
Estimating rigid motion in sparse sequential dynamic imaging: with application to nanoscale fluorescence microscopy

Dissertation zur Erlangung
des mathematisch-naturwissenschaftlichen Doktorgrades
„Doctor rerum naturalium“
der Georg-August-Universität Göttingen
im Promotionsstudiengang „Mathematical Sciences“
der Georg-August University School of Science (GAUSS)

vorgelegt von
Alexander Hartmann
aus Bad Langensalza

Göttingen, 2017

Betreuungsausschuss

Prof. Dr. Axel Munk¹

Prof. Dr. Stephan Huckemann¹

Mitglieder der Prüfungskommission

Referent: Prof. Dr. Axel Munk¹

Korreferent: Prof. Dr. Stephan Huckemann¹

Weitere Mitglieder der Prüfungskommission:

P.D. Dr. Timo Aspelmeier¹

Dr. Alexander Egner²

Prof. Dr. Stefan Halverscheid³

Prof. Dr. Russell Luke⁴

Tag der mündlichen Prüfung: 22.04.2016

¹Institut für Mathematische Stochastik, Georg-August-Universität Göttingen

²Laser-Laboratorium Göttingen

³Mathematisches Institut, Georg-August-Universität Göttingen

⁴Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Acknowledgements

First and foremost, I thank Axel Munk and Stephan Huckemann for all their support over the years. I owe much to their advice and patience with me.

The group of Alexander Egner, especially Oskar Laitenberger and Claudia Geisler, provided important help in understanding the physical background of fluorescence microscopy. Additionally, Oskar Laitenberger constructed the experimental setup and supplied me with the data sets with which I tested my estimation method.

Many thanks go to Timo Aspelmeier, Alexander Egner, Stefan Halverscheid, and Russell Luke for participating in the examination committee.

I gratefully acknowledge support from the Deutsche Forschungsgemeinschaft grant SFB 755.

I am also very obliged to Anne Hein, Benjamin Eltzner, Benjamin Heuer, and Fabian Telschow for lots of helpful suggestions and discussions.

Last but not least, I'd like to thank my lovely wife Julia and my parents for their moral support and for taking the load off me when I needed extra time to write the present work.

Contents

Nomenclature	vii
1 Introduction	1
1.1 What this is all about	1
1.2 On fluorescence microscopy	1
1.3 Estimation of drift, rotation and scaling	4
1.4 A simple model for SMS microscopy	4
1.5 Relation to the literature	6
1.6 Overview	7
2 Fourier transform and Fourier-Mellin transform	13
2.1 Fourier transform	14
2.2 Fourier-Mellin transform	16
2.3 Analytical Fourier-Mellin transform	17
3 Model	19
3.1 Binning	21
4 M-estimation method	25
5 Preliminary Computations	29
6 Main Results	37
6.1 Estimation of rotation and scaling	37
6.2 Drift Estimation	59
7 Model selection and motion blur	83
7.1 Drift blur	83
7.2 Rotation and scaling blur	85
7.3 Motion blur measure in application	85
8 Other estimation methods	87
8.1 Cross Correlation	87
8.2 Fiducial marker tracking	89
9 Simulation Study	93
10 Application to SMS data	107
11 Bootstrap confidence bands	111

12 Conclusion	115
A Mathematical Tools	117
B Well-known theorems	125

Nomenclature

$(\hat{\phi}_T, \hat{\alpha}_T)$	Rotation and scaling parameter estimator, page 26
$ x $	Absolute value of a number $x \in \mathbb{C}$
\mathcal{AFM}_g^{-1}	inverse analytical Fourier-Mellin transform of a function g , page 17
\mathcal{AFM}_g	analytical Fourier-Mellin transform of a function g , page 17
$\alpha \in \mathbb{A}$	Scaling parameter, page 19
$\mathbb{A} \subset \mathbb{R}^{d_3}$	Scaling parameter set, page 19
$\alpha_0 \in \mathbb{A}$	True and unknown scaling parameter, page 19
$\beta_T \in \mathbb{N}$	Bin size for a sequence of T observed frames, page 21
\mathcal{L}^p	\mathcal{L}^p -space, page 13
\mathcal{LP}	Log-polar coordinate transform, page 16
$\mathcal{N}(\mu, \nu^2)$	Gaussian distribution with expectation $\mu \in \mathbb{R}$ and variance $\nu^2 > 0$
\mathcal{P}	Polar coordinate transform, page 16
δ^θ	Drift function, page 19
$\epsilon_j^t \sim \mathcal{N}(0, 1)$	Statistical errors (after binning), page 21
\mathcal{F}_g^{-1}	inverse Fourier transform of a function g , page 14
$\mathcal{F}_g^{\mathbb{X}}$	discrete Fourier transform of a function g on a set \mathbb{X} , page 15
\mathcal{F}_g	Fourier transform of a function g , page 14
\mathcal{FM}_g	Fourier-Mellin transform of a function g , page 16
γ	A constant used for the analytical Fourier Mellin transform, page 17
\hat{f}_T	Image estimator, page 28
$\hat{\theta}_T$	Drift parameter estimator, page 27
i	Imaginary unit, $i = \sqrt{-1} \in \mathbb{C}$
$\Im(x)$	Imaginary part of a number $x \in \mathbb{C}$
\mathbb{I}_B	Indicator function on a set B , page 21
$\mathcal{U}(S)$	Uniform distribution on a set S
$\ \cdot\ $	Euclidean norm, page 14
$\ \cdot\ _1$	1-norm, page 14
$\ \cdot\ _\infty, \ \cdot\ _{\mathcal{L}^\infty}$	\mathcal{L}^∞ -seminorm, page 13
$\ \cdot\ _{\mathcal{L}^p}$	\mathcal{L}^p -seminorm, page 13
$\ \cdot\ _{L^\infty}$	L^∞ -norm, page 13
$\ \cdot\ _{L^p}$	L^p -norm, page 13
ν_j^t	Noise level after binning, page 21
$\phi \in \Phi$	Rotation parameter, page 19
$\Phi \subset \mathbb{R}^{d_2}$	Rotation parameter set, page 19
$\phi_0 \in \Phi$	True and unknown rotation parameter, page 19
$\Re(x)$	Real part of a number $x \in \mathbb{C}$
ρ^ϕ	Rotation function, page 19
σ^α	Scaling function, page 19
\mathbb{T}	Set of observation time points after binning, page 21
$\text{Pois}(\lambda)$	Poisson distribution with parameter λ

$\theta \in \Theta$	Drift parameter, page 19
$\Theta \subset \mathbb{R}^{d_1}$	Drift parameter set, page 19
$\theta_0 \in \Theta$	True and unknown drift parameter, page 19
$\tilde{\epsilon}_j^t \sim \mathcal{N}(0, 1)$	Statistical error at time t and pixel j , page 20
$\tilde{\nu}_j^t$	Noise level at time t and pixel j , page 20
$\tilde{J}^t \subseteq \{1, \dots, n\}$..	Set of observed pixels at time t , page 20
$\widetilde{M}(\phi, \alpha)$	non constant part of $M(\phi, \alpha)$, page 26
$\widetilde{M}_T(\phi, \alpha)$	non constant part of $M_T(\phi, \alpha)$, page 25
$\widetilde{N}(\theta)$	non constant part of $N(\theta)$, page 27
$\widetilde{N}_T(\theta)$	non constant part of $N_T(\theta)$, page 27
\widetilde{O}_j^t	Observations before binning, page 20
TV	Total variation, page 37
$\mathcal{W}^t(\omega)$	Error term of the Fourier-Mellin transformed data, page 23
$\widetilde{\mathbb{T}}$	Set of time points of observations, page 20
$d_{u,v}(\sigma, \rho)$	Rotation and scaling correction term, page 25
f	A grey value image, page 19
$g_1 \star g_2$	Cross correlation of two functions g_1 and g_2 , page 87
H^p	Sobolev space of order p , page 15
$h_\omega(\delta)$	Drift correction term, page 26
J_T^t	Set of observed pixels after binning, page 21
L^p	L^p -space, page 13
$M(\phi, \alpha)$	population contrast functional (for rotation and scaling), page 26
M^0	constant part of $M(\phi, \alpha)$, page 26
m_2	Motion blur measure, page 83
$M_T(\phi, \alpha)$	empirical contrast functional (for rotation and scaling), page 25
M_T^0	constant part of $M_T(\phi, \alpha)$, page 25
$n \in \mathbb{N}$	Total number of observable pixels, page 20
$N(\theta)$	population contrast functional (for drift), page 27
N^0	constant part of $N(\theta)$, page 27
$N_T(\theta)$	empirical contrast functional (for drift), page 27
N_T^0	constant part of $N_T(\theta)$, page 27
O_j^t	Observations after binning, page 21
R_ρ	rotation matrix with angle ρ , page 5
r_T	Cutoff for Fourier transform, page 27
u_T, v_T	Cutoff values for analytical Fourier-Mellin transform, page 25
$V_T^t(\omega)$	For rotation and scaling corrected version of $W^t(\omega)$, page 26
$W^t(\omega)$	Fourier transformed error term, page 23
$Y^t(\omega)$	Fourier transformation of observations at time t , page 23
$Z_T^t(\omega)$	For rotation and scaling corrected version of $Y^t(\omega)$, page 26
a.e.	almost everywhere, page 13

Chapter 1

Introduction

1.1 What this is all about

In this work, we deal with sequences of pixel images (frames) which are noisy shifted, rotated, and scaled versions of some unknown image \tilde{f} . Moreover, those frames are sparse in the sense that they do not show the whole transformed (and noisy) image \tilde{f} but only relatively few pixels (at random locations). If the sequence contains enough frames, it is likely that every pixel is observed in at least one of them, and summing up all frames yields a rather complete version of the unknown image \tilde{f} . However, since the single frames are subject to rigid motions, the result is blurred. This situation comes up in single marker switching (SMS) microscopy on which we elaborate below. Figure 1.1 shows the aggregations of the first 14 and last 14 frames of such a sequence obtained in an SMS experiment. In applications, the frames are often calibrated by tracking the positions of so-called fiducial markers (bright spots that are fixed to the specimen and appear in every frame). This method is technically demanding and has further drawbacks (see below). We propose a purely statistical reconstruction method based on parametric models for the drift, rotation, and scaling functions, where we estimate those parameters by minimizing certain functionals. We prove consistency of our M-estimators, asymptotic normality of the rotation and scaling parameter estimators, and uniform tightness of the drift parameter estimator. Furthermore, we test our M-estimators in a simulation study with various parametric motion models and statistical error models. Last but not least, we apply our method to SMS microscopy data (Figure 1.2 displays the superposition of all frames from the data from Figure 1.1 as well as the reconstructions with our M-estimator and with fiducial marker tracking) and construct bootstrap confidence bands for the drift, rotation, and scaling functions.

1.2 On fluorescence microscopy

In the life sciences, optical fluorescence imaging is an important tool for studying biological molecules at subcellular level. However, for more than a hundred years, the Abbé diffraction barrier implied a physical limitation of spatial resolution for any kind of light microscopy. This barrier means that it is not possible to distinguish between two features that are closer than about 250 nm (approximately half the smallest wavelength of visible light) in lateral and 500 nm in axial direction because they merge into each other. In the last twenty years, the Abbé barrier has been overcome by recording features within such a diffraction limited area consecutively instead of simultaneously by changing their ability to be observed over time (Hell, 2009). Concerning fluorescence microscopy, this means

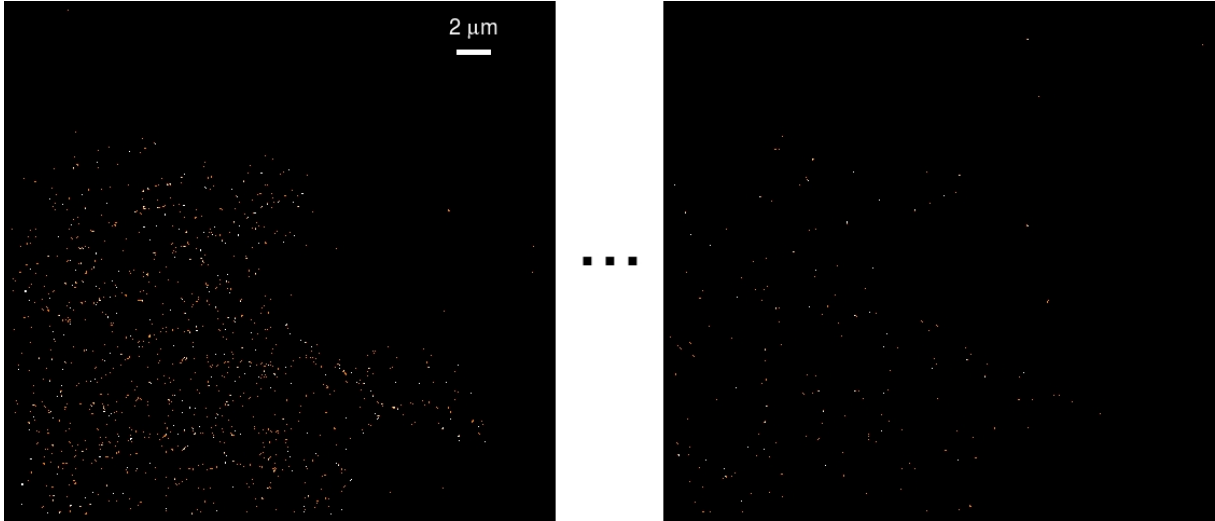


Figure 1.1: *Superimposed position histograms derived from the first (left) and last (right) 14 frames of an SMS experiment. The sample is a HeLa-cell β -tubulin network. The superimposed position histograms of all 29,000 frames of the experiment and a reconstruction derived from them are shown in Figure 1.2.*

