeley (1987).
- [12] ———, *Geometrizing rates of convergence ii*, *Annals of Statistics* **19** (1991), 633–667.
- [13] ———, *Geometrizing rates of convergence iii*, *Annals of Statistics* **19** (1991), 668–701.

- [14] D. L. Donoho and M. G. Low, *Renormalization exponents and optimal point-wise rates of convergence*, Annals of Statistics **20** (1992), no. 2, 944–970.
- [15] D. Zwillinger (ed.), *CRC-standard mathematical tables and formulae*, CRC Press, 1996.
- [16] V. A. Epanechnikov, *Nonparametric estimation of density functions*, Theor. Probability Appl. **14** (1969), 153–158.
- [17] J. Fan, *On the optimal global rates of convergence for nonparametric deconvolution problems*, The Annals of Statistics **19** (1991), no. 3, 1257–1272.
- [18] ———, *Adaptively local one-dimensional subproblems with application to a deconvolution problem*, The Annals of Statistics **21** (1993), no. 2, 600–610.
- [19] S. Geman and D. Geman, *Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI **6** (1984), no. 6, 721–741.
- [20] S. Geman and D. E. McClure, *Bayesian image analysis: An application to Single Photon Emission Tomography*, Proc. Amer. Statist. Assoc. Statist. Comp. Section (1985), 12–18.
- [21] W. C. Giffin, *Transform techniques for probability modeling*, Academic Press, New York, 1975.
- [22] R. Gordon, *A tutorial on ART (Algebraic Reconstruction Technique)*, IEEE Transactions on Nuclear Science NS **21** (1974), 78–93.
- [23] W. Hardle, J.S. Marron, and M.P.Wand, *Bandwidth choice for density derivatives*, J.R. Statist. Soc. B **52** (1990), no. 1, 223–232.
- [24] E. Herrmann, M.P. Wand, J. Engel, and T. Gasser, *A bandwidth selector for bivariate kernel regression*, J.R.Statist. Soc. B **57** (1995), no. 1, 171–180.
- [25] C. H. Hesse, *Deconvolving a density from partially contaminated observations*, Journal of Multivariate Analysis **55** (1995), no. 2, 246–260.
- [26] ———, *Distribution function estimation from noisy observations*, Publ. Inst. Stat. Paris Sud **39** (1995), 21–35.
- [27] ———, *Data-driven deconvolution*, J. Nonparametric Stat. **10** (1999), 343–373.
- [28] ———, *An iterative scheme for density estimation from contaminated observations*, Tech. report, Mathematisches Institut, Universität Stuttgart (1999).

- [29] I. A. Ibragimov and R. Z. Khasminskii, *Statistical estimation, asymptotic theory*, Springer, Berlin, 1981.
- [30] I. A. Ibragimov, A. S. Nemirovskii, and R. Z. Khasminskii, *Some problems on nonparametric estimation in Gaussian white noise*, Theory Probab. Appl. **31** (1986), 391–406.
- [31] I. M. Johnstone and B. W. Silverman, *Speed of estimation in Positron Emission Tomography and related inverse problems*, The Annals of Statistics **18** (1990), no. 1, 251–280.
- [32] J.S.Marron and W.J.Padgett, *Asymptotically optimal bandwidth selection for kernel density estimators from randomly right-censored samples*, Annals of Statistics **15** (1987), no. 4, 1520–1535.
- [33] L. LeCam, *Asymptotic methods in statistical decision theory*, Springer, New York, 1985.
- [34] A. K. Louis and F. Natterer (Ed.), *Mathematical methods in tomography: Proceedings of a conference held in Oberwolfach/Germany*, Springer Verlag, Berlin, New York, 1991.
- [35] M.E.Johnson, *Multivariate statistical simulation*, John Wiley and Sons, 1987.
- [36] F. Natterer, *A Sobolev space analysis of picture reconstruction*, SIAM J. Appl. Math. **39** (1980), 402–411.
- [37] ———, *The mathematics of computerized tomography*, Teubner, Stuttgart; Wiley, Chichester, 1986.
- [38] N.L.Johnson and S. Kotz, *Continuous univariate distributions - 2*, John Wiley and Sons, 1970.
- [39] S. Csörgő, *Multivariate empirical characteristic functions*, Zur Wahrscheinlichkeitstheorie verwandte Gebiete (1981), 203–229.
- [40] ———, *Rates of uniform convergence for the empirical characteristic function*, Acta Sci. Math. **48** (1985), 97–102.
- [41] S. Csörgő and V. Totik, *On how long interval is the empirical characteristic function uniformly consistent*, Acta Sci. Math. **45** (1983), 141–149.
- [42] E. Parzen, *Modern probability theory and its applications*, Wiley, New York, 1966.