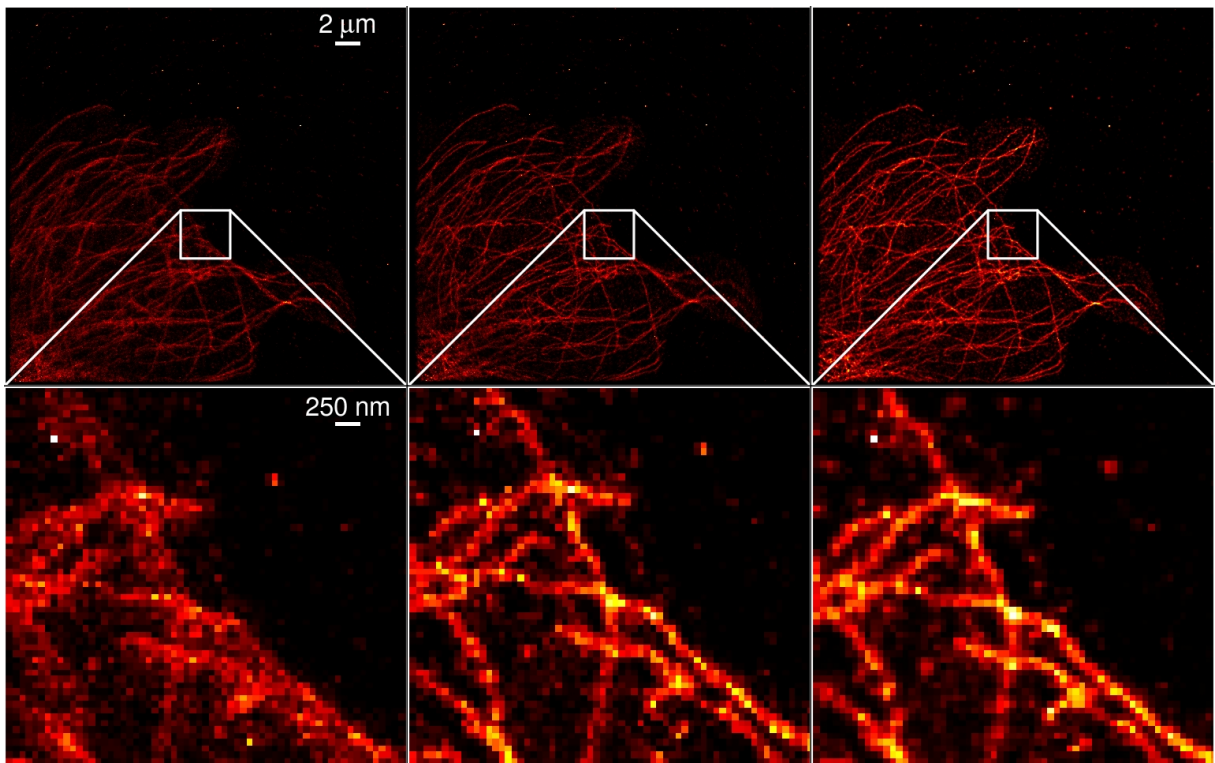


Figure 1.2: *Left column: SMS acquisition of the β -tubulin network I. Top left: motion blurred position histogram. Top middle: reconstructed position histogram under a linear drift model and a quadratic rotation and scaling model. Top right: position histogram reconstructed using fiducial marker tracking. Bottom left: detailed view inside the white box above. Bottom middle and right: reconstructed details after motion corrections.*

changing the fluorophore's ability to emit a fluorescence photon or to change the color or other properties of the emitted photon. Nowadays, there are several such techniques (Hell, 2003; Betzig et al., 2006; Rust et al., 2006; Hess et al., 2006) which has revolutionized the field of cell imaging. It is even possible to observe active biological molecules at a resolution level down to 10 – 20 nm which gives entirely new insights into the signalling and transport processes within cells (see e.g., Westphal et al. (2008); Berning et al. (2012); Jones et al. (2011); Huang et al. (2013), to mention a few).

Roughly speaking, there are two distinct categories of modern nanoscale microscopy. The first is the targeted mode (ensemble based), where the fluorophores (markers) are switched at a known and precisely defined coordinate. It includes techniques such as stimulated emission depletion (STED) (Hell and Wichmann, 1994; Klar et al., 2000; Schmidt et al., 2008), saturated patterned excitation microscopy (SPEM) (Heintzmann et al., 2002) or saturated structured illumination microscopy (SSIM) (Gustafsson, 2005), and reversible saturable optical fluorescence transitions (RESOLFT) (Hofmann et al., 2005; Hell, 2003). Because of the direct targeting, the specimen can usually be scanned in a relatively short time and thus movements are not a major source of blurring.

In contrast, the second category is the stochastic switching (or single marker switching, SMS) mode, where the entire sample is illuminated simultaneously, but with a low switching light intensity. This assures that with high probability only a few (random) markers are in their fluorescent state at any time. Development of these techniques has been rapid during the last years. They include *stochastic optical reconstruction microscopy (STORM)* (Rust et al., 2006; Holden et al., 2011), *photoactivated localization microscopy (PALM)* (Betzig et al., 2006), *fluorescence photoactivation localization microscopy (FPALM)* (Hess et al., 2006), and *PALM with independently running acquisition (PALMIRA)* (Geisler et al., 2007; Egner et al., 2007). See Hell (2007) for a survey.

Due to the fact that each single high resolved image in SMS microscopy yields only very little but sparse information, a long sequence of images (frames) has to be recorded in order to accumulate to a representative view of the specimen. Usually, such sequences are in the range of several tens of thousands of frames. Recently developed methods make explicit use of this sparseness for image reconstruction, for example, employing a sparsity enforcing penalty or prior, see Babcock et al. (2012); Cox et al. (2012); Holden et al. (2011); Zhu et al. (2012); Quan et al. (2011); Hafi et al. (2014). The unknown marker positions are usually determined by calculating the centroid of their observed diffraction patterns, physically enforcing spatial sparseness and rendering more sophisticated deconvolution methods unnecessary. Note that the standard deviation of the average of $N \in \mathbb{N}$ identically and independently distributed random variables is by a factor of $1/\sqrt{N}$ smaller than the standard deviation of only one of these random variables. In our context, N is the average number of detected photons within the individual diffraction patterns, meaning that the localization accuracy can be \sqrt{N} times better than the initial resolution of the microscope (Thompson et al., 2002). The markers localized within each frame are then registered in highly time and space resolved position histograms (see Figure 1.1), the overlay of which represents the final SMS-image.

Nevertheless, since we need to record thousands of frames for this final image, the measurement process typically takes several minutes. Hence, a movement of the specimen over significant distances during this time leads to blurring, which is a major motivation for this work. These movements can have different shapes. Drift may be caused by temperature variations (thermal drift) during the measurement process while rotation might occur due to small vibrations coupled with a rigid specimen that is not perfectly adhesive to the object layer. In this context, even if there are only drift-like effects present,

a large shift at one location and a smaller one at another can lead to a rotation. External systematic movements of the optical device (mechanical drift and rotation) may also be a problem. Moreover, as the optical device heats up during the measurement process, thermal expansion can lead to a decreasing distance between the original focal plane and the ocular, meaning that the image appears scaled up. As can be seen in Figure 1.2 (left upper display) these movements are the major source of blurring. The need to correct for this motion of the object in the sparse position histograms is well known and it is therefore current practice to incorporate fiducial markers (e.g., bright fluorescent microspheres) into the specimen which can be tracked and thus used to align subsequent frames. However, this is not only technically demanding and expensive, often times the fiducials also outshine relevant parts of the image. Hence, developing methods which allow to estimate the drift, rotation, and scaling of the specimen without incorporation of fiducials would be an important achievement.

1.3 Estimation of drift, rotation and scaling

A first attempt at estimating drift has been made by Geisler et al. (2012), who suggested a heuristic correlation method to align subsequent frames properly (see Deschout et al. (2014) for a recent survey on this issue). In Hartmann et al. (2016), we treated this problem in a statistically rigorous way, assuming a parametric model for the drift as a function in time, proposing an M-estimator for this drift parameter and proving its consistency and asymptotic normality. Here, although we will present an expansion of the method of Geisler et al. (2012) to incorporate rotation and scaling in Chapter 8, we focus on a similar expansion of our own M-estimation technique. We argue that a parametric model for the drift, rotation, and scaling functions is often appropriate and we suggest M-estimators for them. See the middle column of Figure 1.2 for the image of the recordings of a β -tubulin network within a HeLa-cell, which was obtained after motion correction with our M-estimators, to be developed in Chapter 4. We will show the asymptotic normality of the rotation and scaling estimator and uniform tightness of the drift estimator as the acquisition time increases, and we argue that this is the “right asymptotics” in SMS microscopy due to relatively long acquisition times which inherently come along with this technique. From this asymptotics we obtain simple bootstrap confidence bands for the drift, rotation, and scaling functions and finally improved estimates of the image together with a measure to access the statistical uncertainty of the aligned images. We stress that our asymptotics is substantially different to that underlying many other image alignment and registration methods where at each time step data from the full image is observed and hence asymptotics concerns the number of pixels tending to infinity.

Finally, our method is compared in real world applications with calibration using cross correlation or fiducial markers. We show that our method is at least as competitive revealing the incorporation of fiducials as not necessary in the analysis and processing of SMS images.

1.4 A simple model for SMS microscopy

In SMS microscopy, data is acquired in a two step process, first, switching marker molecules on, and second, reading out of their fluorescent signal. For details, we refer to Betzig et al. (2006); Geisler et al. (2012); Hell (2009); Hess et al. (2006). However, the data represents single photon counts recorded with an array of photodetectors, where

it is reasonable to assume that different markers emit photons independently. Hence, the data can be described as a spatial (thinned) Poisson process (possibly corrupted by some background noise) with unknown intensity λ which is linked to the unknown marker density \tilde{f} , for example, by a convolution $\lambda = K * \tilde{f}$, where the kernel K is determined by the optical system. In ensemble based microscopy the focal spot is scanned through the sample. This requires an additional deconvolution step which can be helpful to obtain improved resolution (see Vardi et al. (1985); Silverman et al. (1990); Nowak and Kolaczyk (2000); Cavalier and Koo (2002); Antoniadis and Bigot (2006); Zhang et al. (2008); Frick et al. (2013); Bigot et al. (2013) for several Poisson deconvolution methods). In SMS microscopy, however, as considered in this work, the centre of each spot already serves as a very accurate location estimate of the marker molecule because of the enforced sparsity (see Aspelmeier et al. (2015)). Therefore, we adopt current practice, and a sophisticated deconvolution step is not required.

As you may draw from Figure 1.2, indeed the major source of blurring in SMS microscopy comes from sample motion, rather than from optical blurring as the technique is designed to be physically sparse.

Since the number of photon counts in SMS experiments is usually rather high, in the following, we simplify things further by restricting ourselves to use a (heteroscedastic) Gaussian model as an approximation to the Poisson model for large intensity \tilde{f} . Nevertheless, we did some simulations for a Poisson model in our simulation study (see Chapter 9), warranting that approximation appropriate. An approximate model for the above SMS scenario is hence given by (possibly after an offset correction)

$$\tilde{O}_j^t := \begin{cases} \tilde{f}^t(x_j) + \tilde{\nu}_j^t \tilde{\epsilon}_j^t, & \text{if } t \in \tilde{\mathbb{T}}, j \in \tilde{\mathcal{J}}^t, \\ 0, & \text{if } t \in \tilde{\mathbb{T}}, j \in \{1, \dots, n\} \setminus \tilde{\mathcal{J}}^t, \end{cases} \quad (1.1)$$

where $\tilde{f}: \mathbb{R}^2 \rightarrow \mathbb{R}$ represents the marker intensity (which can be thought of as a grey-scale image) and

$$\tilde{f}^t(x_j) := \tilde{f}\left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(x_j - \delta_t^{\theta_0})\right)$$

is a shifted, rotated, and scaled version of \tilde{f} . The scaling factor $\sigma_t^{\alpha_0}$ depends on the observation time t and an unknown parameter $\alpha_0 \in A \subset \mathbb{R}^{d_3}$, where $d_3 \in \mathbb{N}$,

$$R_\rho := \begin{pmatrix} \cos(\rho) & -\sin(\rho) \\ \sin(\rho) & \cos(\rho) \end{pmatrix} \quad (1.2)$$

is the *rotation matrix* with angle ρ , $-\rho_t^{\phi_0}$ is a rotation angle which also depends on time t and an unknown parameter $\phi_0 \in \Phi \subset \mathbb{R}^{d_2}$, where $d_2 \in \mathbb{N}$, and $\delta_t^{\theta_0}$ is a translation (drift) vector depending on time t and an unknown parameter $\theta_0 \in \Theta \subset \mathbb{R}^{d_1}$, where $d_1 \in \mathbb{N}$. The $x_j \in [0, 1]^2$ denote pixel locations. The $\tilde{\epsilon}_j^t \sim \mathcal{N}(0, 1)$ are independent Gaussian errors while the noise levels $\tilde{\nu}_j^t > 0$ model spatial and temporal inhomogeneities and are, in general, unknown. In particular, in a pure Poisson model they would equal the signal \tilde{f}^t . The set $\tilde{\mathbb{T}} = \{0, 1/T, \dots, (T-1)/T\}$ with $T \in \mathbb{N}$ contains equidistant time points (the times at which the images are observed), $n \in \mathbb{N}$ is the total possible number of pixels in the image, and $\tilde{\mathcal{J}}^t \subseteq \{1, \dots, n\}$ indexes the pixels observed at time $t \in \tilde{\mathbb{T}}$.

As stated before, in the low energy stochastic switching scenario we assume that the pixel locations x_j selected by the switch-on process are small in number (i.e., $\tilde{n}_t := \#\tilde{\mathcal{J}}^t$ is small) and sufficiently distant to each other. Consequently, the assumption that the errors $\tilde{\nu}_j^t$ ($t \in \tilde{\mathbb{T}}, j \in \tilde{\mathcal{J}}^t$) are independently distributed for different time points $t_1, t_2 \in \tilde{\mathbb{T}}$, $t_1 \neq t_2$, even if they belong to the same pixel locations, is reasonable. Obviously, the

intensity of the images scales with \tilde{n}_t which means that \tilde{f}^t actually has to be rescaled with the relative amount of total intensity at time t , that is, multiplied by $\tilde{n}_t / \sum_{s \in \tilde{\mathbb{T}}} \tilde{n}_s$, to keep the observed intensity constant. We will suppress this in the following, however, as \tilde{n}_t is observable and any estimate of \tilde{f}^t can thus be rescaled easily.

Note that we need $\tilde{n}_t \geq 1$ for all $t \in \tilde{\mathbb{T}}$, which is, however, always true for SMS microscopy as the sampling rate is never chosen below. We will see that the only additional assumptions required on \tilde{J}_t or \tilde{n}_t are that arbitrarily large unions of subsequent \tilde{J}_t 's converge to $\{1, \dots, n\}$ and that the average observation frequency at each time point t and each pixel location x_j converges to some mean $c^t(x_j)$ (see Assumption 3.8). Both assumptions make sense in our SMS scenario since it is highly unlikely that there is a pixel we never observe in an arbitrarily large number of frames and since the $c^t(x_j)$ are basically determined by the marker intensity $\tilde{f}^t(x_j)$ and the switch-on probability given by the experimental setup. Therefore, even though in SMS microscopy, the \tilde{J}^t 's are strictly speaking random and their exact distributions will depend on the fluorophore characteristics, we will consider the \tilde{J}^t 's to be fixed and we will state our results and conditions accordingly. However, Assumption 3.8 and our main Theorems 6.23 and 6.42 hold similarly for this random situation, for example, when the convergence in Assumption 3.8 is now almost surely, which can be derived from the strong law of large numbers for non identically distributed random variables.

In contrast to the usual asymptotics in imaging, where the pixel number n tends to infinity as the discretization level increases, we consider here the novel scenario where the total pixel number is fixed while the number of time frames T tends to infinity. In SMS nanoscopy, T typically ranges from 10,000 to 40,000, corresponding to a time resolution of several milliseconds.

1.5 Relation to the literature

The asymptotics considered in Chapter 6 require rather involved computations and notably, they are different from various approaches and asymptotics in the literature. Note that, in our SMS microscopy setting, the number \tilde{n}_t of observed pixels at time t in model (1.1) is typically small and, in particular, does not tend to infinity. Hence, our approach is different from *time dynamic imaging* (Feroosh et al., 2002; Huang et al., 2005; Papenberg et al., 2006; Cuzol et al., 2007; Allasonnière et al., 2007; Fleet and Weiss, 2006; Bruhn et al., 2005; Weickert and Schnörr, 2001; Li et al., 2014) where in each time step a (rather) complete sample of the entire image has to be recorded. In strict contrast to this, SMS microscopy provides only a few observed markers at each time. Therefore, our setting is also different from Gamboa et al. (2007) as well as from Bigot et al. (2009), although we borrow the idea of a Procrustes type estimator based on minimizing a suitable contrast functional, see Gower (1975). The two afore mentioned recent references consider a finite and fixed number T of images which are subject to Gaussian noise, each of which comes with an individual unknown similarity transform. More specifically, translated to our setup, Gamboa et al. (2007) consider one-dimensional “images”, each subject to a one-dimensional translation, while Bigot et al. (2009) consider two-dimensional images, each subject to translation, rotation, and scaling. They prove that those unknown similarity transformations can be consistently estimated with asymptotic normality if the number of pixels n tends to infinity which corresponds to an increasing signal-to-noise ratio for each frame. Here, motivated by SMS microscopy, we swap the asymptotics for a finite number of pixels n (of which we observe $\tilde{n}_t \leq n$ at time t), while the number of observation time points T goes to infinity, and allow for a time dependent drift, rotation, and scaling.

We would like to mention that the shift property of the Fourier transform (and a similar property of the analytical Fourier-Mellin transform) which has motivated our approach is crucial in many related methods based on FFT (Reddy and Chatterji, 1996; Bigot et al., 2009).

We note that our work is not limited to SMS microscopy and might be used for other purposes, such as noisy object or particle tracking, when only small parts of the object are registered at each time step as is the case for heavily undersampled magnetic resonance imaging (Li et al., 2014). Extensions to non parametric models for drift, rotation, and scaling are possible and will be the topic of subsequent research. Finally, in a sense, our work is complementary to the issue of testing in fluorescence microscopy whether a protein structure has significantly changed in time as in Bissantz et al. (2009).

1.6 Overview

This thesis is organized as follows.

Chapter 2. In this chapter, we define the Fourier transform

$$\mathcal{F}_g(\omega) = \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, x \rangle} g(x) dx$$

and the (analytical) Fourier-Mellin transform

$$\mathcal{AFM}_g(u, v) = \int_0^\infty \int_0^{2\pi} e^{-2\pi i u \psi} r^{\gamma - i v} (g \circ \mathcal{P})(r, \psi) d\psi \frac{dr}{r}, \quad (u, v) \in \mathbb{Z} \times \mathbb{R},$$

of a function $g: \mathbb{R}^2 \rightarrow \mathbb{R}$, where \mathcal{P} is the polar coordinate transform and $\gamma > 0$ helps with the existence of the integral in the case that $(g \circ \mathcal{P})(r, \psi)$ does not converge to 0 for $r \rightarrow 0$. Due to its shift property

$$\mathcal{F}_{g(\cdot - \delta)}(\omega) = e^{-2\pi i \langle \omega, \delta \rangle} \mathcal{F}_g(\omega), \quad \delta \in \mathbb{R}^2, \quad (1.3)$$

the Fourier transform turns the drift of the image sequence into a phase shift that can be eliminated by taking the absolute values of the Fourier coefficients which allows us to first estimate rotation and scaling. This is done using the Fourier-Mellin transform which has a property similar to (1.3) with respect to rotation and scaling. Key to the estimation method is the Plancherel equality,

$$\int_{\mathbb{R}^2} |g(x)|^2 dx = \int_{\mathbb{R}^2} |\mathcal{F}_g(\omega)|^2 d\omega,$$

since it means that we can minimize distances between images in the Fourier domain instead of the image domain.

Chapter 3. In this chapter, we present our semi-parametric model (3.2). This model is semi-parametric in the sense that it includes an infinite-dimensional parameter \tilde{f} and a finite-dimensional parameter $(\theta, \phi, \alpha) \in \mathbb{R}^{d_1 + d_2 + d_3}$.

To asymptotically decrease the noise level, we average over $\beta_T \in \mathbb{N}$ subsequent frames. This so-called binning gives us the model

$$O_j^t := \frac{1}{\beta_T} \sum_{i=0}^{\beta_T - 1} \tilde{O}_j^{t+i/T} \approx f^t(x_j) + \sqrt{\frac{n_t}{\beta_T}} \nu_j^t \epsilon_j^t, \quad t \in \mathbb{T}, j \in \{1, \dots, n\}, \quad (1.4)$$

where $f^t(x_j) := c^t(x_j)\tilde{f}^t(x_j)$ with an average observation frequency $c^t(x_j)$, $\nu_j^t > 0$ are basically averages of the original noise levels $\tilde{\nu}_j^t$, ϵ_j^t are independent standard normal random variables, and $\mathbb{T} := \{0, \beta_T/T, 2\beta_T/T, \dots, (T - \beta_T)/T\}$. From this, we derive our Fourier model

$$Y^t(\omega) := \mathcal{F}_{f^t}(\omega) + W^t(\omega), \quad t \in \mathbb{T}, \omega \in \mathbb{R}^2,$$

where the $W^t(\omega)$'s are noise terms with centred normal real and imaginary parts. Because of the shift property (1.3) and some additional calculations, with $f := f^0$, we have that

$$|\mathcal{F}_{f^t}|^2(\omega) = (\sigma_t^\alpha)^4 \left| \mathcal{F}_f(\sigma_t^\alpha R_{-\rho_t^\phi} \omega) \right|^2,$$

which does not depend on the drift $\delta_t^{\theta_0}$. In particular, the rotation and scaling in the image domain result in the same rotation and the inverse scaling in the Fourier domain. We will first estimate (ϕ_0, α_0) from the analytical Fourier-Mellin transform of the data,

$$\mathcal{AFM}_{|Y^t|^2}(u, v) = \mathcal{AFM}_{|\mathcal{F}_{f^t}|^2}(u, v) + \mathcal{AFM}_{\mathcal{W}^t}(u, v), \quad t \in \mathbb{T}, (u, v) \in \mathbb{Z} \times \mathbb{R},$$

where $\mathcal{W}^t(\omega) := |W^t(\omega)|^2 + 2\Re(\mathcal{F}_{f^t}(\omega)\overline{W^t(\omega)})$.

Chapter 4. In this chapter, we explain our M-estimation method for the calibration of the image sequence and provide for estimators for the drift, rotation, and scaling parameters as well as for the unknown true image f . This method relies on minimizing so-called contrast functionals. With suitably chosen cutoffs $u_T, v_T > 0$, the contrast functional for rotation and scaling is

$$\begin{aligned} M_T(\phi, \alpha) := & \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| (\sigma_t^\alpha)^{-iv} e^{2\pi i u \rho_t^\phi} \mathcal{AFM}_{|Y^t|^2}(u, v) \right. \\ & \left. - \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} (\sigma_{t'}^\alpha)^{-iv} e^{2\pi i u \rho_{t'}^\phi} \mathcal{AFM}_{|Y^{t'}|^2}(u, v) \right|^2 dv, \end{aligned}$$

from which we derive an estimator $(\hat{\phi}_T, \hat{\alpha}_T) \in \operatorname{argmin}_{(\phi, \alpha) \in \Phi \times \Lambda} \widetilde{M}_T(\phi, \alpha)$ for the parameter (ϕ_0, α_0) . Calibration of the Fourier image sequence $(Y^t)_{t \in \mathbb{T}}$ with $\hat{\phi}_T^{\hat{\phi}_T}$ and $\hat{\sigma}_T^{\hat{\alpha}_T}$ yields the following model,

$$Z_T^t(\omega) := (\hat{\sigma}_T^{\hat{\alpha}_T})^{-2} Y^t \left(1/\hat{\sigma}_T^{\hat{\alpha}_T} \cdot R_{\hat{\phi}_T^{\hat{\phi}_T}} \omega \right).$$

It remains to estimate the drift parameter θ_0 with the minimizer $\hat{\theta}_T$ of the contrast functional

$$N_T(\theta) := \int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| e^{2\pi i \langle \omega, \delta_t^\theta \rangle} Z_T^t(\omega) - \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} e^{2\pi i \langle \omega, \delta_{t'}^\theta \rangle} Z_T^{t'}(\omega) \right|^2 d\omega,$$

with a suitable cutoff $r_T > 0$. By correcting the Z_T^t 's with the resulting drift function estimator $t \mapsto \delta_t^{\hat{\theta}_T}$ and performing a subsequent inverse Fourier transform, we finally get an estimator for the image f ,

$$\hat{f}_T(x_j) := \int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \exp \left(2\pi i \langle \omega, x_j + \delta_t^{\hat{\theta}_T} \rangle \right) Z_T^t(\omega) d\omega, \quad j \in \{1, \dots, n\}.$$

Chapter 5. In this chapter, we deal with some preliminary computations for our main theorems, like the distributions and dependencies of statistical error terms (e.g., $W^t(\omega)$) and derivatives of the rotation and scaling correction term $d_{u,v}(\sigma_t^\alpha, \rho_t^\phi) := (\sigma_t^\alpha)^{-iv} e^{2\pi i u \rho_t^\phi}$ and the drift correction term $h_\omega(\delta_t^\theta) := e^{2\pi i \langle \omega, \delta_t^\theta \rangle}$.

Chapter 6. In this chapter, we present our main theoretical results. In particular, we provide for the main assumptions on the model and derive consistency of the drift, rotation, and scaling parameter estimators $\hat{\theta}_T$, $\hat{\phi}_T$, and $\hat{\alpha}_T$, and the image estimator \hat{f}_T under those assumptions. Moreover, we prove that $\sqrt{T}(\hat{\phi}_T - \phi_0, \hat{\alpha}_T - \alpha_0)$ is asymptotically normal and that $(\sqrt{T}(\hat{\theta}_T - \theta_0))_{T \in \mathbb{N}}$ is uniformly tight, where θ_0 , ϕ_0 , and α_0 are the unknown true parameters.

Chapter 7. In this chapter, we propose a simple model selection method for the parametric drift, rotation, and scaling models based on the values of the contrast functionals. Furthermore, we use the motion blur measure

$$m_2(I) := \log \left(\frac{J(\psi_{\max})}{J(\psi_{\min})} \right)$$

introduced in (Xu et al., 2013) to compare our results with fiducial marker tracking. Here, I is a pixel image,

$$J(\psi) := \sum_{j=1}^{N^2} \left(\Delta I((x_j)_1, (x_j)_2)_\psi \right)^2$$

is the average squared directional derivative of I in direction $(\cos(\psi), \sin(\psi))$, ψ_{\min} is a minimizer of J , and $\psi_{\max} = \psi_{\min} \pm \pi/2$. The idea behind m_2 is that, if the image I is smeared along $(\cos(\psi), \sin(\psi))$, then its intensity will vary little in that direction and much in the perpendicular direction. This motion blur measure can be used to detect the major drift direction and the blurring caused by the drift. Since rotation and scaling become a translation if we transform the image into log-polar coordinates, m_2 can be easily adapted to measure the motion blur created by these movements.