- [43] E. A. Philippe and N. Mullani et al., *Real-time image reconstruction for Time-of-Flight Positron Emission Tomography (ToFPET)*, IEEE Transactions on Nuclear Science NS **29** (1982), no. 1, 524–527.
- [44] D. Pollard, *Convergence of stochastic processes*, Springer-Verlag, New York, 1984.
- [45] D. W. Scott, *Multivariate density estimation, theory, practice and visualization*, Wiley Series in Probability and Mathematical Statistics, New York, 1992.
- [46] B. W. Silverman, *Density estimation for statistics and data analysis*, Chapman and Hall, London, 1986.
- [47] D. L. Snyder, *Some noise comparisons of data-collection arrays for emission tomography systems having time-of-flight measurements*, IEEE Transactions on Nuclear Science NS **29** (1982), no. 1, 1029–1032.
- [48] D. L. Snyder and D. G. Politte, *Image reconstruction from list-mode data in an emission tomography system having time-of-flight measurements*, IEEE Transactions on Nuclear Science NS **20** (1983), no. 3, 1843–1848.
- [49] D. L. Snyder, L. J. Thomas, and M. M. Ter-Pogossian, *A mathematical model for PET systems having time-of-flight measurements*, IEEE Transactions on Nuclear Science NS **28** (1981), no. 3, 3575–3583.
- [50] L. Stefanski, *Rates of convergence of some estimators in a class of deconvolution problems*, Statistics and Probability Letters **9** (1990), 229–235.
- [51] L. Stefanski and R. J. Carroll, *Deconvoluting kernel density estimators*, Statistics **21** (1990), no. 2, 169–184.
- [52] C. J. Stone, *Optimal rates of convergence for nonparametric estimators*, The Annals of Statistics **8** (1980), no. 6, 1348–1360.
- [53] ———, *Optimal global rates of convergence for nonparametric regression*, The Annals of Statistics **10** (1982), no. 4, 1040–1053.
- [54] ———, *Optimal uniform rate of convergence for nonparametric estimators of a density function or its derivatives*, Recent Advances in Statistics (1983), 393–406.
- [55] R. A. Tapia and J. R. Thompson, *Nonparametric probability density estimation*, John Hopkins University Press, Baltimore, 1978.

- [56] H. Triebel, *Interpolation theory, function spaces, differential operators*, North Holland, 1978.
- [57] C. Werthenbach und E. Herrmann, *A fast and stable updating algorithm for bivariate nonparametric curve estimation*, J. Comp. and Graph. Statistics **7** (1998), no. 1, 61–76.
- [58] Y. Vardi, L. A. Shepp, and L. Kaufman, *A statistical model for Positron Emission Tomography*, JASA **80** (1985), no. 389, 8–37.
- [59] G. Wahba, *Optimal convergence properties of variable knot, kernel and orthogonal series methods for density estimation*, The Annals of Statistics **3** (1975), 15–29.
- [60] J. Wloka, *Partial differential equations*, Cambridge University Press, 1987.
- [61] H. P. Zhang, *On deconvolution using time-of-flight information in Positron Emission Tomography*, Statistica Sinica **2** (1992), 553–575.

# Lebenslauf

Silke Kuball, geborene Ludwig,

geboren am 11.07.1968 in Neunkirchen/Saar als Tochter von Gerd Ludwig und Margot Ludwig, geborene Schütz.

## Werdegang:

- |                 |   |
|-----------------|---|
| 1974–1978       | Grundschule in Wiebelskirchen   |
| 1978–1987       | Gymnasium am Steinwald in Neunkirchen<br>Abitur mit Auszeichnung  |
| 1987–1993       | Studium der Technomathematik, Universität Kaiserslautern  |
| 03/1993         | Diplom in Technomathematik mit Auszeichnung<br>Betreuer der Diplomarbeit: Prof. Dr. Jürgen Franke<br>Thema: <i>Eine Anwendung des Kalman-Filters zur Vorhersage der Bestandsentwicklung von Waldflächen</i> |
| 04/1993         | Aufnahme der Doktorarbeit bei Prof. Dr. Christian H. Hesse,<br>Mathematisches Institut A der Universität Stuttgart  |
| 04/1993–04/1996 | Stipendiatin des Graduiertenkollegs für Parallele<br>und Verteilte Systeme der Universität Stuttgart  |
| 04/1996–09/1997 | Aufenthalt in Providence, Rhode Island, USA;<br>Gast an der Brown University und Forschungsstudentin<br>am Fraunhofer Institut für Forschung in Computergraphik   |
| seit 10/1997    | Visiting Fellow in der Arbeitsgruppe Statistik,<br>Universität Bristol, UK, unter der Leitung von<br>Professor Peter Green  |
| seit 02/1998    | Anstellung als Research Associate;<br>Safety Systems Research Centre, Fachbereich Informatik,<br>Universität Bristol, UK.   |