Chapter 8. In this chapter, we describe two alternative methods for the estimation of drift, rotation and scaling in image sequences (without theoretical results).

The first one is based on the work of Geisler et al. (2012) who deal with drift only and maximize the cross correlation of two frames,

$$(O^t \star O^{t'}) (\tau) := \sum_{j=1}^n S_\tau(O_j^t) O_j^{t'},$$

to find the optimal lag $\tau^{t,t'}$ which should be close to the true translation vector $\delta_{t'}^{\theta_0} - \delta_t^{\theta_0}$ between the t -th and the t' -th frame. Here, S_τ denotes the translation by τ . Then, $\bar{\tau}^{t,\bullet} := \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} \tau^{t,t'}$ estimates the average translation of the t -th single frame O^t with respect to the entire sequence. Hence, we can correct the image sequence for drift by shifting each O^t by $-\bar{\tau}^{t,\bullet}$. Since the squared Fourier magnitudes $|Y^t|^2$ are devoid of drift but retain the rotation and (inverse) scaling of the original frames, and transformation into log-polar coordinates converts those into translations, this cross correlation method can easily be adapted to estimate rotation and scaling. As with the M-estimation technique,

we can calibrate the images for rotation and scaling and, in a second step, estimate the drift by employing the cross correlation procedure again.

The second method is the afore-mentioned fiducial marker tracking which is commonly used in experiments. It is based on tracking the positions of at least two bright fluorescent spheres that have to be attached to the specimen and then appear in every frame. We denote the positions of these two markers in the t -th frame with μ_1^t and μ_2^t . Since the translation by $\delta_t^{\theta_0}$, the rotation by $\rho_t^{\phi_0}$, and the scaling by $\sigma_t^{\alpha_0}$ are assumed to act globally on the entire image, we have that $\mu_i^t = \sigma_t^{\alpha_0} R_{\rho_t^{\phi_0}}(\mu_i^0 + \delta_t^{\theta_0})$, $i \in \{1, 2\}$. Thus,

$$\mu_1^t - \mu_2^t = \sigma_t^{\alpha_0} R_{\rho_t^{\phi_0}}(\mu_1^0 - \mu_2^0),$$

which means that we can see the rotation and scaling directly from comparing the direction and length of the difference vector $\mu_1^t - \mu_2^t$ with those of $\mu_1^0 - \mu_2^0$. After calibrating the frames with these quantities, we can correct them for drift by aligning the marker positions in each frame with their respective starting points in the first frame.

Chapter 9. In this chapter, we illustrate the M-estimation method proposed in Chapter 4 in a simulation study, using polynomial drift, rotation, and scaling models. In particular, we study the robustness of our M-estimator with respect to outliers by exchanging the Gaussian distribution of the errors ϵ_j^t with a Student- t -distribution with two degrees of freedom. In both cases, the sparsity of the images is enforced by multiplying the (noisy) grey values at all possible pixel locations with independent Bernoulli random variables with parameter $p \in (0, 1)$. Moreover, we employ a Poisson model, that is, $O_j^t \sim \text{Pois}(p f^t(x_j))$ independently in j and t , for which our method still works fine. Here, the Bernoulli probability p yields a comparable average image intensity with the two other error models. Furthermore, we consider a piecewise linear drift, rotation, and scaling model with a jump at an unknown time point.

Chapter 10. In this chapter, we apply our M-estimation method to SMS nanoscopy data and give a detailed discussion of the results including bootstrap confidence regions. In particular, we demonstrate the model selection method from Chapter 7 and compare our M-estimators to results from cross correlation and fiducial marker tracking (see Chapter 8).

Chapter 11. In this final chapter, we explain how to construct bootstrap confidence bands for the drift, rotation, and scaling functions. Bootstrapping is a so-called re-sampling method which means that we create additional ‘‘artificial data’’ by sampling from the original data in some way. We employ a residual bootstrap oriented on the model (1.4), simulating $B \in \mathbb{N}$ bootstrap samples of the form

$$(O_j^t)^{(b)} := \hat{f}_T \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}}(x_j - \delta_t^{\hat{\theta}_T}) \right) + (\epsilon_j^t)^{(b)}, \quad t \in \mathbb{T}, j \in \{1, \dots, n\},$$

where $b \in \{1, \dots, B\}$ and the $(\epsilon_j^t)^{(b)}$'s are independently drawn from the uniform distribution on the set of residuals,

$$(\epsilon_j^t)^{(b)} \sim \mathcal{U} \left\{ O_j^t - \hat{f}_T \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}}(x_j - \delta_t^{\hat{\theta}_T}) \right) \mid t \in \mathbb{T}, j \in \{1, \dots, n\} \right\}.$$

Performing our M-estimation method with $(O_j^t)^{(b)}$ instead of O_j^t gives us bootstrap replicats $\hat{\theta}_T^{(b)}$, $\hat{\phi}_T^{(b)}$, and $\hat{\alpha}_T^{(b)}$ of the estimators $\hat{\theta}_T$, $\hat{\phi}_T$, and $\hat{\alpha}_T$, B of each, which correspond to

drift component functions $(\delta_t^{\hat{\theta}^{(b)}})_1$, $(\delta_t^{\hat{\theta}^{(b)}})_2$, rotation functions $\rho_t^{\hat{\phi}^{(b)}}$, and scaling functions $\sigma_t^{\hat{\alpha}^{(b)}}$. Based on the work of Hall and Pittelkow (1990), we construct a 95%-confidence band for the drift component function $(\delta^{\theta_0})_1$ by choosing optimal $u_+, u_- > 0$ such that $u_+ + u_-$ is minimal under the condition that

$$(\delta_t^{\hat{\theta}^{(b)}})_1 \in \left[(\delta_t^{\hat{\theta}^{(b)}})_1 - u_- t, (\delta_t^{\hat{\theta}^{(b)}})_1 + u_+ t \right] \quad \text{for all } t \in [0, 1],$$

for all b in a set that includes at least 95% of the elements of $\{1, \dots, B\}$. The boundaries of the confidence band are then given by the functions $t \mapsto (\delta_t^{\hat{\theta}^{(b)}})_1 - u_- t$ and $t \mapsto (\delta_t^{\hat{\theta}^{(b)}})_1 + u_+ t$. Bootstrap confidence bands for $(\delta_t^{\hat{\theta}^{(b)}})_2$, $\rho_t^{\hat{\phi}^{(b)}}$, and $\sigma_t^{\hat{\alpha}^{(b)}}$ are achieved similarly.

Chapter 2

Fourier transform and Fourier-Mellin transform

In this chapter, we define the Fourier transform and the related (analytical) Fourier-Mellin transform which are crucial for this work because of their properties proved in the Lemmas 2.5 and 2.18. For more details on the definitions, lemmas, and theorems presented in this chapter, see (Rudin, 1990). First, we need the following definition.

Definition 2.1 (\mathcal{L}^p -spaces with seminorm). *Let $(\Omega, \mathcal{A}, \mu)$ a measure space. For values $p \in [1, \infty)$, we define the \mathcal{L}^p -space*

$$\mathcal{L}^p(\Omega, \mathcal{A}, \mu) := \{g: \Omega \rightarrow \mathbb{C} \mid g \text{ is } \mu\text{-measurable and } \|g\|_{\mathcal{L}^p} < \infty\},$$

with the \mathcal{L}^p -seminorm

$$\|\cdot\|_{\mathcal{L}^p}: \mathcal{L}^p \rightarrow [0, \infty], \quad g \mapsto \left(\int_{\Omega} |g(x)|^p \mu(dx) \right)^{1/p}.$$

For the special case $p = \infty$, we define

$$\mathcal{L}^{\infty}(\Omega, \mathcal{A}, \mu) := \{g: \Omega \rightarrow \mathbb{C} \mid g \text{ is } \mu\text{-measurable and } \|g\|_{\mathcal{L}^{\infty}} < \infty\},$$

with the \mathcal{L}^{∞} -seminorm

$$\|\cdot\|_{\mathcal{L}^{\infty}} := \inf_{N \in \mathcal{A}, \mu(N)=0} \sup_{x \in \Omega \setminus N} |g(x)|.$$

We will denote $\|\cdot\|_{\mathcal{L}^{\infty}}$ also with $\|\cdot\|_{\infty}$.

Note, that $\|\cdot\|_{\mathcal{L}^p}$ is only a seminorm. For example, if we consider $\mathcal{L}^p(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, where $\mathcal{B}(\mathbb{R})$ is the Borel- σ -algebra on \mathbb{R} and λ is the Lebesgue-measure on $\mathcal{B}(\mathbb{R})$, and $p \in [0, \infty]$, then we have for

$$g: \mathbb{R} \rightarrow \mathbb{C}, \quad x \mapsto \begin{cases} 1, & \text{if } x \in \mathbb{Q}, \\ 0, & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$$

that $g \in \mathcal{L}^p(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ with $\|g\|_{\mathcal{L}^p} = 0$, but $g \neq 0$. This is because $\lambda(\mathbb{Q}) = 0$. Identifying functions in $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ if they are μ -almost everywhere (μ -a.e.) equal (i.e., they differ only on a set $N \in \mathcal{A}$ with $\mu(N) = 0$) leads to the following normed spaces.

Definition 2.2 (L^p -spaces with norm). *Let $(\Omega, \mathcal{A}, \mu)$ a measure space, $p \in [1, \infty]$, and*

$$\mathcal{N}^p := \{g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu) \mid \|g\|_{\mathcal{L}^p} = 0\} = \{g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu) \mid g = 0 \text{ } \mu\text{-a.e.}\}.$$

Then, the L^p -space $L^p(\Omega, \mathcal{A}, \mu) := \mathcal{L}^p(\Omega, \mathcal{A}, \mu) / \mathcal{N}^p$, together with the L^p -norm

$$\|\cdot\|_{L^p}: L^p(\Omega, \mathcal{A}, \mu) \rightarrow [0, \infty), \quad [g] \mapsto \|g\|_{\mathcal{L}^p},$$

is a normed space.

2.1 Fourier transform

Definition 2.3 (Fourier transform). Let $g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$, where $\mathcal{B}(\mathbb{R}^2)$ is the Borel- σ -algebra on \mathbb{R}^2 . The Fourier transform of g is defined as

$$\mathcal{F}_g: \mathbb{R}^2 \rightarrow \mathbb{C}, \quad \omega \mapsto \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, x \rangle} g(x) dx.$$

Moreover, we define the inverse Fourier transform as

$$\mathcal{F}_g^{-1}: \mathbb{R}^2 \rightarrow \mathbb{C}, \quad x \mapsto \int_{\mathbb{R}^2} e^{2\pi i \langle \omega, x \rangle} g(\omega) d\omega.$$

The following lemma can be found in (Rudin, 1990, page 22)

Lemma 2.4 (Inverse Fourier transform). Let $g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$. If we have that $\mathcal{F}_g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$, then $\mathcal{F}_{\mathcal{F}_g}^{-1} = g$ a.e.

Lemma 2.5 (Generalized shift property). Let $g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$, $\delta \in \mathbb{R}^2$, $\rho \in [0, 2\pi)$, and $\sigma > 0$. With

$$\tilde{g}: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad x \mapsto g(1/\sigma \cdot R_{-\rho}(x - \delta)),$$

where $R_{-\rho}$ is the rotation by $-\rho$, we have

$$\mathcal{F}_{\tilde{g}}(\omega) = \sigma^2 e^{-2\pi i \langle \omega, \delta \rangle} \mathcal{F}_g(\sigma R_{-\rho} \omega).$$

Proof. With the substitution $y := 1/\sigma \cdot R_{-\rho}(x - \delta)$, we get

$$\begin{aligned} \mathcal{F}_{\tilde{g}}(\omega) &= \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, x \rangle} \tilde{g}(x) dx \\ &= \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, x \rangle} g(1/\sigma \cdot R_{-\rho}(x - \delta)) dx \\ &= \int_{\mathbb{R}^2} \exp(-2\pi i \langle \omega, \sigma R_{\rho} y + \delta \rangle) g(y) \sigma^2 dy \\ &= \sigma^2 e^{-2\pi i \langle \omega, \delta \rangle} \int_{\mathbb{R}^2} \exp(-2\pi i \langle \omega, \sigma R_{\rho} y \rangle) g(y) dy \\ &= \sigma^2 e^{-2\pi i \langle \omega, \delta \rangle} \int_{\mathbb{R}^2} \exp(-2\pi i \langle \sigma R_{-\rho} \omega, y \rangle) g(y) dy \\ &= \sigma^2 e^{-2\pi i \langle \omega, \delta \rangle} \mathcal{F}_g(\sigma R_{-\rho} \omega). \end{aligned}$$

□

This means that a translation in the image domain becomes a phase shift in the Fourier domain (this is the so-called *shift property* of the Fourier transform), a rotation leads to the same rotation in the Fourier domain, and scaling in the image domain translates to the inverse scaling in the Fourier domain as well as a multiplication of the Fourier magnitude.

In what follows, we will need some standard norms that we define below.

Definition 2.6 (Euclidean norm and 1-norm). Let $d \in \mathbb{N}$, $b = (b_1, \dots, b_d)^\top \in \mathbb{R}^d$, and $A = (a_{i,j})_{i,j=1}^d \in \mathbb{R}^{d \times d}$. We define the Euclidean norm

$$\|b\| := \sqrt{\sum_{i=1}^d b_i^2}$$

and the 1-norm

$$\|b\|_1 := \sum_{i=1}^d |b_i|, \quad \|A\|_1 := \sum_{i,j=1}^d |a_{i,j}|.$$

The following well-known properties of the 1-norm will be useful throughout this work.

Lemma 2.7 (Properties of $\|\cdot\|_1$). *Let $d \in \mathbb{N}$, $x, y \in \mathbb{R}^d$, and $A \in \mathbb{R}^{d \times d}$. Then,*

$$\|x\| \leq \|x\|_1 \leq \sqrt{d} \|x\|, \quad \|xy^\top\|_1 = \|x\|_1 \|y\|_1 \quad \text{and} \quad \|Ax\|_1 \leq \|A\|_1 \|x\|_1.$$

The following definition will come in handy later on. It is linked to an integrable function $g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$ being p times differentiable in a weak sense (see e.g., Evans (1998), page 245).

Definition 2.8 (Sobolev space). *For $p > 0$, we call*

$$H^p(\mathbb{R}^2) := \left\{ g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda) \mid \int_{\mathbb{R}^2} (1 + \|\omega\|^2)^p |\mathcal{F}_g(\omega)|^2 d\omega < \infty \right\}$$

the Sobolev space of order p .

By Lemma 2.5, a translation by a vector δ in the image domain becomes a multiplicative phase shift $e^{-2\pi i \langle \omega, \delta \rangle}$ in the Fourier domain. Because $|e^x| = 1$ for all $x \in \mathbb{C}$, the Fourier magnitude defined below helps to get rid of the drift so we can estimate rotation and scaling first.

Definition 2.9 (Fourier magnitude). *For any function g , we define its Fourier magnitudes as*

$$|\mathcal{F}_g| : \mathbb{R}^2 \rightarrow [0, \infty), \quad \omega \mapsto |\mathcal{F}_g(\omega)|.$$

Lemma 2.10 (π -periodicity of Fourier magnitude). *For integrable $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, the Fourier magnitude $|\mathcal{F}_g|$ of g is invariant to rotation by π .*

Proof. Let $\tilde{g} : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto g(R_\pi x)$, where R_π is the rotation by π . Because $R_\pi x = -x$ for all $x \in \mathbb{R}^2$, we have

$$\begin{aligned} |\mathcal{F}_{\tilde{g}}(\omega)|^2 &= \left| \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, x \rangle} g(-x) dx \right|^2 = \left| \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, -y \rangle} g(y) dy \right|^2 \\ &= \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, (-y) - (-y') \rangle} g(y) g(y') dy dy' \\ &= \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, y' - y \rangle} g(y') g(y) dy' dy \\ &= \left| \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, y \rangle} g(y) dy \right|^2 = |\mathcal{F}_g(\omega)|^2, \end{aligned}$$

with $y := R_\pi x$. □

Since we deal with data in the form of pixel images, we will also need the discrete Fourier transform.

Definition 2.11 (Discrete Fourier transform). *For a finite set J , $\mathbb{X} = \{x_j \mid j \in J\} \subset \mathbb{R}^2$, and a function $g : \mathbb{X} \rightarrow \mathbb{R}$, we call*

$$\mathcal{F}_g^{\mathbb{X}}(\omega) := \frac{1}{\#J} \sum_{j \in J} e^{-2\pi i \langle \omega, x_j \rangle} g(x_j), \quad \omega \in \mathbb{R}^2,$$

the discrete Fourier transform of g .

The following theorem is essential for our M-estimation method which we will describe in detail in Chapter 4. In our context, it says that the distance between two images f^t and $f^{t'}$ is the same as the distance between their Fourier transforms \mathcal{F}_{f^t} and $\mathcal{F}_{f^{t'}}$, so that we can minimize such distances in Fourier space instead. For a proof, see (Rudin, 1990, page 26).

Theorem 2.12 (Plancherel). *Let $d \in \mathbb{N}$. There is an isometry*

$$\Psi: L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda) \rightarrow L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda)$$

which is unitary (i.e., for all $g_1, g_2 \in L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda)$ we have that $\langle \Psi g_1, \Psi g_2 \rangle = \langle g_1, g_2 \rangle$) and uniquely defined by $\Psi(g) = \mathcal{F}_g$ for all $g \in \mathcal{S}(\mathbb{R}^d)$, where the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is a certain subset of the set of smooth functions that is dense in $L^p(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda)$ for all $p \in [0, \infty)$.

2.2 Fourier-Mellin transform

As we have seen in Lemma 2.5, the Fourier magnitude is invariant to translations of the image which will prove useful. Various Fourier-type transformations have been developed for handling different groups of transformations in image analysis (see e.g., Derrode and Ghorbel (2004); Ghorbel (1994); Lenz (1990); Gauthier et al. (1991); Segman et al. (1992)). As explained in the following, the Fourier-Mellin transformation has a “shift property” with respect to rotation and scaling. We refer to (Rudin, 1990) for more details.

Let \mathbb{Z} the additive group of integers, \mathbb{R} the additive group of real numbers, \mathbb{R}_+^* the multiplicative group of strictly positive real numbers, and S^1 the unit sphere in \mathbb{R}^2 , all of which are locally compact groups. We will identify points in S^1 (as vectors in \mathbb{R}^2) with their angle to the x -axis in $[0, 2\pi)$. Then, the direct product $\mathbb{R}_+^* \times S^1$ is also locally compact and represents scaling factors and rotation angles in the real plane. Let dr and $d\psi$ denote the standard Lebesgue measures on \mathbb{R}_+^* and $[0, 2\pi)$, respectively. Then, the Haar measure of $\mathbb{R}_+^* \times S^1$ is $dr/r \cdot d\psi/(2\pi)$, which is positive and invariant (because it stems from the Lebesgue measures). Furthermore, the dual group of $\mathbb{R}_+^* \times S^1$ is $\mathbb{Z} \times \mathbb{R}$. Therefore, we can define a Fourier transform on $\mathbb{R}_+^* \times S^1$ as described in Definition 2.14. For that, we first need to define the polar (and log-polar) coordinate transformations.

Definition 2.13 (Polar and log-polar coordinate transforms). *We define the polar coordinate transform \mathcal{P} and the log-polar coordinate transform \mathcal{LP} as*

$$\begin{aligned} \mathcal{P}: [0, \infty) \times [0, 2\pi) &\rightarrow \mathbb{R}^2, & (r, \psi) &\mapsto (r \cos(\psi), r \sin(\psi)), \\ \mathcal{LP}: \mathbb{R} \times [0, 2\pi) &\rightarrow \mathbb{R}^2, & (l, \psi) &\mapsto (e^l \cos(\psi), e^l \sin(\psi)). \end{aligned}$$

Definition 2.14 (Fourier-Mellin transform). *Let $g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$. The Fourier-Mellin transform of g is defined as*

$$\mathcal{FM}_g: \mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{C}, \quad (u, v) \mapsto \int_0^\infty \int_0^{2\pi} (g \circ \mathcal{P})(r, \psi) r^{-iv} e^{-iu\psi} d\psi \frac{dr}{r}.$$

Remark 2.15 (Connection between Fourier transform and Fourier-Mellin transform). *The Fourier-Mellin transform is a Fourier-type transform on the similarity group $\mathbb{Z} \times \mathbb{R}$. This will later allow us to use the Fast Fourier Transform algorithm (FFT, see e.g., Cooley and Tukey (1965)) to efficiently compute the Fourier-Mellin transform.*

2.3 Analytical Fourier-Mellin transform

The Fourier-Mellin transform is not well-defined if the function g does not vanish for $r \rightarrow 0$ (see e.g., Derrode and Ghorbel (2004)). In particular, pixel images usually do not assume the value 0 at their centre. To fix this problem, we multiply $g(r, \psi)$ with r^γ with a suitable $\gamma > 0$. This leads to the following definition.

Definition 2.16 (Analytical Fourier-Mellin transform). *Let $g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$ and let $\gamma > 0$. The analytical Fourier-Mellin transformation of g is defined as*

$$\mathcal{AFM}_g: \mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{C}, \quad (u, v) \mapsto \int_0^\infty \int_0^{2\pi} e^{-2\pi i u \psi} r^{\gamma - i v} (g \circ \mathcal{P})(r, \psi) \, d\psi \frac{dr}{r}.$$

Furthermore, we define the inverse analytical Fourier-Mellin transform as

$$\mathcal{AFM}_g^{-1}: [0, \infty) \times [0, 2\pi), \quad (r, \psi) \mapsto \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} e^{2\pi i u \psi} r^{i v} g(u, v) \, dv.$$

Lemma 2.17 (Inverse analytical Fourier-Mellin transform). *Let $g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$ and $\gamma > 0$. If $\mathcal{AFM}_g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$, then $\mathcal{AFM}_{\mathcal{AFM}_g}^{-1}(r, \psi) = r^\gamma g(r, \psi)$ for all $(r, \psi) \in [0, \infty) \times [0, 2\pi)$.*

The analytical Fourier-Mellin transform fulfils something similar to the shift property of the Fourier transform.

Lemma 2.18. *Let $g \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$ and \tilde{g} a rotated and scaled version of g , that is,*

$$(\tilde{g} \circ \mathcal{P})(r, \psi) = (g \circ \mathcal{P})(\sigma r, \psi + \rho),$$

with a scaling factor $\sigma > 0$ and a rotation angle $\rho \in [0, 2\pi)$. Then,

$$\mathcal{AFM}_{\tilde{g}}(u, v) = \sigma^{-\gamma + i v} e^{2\pi i u \rho} \mathcal{AFM}_g(u, v), \quad (u, v) \in \mathbb{Z} \times \mathbb{R}.$$

Proof. We have

$$\begin{aligned} \mathcal{AFM}_{\tilde{g}}(u, v) &= \int_0^\infty \int_0^{2\pi} e^{-2\pi i u \psi} r^{\gamma - i v} (\tilde{g} \circ \mathcal{P})(r, \psi) \, d\psi \frac{dr}{r} \\ &= \int_0^\infty \int_0^{2\pi} e^{-2\pi i u \psi} r^{\gamma - i v} (g \circ \mathcal{P})(\sigma r, \psi + \rho) \, d\psi \frac{dr}{r} \\ &= \int_0^\infty \int_\rho^{2\pi + \rho} e^{-2\pi i u (\tilde{\psi} - \rho)} (\tilde{r}/\sigma)^{\gamma - i v} (g \circ \mathcal{P})(\tilde{r}, \tilde{\psi}) \, d\tilde{\psi} \frac{d\tilde{r}}{\tilde{r}} \\ &= \sigma^{-\gamma + i v} e^{2\pi i u \rho} \int_0^\infty \int_0^{2\pi} e^{-2\pi i u \tilde{\psi}} \tilde{r}^{\gamma - i v} (g \circ \mathcal{P})(\tilde{r}, \tilde{\psi}) \, d\tilde{\psi} \frac{d\tilde{r}}{\tilde{r}} \\ &= \sigma^{-\gamma + i v} e^{2\pi i u \rho} \mathcal{AFM}_g(u, v), \end{aligned}$$

where $\tilde{r} := \sigma r$, $\tilde{\psi} := \psi + \rho$, and g is 2π -periodic in ψ . □

Remark 2.19. *Note, that by Remark 2.15, we get $\mathcal{AFM}_g(u, v) = \mathcal{F}_{\tilde{g}}(u, v)$ with*

$$\tilde{g}: [0, \infty) \times [0, 2\pi), \quad (r, \psi) \mapsto r^\gamma (g \circ \mathcal{LP})(r, \psi).$$

In particular, the Plancherel theorem (Theorem 2.12) holds also for the analytical Fourier-Mellin transform.

Chapter 3

Model

Now, we introduce our model for the image sequences from SMS microscopy. Note that we initially use notations like \tilde{O}_j^t for the observations or $\tilde{\epsilon}_j^t$ for the statistical errors because we will soon consider binned (i.e., averaged in time) versions of the data frames which we will mainly work with (see Section 3.1). The resulting variables will then be called O_j^t , ϵ_j^t , and so on.

Let a function $\tilde{f}: \mathbb{R}^2 \rightarrow \mathbb{R}$ represent a grey value image. We will consider a shifted, rotated and scaled version of \tilde{f} , for which we define the following functions.

Definition 3.1 (Drift, rotation, and scaling functions). *Let $d_1, d_2, d_3 \in \mathbb{N}$ and let $\Theta \subseteq \mathbb{R}^{d_1}$, $\Phi \subseteq \mathbb{R}^{d_2}$, $A \subseteq \mathbb{R}^{d_3}$ compact subsets. Let $0 < \sigma_{\min} < 1 < \sigma_{\max} < \infty$.*

- For $\theta \in \Theta$, we call $\delta^\theta: [0, 1] \rightarrow \mathbb{R}^2$, $t \mapsto \delta_t^\theta$, a drift function.
- For $\phi \in \Phi$, we call $\rho^\phi: [0, 1] \rightarrow (-\pi/2, \pi/2]$, $t \mapsto \rho_t^\phi$, a rotation function.
- For $\alpha \in A$, we call $\sigma^\alpha: [0, 1] \rightarrow [\sigma_{\min}, \sigma_{\max}]$, $t \mapsto \sigma_t^\alpha$, a scaling function.

We denote the true and unknown parameters by $\theta_0 \in \Theta$, $\phi_0 \in \Phi$, and $\alpha_0 \in A$.

Remark 3.2 (Interpretation). *The interval $[0, 1]$ represents a (standardised) time period, in which we observe a sequence of transformed images. For each $t \in [0, 1]$, $\theta \in \Theta$, $\phi \in \Phi$, and $\alpha \in A$,*

- $\delta_t^\theta \in \mathbb{R}^2$ is a translation vector,
- $\rho_t^\phi \in (-\pi/2, \pi/2]$ is a rotation angle,
- $\sigma_t^\alpha \in [\sigma_{\min}, \sigma_{\max}]$ is a scaling factor,

such that, with R_ρ from (1.2),

$$x \mapsto \tilde{f}\left(1/\sigma_t^\alpha \cdot R_{-\rho_t^\phi}(x - \delta_t^\theta)\right)$$

is a version of the image \tilde{f} which is first shifted by δ_t^θ and then rotated by ρ_t^ϕ and scaled by σ_t^α . The bounds on the scaling function σ_t^α are useful for technical reasons (see the proofs in Chapter 6) but are also naturally given by the experimental setting, σ_{\min} by the resolution of the microscope (or the pixel size) and σ_{\max} by the size of the observation window. The values of the rotation function ρ_t^ϕ have to be in an interval of length π to ensure that we get unique estimates of the rotation angles, which will be explained later in Remark 4.4. Of course, that interval has to contain 0 to allow for $\rho_0^\phi = 0$ as set in Assumption 3.3. We choose the interval $(-\pi/2, \pi/2]$ because it is symmetric and we want to allow clockwise and counter-clockwise rotations alike.

Assumption 3.3. *Since we do not expect drift, rotation, or scaling at time $t = 0$, we assume that $\delta_0^\theta = 0$, $\rho_0^\phi = 0$, and $\sigma_0^\alpha = 1$ for all $\theta \in \Theta$, $\phi \in \Phi$, and $\alpha \in \Lambda$.*

For each $t \in [0, 1]$, we define

$$\tilde{f}^t: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad x \mapsto \tilde{f}\left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(x - \delta_t^{\theta_0})\right). \quad (3.1)$$

We observe

$$\tilde{O}_j^t := \begin{cases} \tilde{f}^t(x_j) + \tilde{\nu}_j^t \tilde{\epsilon}_j^t, & \text{if } t \in \tilde{\mathbb{T}}, j \in \tilde{\mathcal{J}}^t, \\ 0, & \text{if } t \in \tilde{\mathbb{T}}, j \in \{1, \dots, n\} \setminus \tilde{\mathcal{J}}^t, \end{cases} \quad (3.2)$$

where $\tilde{\mathbb{T}} := \{0, 1/T, \dots, (T-1)/T\}$ is a set of equidistant time points with $T \in \mathbb{N}$, $x_1, \dots, x_n \in [0, 1]^2$ are pixel locations, $\tilde{\mathcal{J}}^t \subseteq \{1, \dots, n\}$ with a total number of pixels $n \in \mathbb{N}$, $\tilde{\nu}_j^t > 0$, and $\tilde{\epsilon}_j^t \sim \mathcal{N}(0, 1)$ such that $\tilde{\epsilon}_j^t$ is independent of $\tilde{\epsilon}_{j'}^{t'}$, except if $j = j'$ and $t = t'$.

Definition 3.4 (Frames). *For each $t \in \tilde{\mathbb{T}}$, we call the observed image at time t , $(\tilde{O}_j^t)_{j \in \tilde{\mathcal{J}}^t}$, a frame.*

Remark 3.5 (Order of transformations). *We choose the specific order in which we apply the affine transformations in (3.1) because we will correct the observed images in two steps: first for rotation and scaling and then for drift (see Chapter 4). In practice, however, this is not necessarily a restriction. Consider the alternatively transformed image*

$$\tilde{f}_{\text{alt}}^t: x \mapsto \tilde{f}\left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}x - \delta_t^{\theta_0}\right), \quad (3.3)$$

which is first rotated and scaled and then shifted. Since the rotation and scaling is easily invertible, we can define a new family of drift functions

$$\left\{ \tilde{\delta}^{\theta; \phi; \alpha}: [0, 1] \rightarrow \mathbb{R}^2, t \mapsto \sigma_t^\alpha \cdot R_{\rho_t^\phi} \delta_t^\theta \mid \theta \in \Theta, \phi \in \Phi, \alpha \in \Lambda \right\},$$

such that

$$\tilde{f}_{\text{alt}}^t(x) = \tilde{f}\left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(x - \tilde{\delta}_t^{\theta_0; \phi_0; \alpha_0})\right),$$

which has the form given in (3.1). Obviously, $\tilde{\delta}^{\theta; \phi; \alpha}$ depends on the rotation and scaling of the image sequence and will usually not have the same form as the originally proposed δ^θ . Hence, it might be useful to implement a suitably flexible drift model δ^θ in (3.1) such that it can reasonably approximate $\tilde{\delta}^{\theta; \phi; \alpha}$ in the case that the actual model looks like (3.3).

Remark 3.6 (Center of rotation and scaling). *Note that, in model (3.1), we assume that the center of rotation and scaling is always in the middle of the image domain. Considering a more complex model featuring a shifting center point will be the subject of further research.*

3.1 Binning

In SMS-microscopy, each frame \tilde{O}^t typically contains very little information because the number of observed pixels $\#\tilde{J}^t$ is small, whereas the length T of the image sequence is quite huge. It might thus be interesting to *bin* subsequent frames, that is, take the point-wise average of them, to increase the information per frame or just to reduce the memory needed to process calibration methods on the sequence. Here, however, binning serves mainly as a means to asymptotically reduce the noise level of the data.

Definition 3.7 (Indicator function). *For a set A and a subset $B \subseteq A$, we define the indicator function*

$$\mathbb{I}_B: A \rightarrow \{0, 1\}, \quad x \mapsto \begin{cases} 1, & \text{if } x \in B, \\ 0, & \text{if } x \in A \setminus B. \end{cases}$$

Note, that we omit the set A from the notation \mathbb{I}_B . When we use the indicator function, the identity of A will be clear from the context.

For all $T \in \mathbb{N}$, we define a *bin size* $\beta_T \in \mathbb{N}$ such that $T/\beta_T \in \mathbb{N}$. We construct a new image sequence of length T/β_T by averaging over β_T subsequent frames,

$$O_j^t := \frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \tilde{O}_j^{t+i/T}, \quad t \in \mathbb{T}, j \in \{1, \dots, n\}, \quad (3.4)$$

where

$$\mathbb{T} := \left\{ 0, \frac{1}{T/\beta_T}, \frac{2}{T/\beta_T}, \dots, \frac{T/\beta_T - 1}{T/\beta_T} \right\} = \left\{ 0, \frac{\beta_T}{T}, \frac{2\beta_T}{T}, \dots, \frac{T - \beta_T}{T} \right\}.$$

For T large and β_T small enough (i.e., $\beta_T = o(T)$), we assume that the drift, rotation, and scaling functions are approximately constant over β_T subsequent frames (i.e., $\tilde{f}^s = \tilde{f}^{s'}$ for all $s, s' \in \{t + i/T \mid i \in \{0, \dots, \beta_T - 1\}\}$). Thus,

$$\begin{aligned} O_j^t &= \frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}_{t+i/T}}(j) \tilde{f}^{t+i/T}(x_j) + \frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}_{t+i/T}}(j) \tilde{\nu}_j^{t+i/T} \tilde{\epsilon}_j^{t+i/T} \\ &= \frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}_{t+i/T}}(j) \tilde{f}^{t+i/T}(x_j) + \sqrt{\frac{n_t}{\beta_T}} \nu_j^t \epsilon_j^t, \quad t \in \mathbb{T}, j \in \{1, \dots, n\}, \end{aligned}$$

where $J_T^t := \bigcup_{i=0}^{\beta_T-1} \tilde{J}^{t+i/T}$, $n_t := \#J_T^t$, $\epsilon_j^t \sim \mathcal{N}(0, 1)$ are pairwise independent, and

$$(\nu_j^t)^2 := \frac{1}{n_t \beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}_{t+i/T}}(j) (\tilde{\nu}_j^{t+i/T})^2.$$

Let $\mathbb{X}^t := \{x_j \mid j \in J_T^t\}$ and $g^t: \mathbb{X}^t \rightarrow \mathbb{R}$ with

$$g^t(x_j) := \frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}_{t+i/T}}(j) \tilde{f}^{t+i/T}(x_j).$$

The discrete Fourier transform of g^t is

$$\mathcal{F}_{g^t}^{\mathbb{X}^t}(\omega) = \frac{1}{n_t} \sum_{j \in J_T^t} e^{-2\pi i \langle \omega, x_j \rangle} \frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}_{t+i/T}}(j) \tilde{f}^{t+i/T}(x_j). \quad (3.5)$$

Assumption 3.8. Assume that $\beta_T \xrightarrow{T \rightarrow \infty} \infty$, $\beta_T = o(T)$, $J_T^t \nearrow \{1, \dots, n\}$ as $T \rightarrow \infty$, and that

$$\frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}^{t+i/T}}(j) \xrightarrow{T \rightarrow \infty} c^t(x_j)$$

for all $t \in [0, 1]$, $j \in \{1, \dots, n\}$, with functions $c^t: \mathbb{R}^2 \rightarrow (0, 1)$. Assume further that the drift, rotation, and scaling functions are right-continuous in t and that \tilde{f} is continuous.

Lemma 3.9. Under Assumption 3.8, $\mathcal{F}_{f^t}^{\mathbb{X}}(\omega)$ converges to

$$\mathcal{F}_{f^t}^{\mathbb{X}}(\omega) = \frac{1}{n} \sum_{j=1}^n e^{-2\pi i \langle \omega, x_j \rangle} f^t(x_j),$$

where $\mathbb{X} := \{x_j \mid j \in \{1, \dots, n\}\}$ and $f^t := c^t \tilde{f}^t$.

Proof. Because of the continuity of \tilde{f} , the right-continuity of the transformations, and the fact that $i < \beta_T = o(T)$, we have for a fixed $t \in [0, 1]$ that

$$\tilde{f}^{t+i/T}(x_j) - \tilde{f}^t(x_j) \xrightarrow{T \rightarrow \infty} 0 \quad \text{uniformly in } i.$$

Since we observe \tilde{f} only on a compact subset of \mathbb{R}^2 and a continuous function is always uniformly continuous on a compact set, it follows that

$$\begin{aligned} & \frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}^{t+i/T}}(j) \tilde{f}^{t+i/T}(x_j) \\ &= \frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}^{t+i/T}}(j) (\tilde{f}^{t+i/T}(x_j) - \tilde{f}^t(x_j)) + \frac{1}{\beta_T} \sum_{i=0}^{\beta_T-1} \mathbb{I}_{\tilde{J}^{t+i/T}}(j) \tilde{f}^t(x_j) \\ & \xrightarrow{T \rightarrow \infty} c^t(x_j) \tilde{f}^t(x_j). \end{aligned}$$

The assertion now follows from (3.5) and the fact that $J_T^t \nearrow \{1, \dots, n\}$. \square

Remark 3.10 (Asymptotic order of β_T). *The asymptotic behaviour $\beta_T \rightarrow \infty$ ensures that, for larger T , we bin more frames and thus, the sizes of the J_T^t 's tend to increase. On the other hand, the fact that $\beta_T = o(T)$ ensures that the time period represented by each binned frame decreases. Note that we can assure these asymptotics for β_T only for T in a subsequence of \mathbb{N} if we also require $\beta_T \in \mathbb{N}$ and $T/\beta_T \in \mathbb{N}$. However, dropping one of the latter requirements would lead to a lot of additional notation and we will therefore keep them for the sake of simplicity.*

Remark 3.11 (Interpretation of c^t and f^t). *In SMS-microscopy, the probability of registering something at time t and pixel location x_j depends mainly on the marker density $\tilde{f}^t(x_j)$ at this point in time and space and on the switching-light intensity. Hence, $c^t(x_j)$ is proportional to $\tilde{f}^t(x_j)$, meaning that we basically observe a noisy version of $(\tilde{f}^t)^2$ instead of \tilde{f}^t .*

If we instead consider a Bernoulli model meaning that each pixel location x_j is active at time t with some probability $p_j^t \in [0, 1]$ and inactive with probability $1 - p_j^t$, and if $t \mapsto p_j^t$ is right-continuous for each $j \in \{1, \dots, n\}$, then we have $c^t(x_j) = p_j^t$.

Whatever the truth is, we cannot observe \tilde{f}^t without c^t and will therefore reconstruct $f := c^0 \tilde{f}$ instead of \tilde{f} itself.

For simplicity's sake, we assume that

$$\mathcal{F}_{f^t}^{\mathbb{X}}(\omega) \approx \mathcal{F}_{f^t}(\omega). \quad (3.6)$$

From Lemma 2.5, we get

$$\mathcal{F}_{f^t}(\omega) = (\sigma_t^{\alpha_0})^2 e^{-2\pi i \langle \omega, \delta_t^{\theta_0} \rangle} \mathcal{F}_f(\sigma_t^{\alpha_0} R_{-\rho_t \phi_0} \omega). \quad (3.7)$$

Instead of the original observations $O^t: j \mapsto O_j^t$, we now consider their squared Fourier magnitudes $|\mathcal{F}_{O^t}^{\mathbb{X}^t}(\omega)|^2$, where

$$\mathcal{F}_{O^t}^{\mathbb{X}^t}(\omega) = \frac{1}{n_t} \sum_{j \in J_T^t} e^{-2\pi i \langle \omega, x_j \rangle} O_j^t = \mathcal{F}_{g^t}^{\mathbb{X}^t}(\omega) + W^t(\omega)$$

with

$$W^t(\omega) := \frac{1}{\sqrt{n_t} \beta_T} \sum_{j \in J_T^t} e^{-2\pi i \langle \omega, x_j \rangle} \nu_j^t \epsilon_j^t. \quad (3.8)$$

Because of Lemma 3.9, (3.6), and (3.7), in the following, we will work with the approximation

$$\mathcal{F}_{O^t}^{\mathbb{X}^t}(\omega) \approx Y^t(\omega),$$

where $Y^t(\omega)$ is defined below.

Definition 3.12 (Fourier Model). *For $t \in \mathbb{T}$ and $\omega \in \mathbb{R}^2$, we define*

$$Y^t(\omega) := \mathcal{F}_{f^t}(\omega) + W^t(\omega), \quad (3.9)$$

with $W^t(\omega)$ from (3.8).

By Lemma 2.5,

$$\mathcal{F}_{f^t}(\omega) = (\sigma_t^{\alpha_0})^2 e^{-2\pi i \langle \omega, \delta_t^{\theta_0} \rangle} \mathcal{F}_f(\sigma_t^{\alpha_0} R_{-\rho_t \phi_0} \omega),$$

which implies

$$|\mathcal{F}_{f^t}|^2(\omega) = (\sigma_t^{\alpha_0})^4 \left| \mathcal{F}_f(\sigma_t^{\alpha_0} R_{-\rho_t \phi_0} \omega) \right|^2. \quad (3.10)$$

Note that the $|\mathcal{F}_{f^t}|^2$'s do not depend on the drift δ^θ . We aim to estimate the rotation parameter ϕ and the scaling parameter α from $\{|Y^t|^2\}_{t \in \mathbb{T}}$. Then, we can calibrate the images f^t with the estimated rotation and scaling, leaving only the drift to be estimated.

Because of (3.9), the analytical Fourier-Mellin transform of $|Y^t|^2$ is

$$\begin{aligned} \mathcal{AFM}_{|Y^t|^2}(u, v) &= \int_0^\infty \int_0^{2\pi} e^{-2\pi i u \psi} r^{-iv} r^\gamma (|Y^t|^2 \circ \mathcal{P})(r, \psi) d\psi \frac{dr}{r} \\ &= \int_0^\infty \int_0^{2\pi} e^{-2\pi i u \psi} r^{-iv} r^\gamma (|\mathcal{F}_{f^t}|^2 \circ \mathcal{P} + W^t \circ \mathcal{P})(r, \psi) d\psi \frac{dr}{r} \\ &= \mathcal{AFM}_{|\mathcal{F}_{f^t}|^2}(u, v) + \mathcal{AFM}_{W^t}(u, v), \end{aligned} \quad (3.11)$$

where

$$W^t(\omega) := |W^t(\omega)|^2 + 2\Re(\mathcal{F}_{f^t}(\omega) \overline{W^t(\omega)}). \quad (3.12)$$

In the next chapter, we explain how to estimate the unknown drift, rotation, and scaling parameters θ_0 , ϕ_0 , and α_0 .

Chapter 4

M-estimation method

In this chapter, we describe a method for the estimation of the drift, rotation, and scaling parameters θ_0 , ϕ_0 , and α_0 based on M-estimation. This means that we define certain functions (called contrast functionals) depending on the data which we minimize with respect to the parameters θ , ϕ , and α .

From Lemma 2.18 and (3.10), we get

$$\mathcal{AFM}_{|\mathcal{F}_t|}^2(u, v) = d_{u,v}(1/\sigma_t^{\alpha_0}, -\rho_t^{\phi_0})F^t(u, v), \quad (4.1)$$

where

$$d_{u,v}(\sigma, \rho) := \sigma^{-iv} e^{2\pi i u \rho}, \quad (4.2)$$

$$F^t(u, v) := (\sigma_t^{\alpha_0})^{4-\gamma} \mathcal{AFM}_{|\mathcal{F}_t|}^2(u, v). \quad (4.3)$$

This, together with (3.11), gives the idea that (if the error terms get small)

$$d_{u,v}(\sigma_t^{\alpha_0}, \rho_t^{\phi_0}) \mathcal{AFM}_{|Y_t|^2}(u, v) \approx F^t(u, v).$$

Based on this, we define an estimator for the scaling and rotation parameters as a minimizer of a certain functional.

Definition 4.1 (Contrast functionals for rotation and scaling). *With suitable cutoffs $u_T, v_T > 0$ (as specified in the Theorems of Chapter 6), we define the empirical contrast functional (for rotation and scaling),*

$$\begin{aligned} M_T(\phi, \alpha) &:= \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| d_{u,v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{AFM}_{|Y_t|^2}(u, v) \right. \\ &\quad \left. - \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} d_{u,v}(\sigma_{t'}^\alpha, \rho_{t'}^\phi) \mathcal{AFM}_{|Y_{t'}|^2}(u, v) \right|^2 dv \\ &= M_T^0 + \widetilde{M}_T(\phi, \alpha), \end{aligned}$$

with

$$\begin{aligned} M_T^0 &:= \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| \mathcal{AFM}_{|Y_t|^2}(u, v) \right|^2 dv, \\ \widetilde{M}_T(\phi, \alpha) &:= - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{AFM}_{|Y_t|^2}(u, v) \right|^2 dv, \end{aligned}$$

where we used Corollary A.3 in the Appendix and $|d_{u,v}(\sigma, \rho)| = 1$. Similarly, we define the population contrast functional (for rotation and scaling),

$$\begin{aligned} M(\phi, \alpha) &:= \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \int_0^1 \left| d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) \right. \\ &\quad \left. - \int_0^1 d_{u,v}(\sigma_{t'}^\alpha / \sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0}) F^{t'}(u, v) dt' \right|^2 dt dv \\ &= M^0 + \widetilde{M}(\phi, \alpha), \end{aligned}$$

with

$$\begin{aligned} M^0 &:= \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \int_0^1 |F^t(u, v)|^2 dt dv, \\ \widetilde{M}(\phi, \alpha) &:= - \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \int_0^1 d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) dt \right|^2 dv, \end{aligned}$$

where we used Corollary A.3 in the Appendix and $|d_{u,v}(\sigma, \rho)| = 1$, again.

Remark 4.2. Note, that M_T^0 and M^0 are constant in (ϕ, α) . Minimizing M_T or M is therefore equivalent to minimizing \widetilde{M}_T or \widetilde{M} , respectively.

Definition 4.3. (Scaling and rotation parameter estimator) We define

$$(\hat{\phi}_T, \hat{\alpha}_T) \in \operatorname{argmin}_{(\phi, \alpha) \in \mathbb{A} \times \Phi} \widetilde{M}_T(\phi, \alpha).$$

Remark 4.4 (Bounds on rotation angles). We can now explain the restriction of the rotation function to values in an interval of length π . By Lemma 2.10, the squared Fourier magnitude $|\mathcal{F}_f|^2$ (and hence its analytical Fourier-Mellin transform) does not change if we rotate f by π . Thus, allowing ρ_t^ϕ to assume values in an interval of length larger than π could result in an estimated rotation angle that is shifted by π which in turn would produce frames that are turned upside down.

In the next section, the Theorems 6.13 and 6.23 will show that, under some assumptions, $(\hat{\phi}_T, \hat{\alpha}_T)$ is a consistent estimator for (ϕ_0, α_0) and asymptotically centred normal with rate \sqrt{T} .

The next step is to calibrate the Fourier data Y^t with those estimators which leads to the following model.

Definition 4.5 (Fourier model after rotation and scaling correction). We define

$$\begin{aligned} Z_T^t(\omega) &:= (\sigma_t^{\hat{\alpha}_T})^{-2} Y^t \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{\rho_t^{\hat{\phi}_T}} \omega \right) \\ &= h_\omega \left(-1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^{\theta_0} \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) + V_T^t(\omega), \end{aligned} \quad (4.4)$$

where

$$h_\omega(\delta) := e^{2\pi i \langle \omega, \delta \rangle} \quad (4.5)$$

and

$$\begin{aligned} V_T^t(\omega) &:= (\sigma_t^{\hat{\alpha}_T})^{-2} W^t \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{\rho_t^{\hat{\phi}_T}} \omega \right) \\ &= \frac{1}{\sqrt{n_t} \beta_T (\sigma_t^{\hat{\alpha}_T})^2} \sum_{j \in J_T^t} \exp \left(-2\pi i \left\langle 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{\rho_t^{\hat{\phi}_T}} \omega, x_j \right\rangle \right) \nu_j^t \epsilon_j^t. \end{aligned} \quad (4.6)$$

Similar to the estimation of the rotation and scaling parameters, we minimize a certain contrast functional to estimate the true drift parameter θ_0 .

Definition 4.6 (Contrast functionals for drift). *For a suitable cutoff $r_T > 0$ (to be specified in Chapter 6), let*

$$\Omega_T := \{\omega \in \mathbb{R}^2 \mid \|\omega\| < r_T\} \quad (4.7)$$

the (open) Euclidean ball with centre $0 \in \mathbb{R}^2$ and radius r_T and define the empirical contrast functional (for drift),

$$\begin{aligned} N_T(\theta) &:= \int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^\theta \right) Z_T^t(\omega) \right. \\ &\quad \left. - \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} h_\omega \left(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^\theta \right) Z_{T'}^{t'}(\omega) \right|^2 d\omega \\ &= N_T^0 + \tilde{N}_T(\theta), \end{aligned}$$

with

$$\begin{aligned} N_T^0 &:= \int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} |Z_T^t(\omega)|^2 d\omega, \\ \tilde{N}_T(\theta) &:= - \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^\theta \right) Z_T^t(\omega) \right|^2 d\omega, \end{aligned}$$

where we used Corollary A.3 in the Appendix and $|h_\omega(\delta)| = 1$. Similarly, we define the population contrast functional (for drift),

$$\begin{aligned} N(\theta) &:= \int_{\mathbb{R}^2} \int_0^1 \left| h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \mathcal{F}_f(\omega) \right. \\ &\quad \left. - \int_0^1 h_\omega \left(1/\sigma_{t'}^{\alpha_0} \cdot R_{-\rho_{t'}^{\phi_0}} (\delta_{t'}^\theta - \delta_{t'}^{\theta_0}) \right) \mathcal{F}_f(\omega) dt' \right|^2 dt d\omega \\ &= N^0 + \tilde{N}(\theta), \end{aligned}$$

where

$$\begin{aligned} N^0 &:= \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 d\omega, \\ \tilde{N}(\theta) &:= - \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \left| \int_0^1 h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) dt \right|^2 d\omega, \end{aligned}$$

where we used Corollary A.3 in the Appendix and $|h_\omega(\delta)| = 1$, again.

Remark 4.7. *Note, that N_T^0 and N^0 are constant in θ . Minimizing N_T or N is therefore equivalent to minimizing \tilde{N}_T or \tilde{N} , respectively.*

We can now define estimators for the drift parameter θ_0 and the unknown image f .

Definition 4.8. (Drift parameter estimator) *We define*

$$\hat{\theta}_T := \operatorname{argmin}_{\theta \in \Theta} \tilde{N}_T(\theta).$$

Definition 4.9. (*Image estimator*) We define the estimator for f as the inverse Fourier transform of the calibrated Fourier data,

$$\hat{f}_T(x_j) := \int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} e^{2\pi i \langle \omega, x_j \rangle} h_\omega(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^{\hat{\theta}_T}) Z_T^t(\omega) d\omega, \quad j \in \{1, \dots, n\}.$$

Theorem 6.33 in Chapter 6 will show that $\hat{\theta}_T$ and \hat{f}_T are consistent estimators for θ_0 and f , respectively. Furthermore, Theorem 6.42 will show that $(\sqrt{T}(\hat{\theta}_T - \theta_0))_{T \in \mathbb{N}}$ is uniformly tight.

We summarize the M-estimation procedure in the following algorithm.

Algorithm 4.10 (Image correction via M-estimation). Choose cutoffs $r_T, u_T, v_T > 0$, $\gamma > 0$, and parametric models for the drift function δ^θ , the rotation function ρ^ϕ , and the scaling function σ^α .

1. Given a sequence of frames $(O^t)_{t \in \mathbb{T}}$, approximate the squared Fourier magnitudes $|Y^t|^2$ with $|\mathcal{F}_{O^t}^\times|^2$, $t \in \mathbb{T}$.
2. Compute $\mathcal{AFM}_{|Y^t|^2}^T$, $t \in \mathbb{T}$.
3. Estimate (ϕ_0, α_0) with the minimizer $(\hat{\phi}_T, \hat{\alpha}_T)$ of $\widetilde{M}_T(\phi, \alpha)$.
4. Calibrate the Fourier frames $(Y^t)_{t \in \mathbb{T}}$ with those estimators, to get

$$Z_T^t(\omega) := (\sigma_t^{\hat{\alpha}_T})^{-2} Y^t \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{\rho_t^{\hat{\phi}_T}} \omega \right).$$

5. Estimate θ_0 with the minimizer $\hat{\theta}_T$ of $\widetilde{N}_T(\theta)$.
6. Calibrate the $(Z_T^t)_{t \in \mathbb{T}}$ with those estimators.
7. Compute the image estimator \hat{f}_T as the inverse Fourier transform of the resulting calibrated Fourier frames.

Remark 4.11. In the simulation study and the application to SMS data (Chapters 9 and 10), we will use the well-known fast Fourier transform algorithm (FFT) to compute the Fourier transforms of the frames. Moreover, we will employ Remark 2.15 to compute the analytical Fourier-Mellin transform via a log-polar coordinate transform of (the squared Fourier magnitudes of) the frames and a subsequent FFT.

Chapter 5

Preliminary Computations

In this chapter, we prove some properties of the error terms $W^t(\omega)$, $\mathcal{AFM}_{\mathcal{W}^t}^T(u, v)$, and $V_T^t(\omega)$. Moreover, we summarize some properties of the terms $d_{(u,v)}(\sigma, \rho)$ and $h_\omega(\delta)$ and compute the first and second derivatives of $d_{(u,v)}(\sigma_t^\alpha, \rho_t^\phi)$ with respect to (ϕ, α) and the derivatives of $h_\omega(\delta_t^\theta)$ with respect to θ .

Lemma 5.1 (Properties of $W^t(\omega)$). *Recall $W^t(\omega)$ from (3.8).*

1. *The real and imaginary parts of $W^t(\omega)$ are centred normal,*

$$\begin{aligned}\Re(W^t(\omega)) &\sim \mathcal{N}\left(0, \frac{1}{n_t \beta_T} \sum_{j \in J_T^t} \left(\cos(2\pi \langle \omega, x_j \rangle)\right)^2 (\nu_j^t)^2\right), \\ \Im(W^t(\omega)) &\sim \mathcal{N}\left(0, \frac{1}{n_t \beta_T} \sum_{j \in J_T^t} \left(\sin(2\pi \langle \omega, x_j \rangle)\right)^2 (\nu_j^t)^2\right).\end{aligned}$$

2. *We have $\mathbb{E}\left(|W^t(\omega)|^2\right) = \frac{1}{n_t \beta_T} \sum_{j \in J_T^t} (\nu_j^t)^2$.*

3. *$W^t(\omega)$ and $W^{t'}(\omega')$ are independent unless $t = t'$,*

4. *For $t \in [0, 1]$, $j, j' \in J_T^t$, and $\omega \in \mathbb{R}^2$, define*

$$\Re_\omega^t := \Re(\mathcal{F}_{f^t}(\omega) \overline{W^t(\omega)}), \quad S_\omega^t := |W^t(\omega)|^2, \quad \cos_\omega^{j,j'} := \cos(2\pi \langle \omega, x_j - x_{j'} \rangle).$$

With this, we have $\mathbb{E}(\Re_\omega^t S_{\omega'}^t) = 0$ and

$$\begin{aligned}\mathbb{E}(\Re_\omega^t \Re_{\omega'}^t) &= \frac{1}{n_t^3 \beta_T} \sum_{j, j', j'' \in J_T^t} \cos_\omega^{j,j'} \cos_{\omega'}^{j'',j'} f^t(x_j) f^t(x_{j''}) (\nu_j^t)^2, \\ \mathbb{E}(S_\omega^t S_{\omega'}^t) &= \frac{1}{n_t^2 \beta_T^2} \left[3 \sum_{j \in J_T^t} (\nu_j^t)^4 + \sum_{j \neq j'} \left(1 + 2 \cos_\omega^{j,j'} \cos_{\omega'}^{j,j'}\right) (\nu_j^t)^2 (\nu_{j'}^t)^2 \right].\end{aligned}$$

Proof.

1. Since they are linear combinations of independent centred Gaussian random variables $\{\epsilon_j^t \mid j \in J_T^t\}$, $\Re(W^t(\omega))$ and $\Im(W^t(\omega))$ are also centred Gaussian. Because

the ϵ_j^t are standard normal and independent, we get

$$\begin{aligned}\text{Var}(\Re(W^t(\omega))) &= \frac{1}{n_t \beta_T} \sum_{j \in J_T^t} \text{Var}\left(\cos(2\pi \langle \omega, x_j \rangle) \nu_j^t \epsilon_j^t\right) \\ &= \frac{1}{n_t \beta_T} \sum_{j \in J_T^t} \left(\cos(2\pi \langle \omega, x_j \rangle)\right)^2 (\nu_j^t)^2,\end{aligned}$$

and, similarly,

$$\begin{aligned}\text{Var}(\Im(W^t(\omega))) &= \frac{1}{n_t \beta_T} \sum_{j \in J_T^t} \text{Var}\left(-\sin(2\pi \langle \omega, x_j \rangle) \nu_j^t \epsilon_j^t\right) \\ &= \frac{1}{n_t \beta_T} \sum_{j \in J_T^t} \left(\sin(2\pi \langle \omega, x_j \rangle)\right)^2 (\nu_j^t)^2,\end{aligned}$$

2. This follows from

$$\begin{aligned}\mathbb{E}\left(|W^t(\omega)|^2\right) &= \mathbb{E}\left(\Re(W^t(\omega))^2\right) + \mathbb{E}\left(\Im(W^t(\omega))^2\right) \\ &= \text{Var}\left(\Re(W^t(\omega))\right) + \text{Var}\left(\Im(W^t(\omega))\right)\end{aligned}$$

and the fact that $\cos^2(x) + \sin^2(x) = 1$ for all $x \in \mathbb{R}$.

3. Since $\{\epsilon_j^t \mid j \in J_T^t\}$ and $\{\epsilon_{j'}^{t'} \mid j' \in J_T^{t'}\}$ are independent for $t \neq t'$, so are $W^t(\omega)$ and $W^{t'}(\omega')$.

4. First, note that

$$\begin{aligned}\Re(\mathcal{F}_{f^t}(\omega) \overline{W^t(\omega)}) &= \Re\left(n_t^{-3/2} \beta_T^{-1/2} \sum_{j, j' \in J_T^t} e^{-2\pi i \langle \omega, x_j \rangle} f^t(x_j) e^{2\pi i \langle \omega, x_{j'} \rangle} \nu_{j'}^t \epsilon_{j'}^t\right) \\ &= n_t^{-3/2} \beta_T^{-1/2} \sum_{j, j' \in J_T^t} \cos_{\omega}^{j, j'} f^t(x_j) \nu_{j'}^t \epsilon_{j'}^t.\end{aligned}$$

Since $\mathbb{E}(\epsilon_{j'}^t \epsilon_{j''}^t) = 0$ unless $j' = j''$ and $\mathbb{E}((\epsilon_{j'}^t)^2) = 1$, it follows that

$$\begin{aligned}\mathbb{E}\left[\Re(\mathcal{F}_{f^t}(\omega) \overline{W^t(\omega)}) \Re(\mathcal{F}_{f^{t'}}(\omega') \overline{W^{t'}(\omega')})\right] \\ = n_t^{-3} \beta_T^{-1} \sum_{j, j', j'' \in J_T^t} \cos_{\omega}^{j, j'} \cos_{\omega'}^{j'', j'} f^t(x_j) f^{t'}(x_{j''}) (\nu_{j'}^t)^2.\end{aligned}$$

Furthermore, using Theorem B.3 in the Appendix,

$$\begin{aligned}|W^t(\omega')|^2 &= \left(\frac{1}{\sqrt{n_t \beta_T}} \sum_{j'' \in J_T^t} \cos(2\pi \langle \omega', x_{j''} \rangle) \nu_{j''}^t \epsilon_{j''}^t\right)^2 \\ &\quad + \left(\frac{1}{\sqrt{n_t \beta_T}} \sum_{j'' \in J_T^t} \sin(2\pi \langle \omega', x_{j''} \rangle) \nu_{j''}^t \epsilon_{j''}^t\right)^2 \\ &= \frac{1}{n_t \beta_T} \sum_{j'', j''' \in J_T^t} \cos(2\pi \langle \omega', x_{j''} \rangle) \cos(2\pi \langle \omega', x_{j'''} \rangle) \nu_{j''}^t \nu_{j'''}^t \epsilon_{j''}^t \epsilon_{j'''}^t \\ &\quad + \frac{1}{n_t \beta_T} \sum_{j'', j''' \in J_T^t} \sin(2\pi \langle \omega', x_{j''} \rangle) \sin(2\pi \langle \omega', x_{j'''} \rangle) \nu_{j''}^t \nu_{j'''}^t \epsilon_{j''}^t \epsilon_{j'''}^t,\end{aligned}$$

which implies

$$|W^t(\omega')|^2 = \frac{1}{n_t \beta_T} \sum_{j'', j''' \in J_T^t} \cos_{\omega'}^{j'', j'''} \nu_{j''}^t \nu_{j'''}^t \epsilon_{j''}^t \epsilon_{j'''}^t.$$

Because $\mathbb{E}(\epsilon_{j'}^t \epsilon_{j''}^t \epsilon_{j'''}^t) = 0$ even if $j' = j'' = j'''$, we get

$$\begin{aligned} & \mathbb{E} \left[\Re(\mathcal{F}_{f^t}(\omega) \overline{W^t(\omega)}) |W^t(\omega')|^2 \right] \\ &= n_t^{-5/2} \beta_T^{-3/2} \sum_{j, j', j'', j''' \in J_T^t} \cos_{\omega}^{j, j'} \cos_{\omega'}^{j'', j'''} f^t(x_j) \nu_j^t \nu_{j'}^t \nu_{j''}^t \nu_{j'''}^t \mathbb{E} [\epsilon_j^t \epsilon_{j'}^t \epsilon_{j''}^t \epsilon_{j'''}^t] = 0. \end{aligned}$$

Finally, because $\mathbb{E}((\epsilon_j^t)^4) = 3$,

$$\begin{aligned} & \mathbb{E}(|W^t(\omega)|^2 |W^t(\omega')|^2) \\ &= n_t^{-2} \beta_T^{-2} \sum_{j, j', j'', j''' \in J_T^t} \cos_{\omega}^{j, j'} \cos_{\omega'}^{j'', j'''} \nu_j^t \nu_{j'}^t \nu_{j''}^t \nu_{j'''}^t \mathbb{E}(\epsilon_j^t \epsilon_{j'}^t \epsilon_{j''}^t \epsilon_{j'''}^t) \\ &= n_t^{-2} \beta_T^{-2} \left(\sum_{j=j'=j''=j'''} + \sum_{j=j' \neq j''=j'''} + \sum_{j=j'' \neq j'=j'''} + \sum_{j=j''' \neq j'=j''} \right) \\ & \quad \left(\cos_{\omega}^{j, j'} \cos_{\omega'}^{j'', j'''} \nu_j^t \nu_{j'}^t \nu_{j''}^t \nu_{j'''}^t \mathbb{E}(\epsilon_j^t \epsilon_{j'}^t \epsilon_{j''}^t \epsilon_{j'''}^t) \right) \\ &= n_t^{-2} \beta_T^{-2} \left(\sum_{j \in J_T^t} 3(\nu_j^t)^4 + \sum_{j \neq j''} (\nu_j^t)^2 (\nu_{j''}^t)^2 + 2 \sum_{j \neq j'} \cos_{\omega}^{j, j'} \cos_{\omega'}^{j, j'} (\nu_j^t)^2 (\nu_{j'}^t)^2 \right) \\ &= n_t^{-2} \beta_T^{-2} \left(3 \sum_{j \in J_T^t} (\nu_j^t)^4 + \sum_{j \neq j'} \left(1 + 2 \cos_{\omega}^{j, j'} \cos_{\omega'}^{j, j'} (\nu_j^t)^2 (\nu_{j'}^t)^2 \right) \right). \end{aligned}$$

□

Assumption 5.2. *There is a constant $\nu_{\max} \in (0, \infty)$ such that $\nu_j^t \leq \nu_{\max}$ for all $t \in [0, 1]$ and $j \in \{1, \dots, n\}$.*

Lemma 5.3 (Properties of $\mathcal{AFM}_{\mathcal{W}^t}^T(u, v)$). *Recall \mathcal{W}^t from (3.12). Under Assumption 5.2, we have*

$$\mathbb{E} \left[|\mathcal{AFM}_{\mathcal{W}^t}^T(u, v)|^2 \right] = \mathcal{O} \left(\frac{r_T^{2\gamma}}{\beta_T} \right). \quad (5.1)$$

Proof. With Lemma 5.1, we get

$$\begin{aligned} & \mathbb{E} \left[|\mathcal{AFM}_{\mathcal{W}^t}^T(u, v)|^2 \right] \\ &= \mathbb{E} \left[\left| \int_0^{r_T} \int_0^{2\pi} e^{-2\pi i u \psi} r^{\gamma - i v} (\mathcal{W}^t \circ \mathcal{P})(r, \psi) \, d\psi \frac{dr}{r} \right|^2 \right] \\ &= \int_0^{r_T} \int_0^{r_T} \int_0^{2\pi} \int_0^{2\pi} e^{-2\pi i u (\psi - \psi')} r^{\gamma - i v} (r')^{\gamma + i v} \\ & \quad \cdot \mathbb{E} \left[\left(|W^t(\mathcal{P}(r, \psi))|^2 + 2 \Re \left((\mathcal{F}_{f^t} \overline{W^t})(\mathcal{P}(r, \psi)) \right) \right) \right. \\ & \quad \left. \cdot \left(|W^t(\mathcal{P}(r', \psi'))|^2 + 2 \Re \left((\mathcal{F}_{f^t} \overline{W^t})(\mathcal{P}(r', \psi')) \right) \right) \right] \, d\psi \, d\psi' \frac{dr}{r} \frac{dr'}{r'} \end{aligned}$$

$$\begin{aligned}
&= \int_0^{r_T} \int_0^{r_T} \int_0^{2\pi} \int_0^{2\pi} e^{-2\pi i u(\psi-\psi')} r^{\gamma-iv} (r')^{\gamma+iv} \left\{ \mathbb{E} \left[|W^t(\mathcal{P}(r, \psi))|^2 |W^t(\mathcal{P}(r', \psi'))|^2 \right] \right. \\
&\quad \left. + 4 \mathbb{E} \left[\Re \left((\mathcal{F}_{f^t} \overline{W^t})(\mathcal{P}(r, \psi)) \right) \Re \left((\mathcal{F}_{f^t} \overline{W^t})(\mathcal{P}(r', \psi')) \right) \right] \right\} d\psi d\psi' \frac{dr}{r} \frac{dr'}{r'} \\
&= \int_0^{r_T} \int_0^{r_T} \int_0^{2\pi} \int_0^{2\pi} e^{-2\pi i u(\psi-\psi')} r^{\gamma-iv} (r')^{\gamma+iv} \left\{ \frac{1}{n_t^2 \beta_T^2} \left[3 \sum_{j \in J_T^t} (\nu_j^t)^4 \right. \right. \\
&\quad \left. \left. + \sum_{j \neq j'} \left(1 + 2 \cos \left(2\pi \langle \mathcal{P}(r, \psi), x_j - x_{j'} \rangle \right) \cos \left(2\pi \langle \mathcal{P}(r', \psi'), x_j - x_{j'} \rangle \right) \right) (\nu_j^t)^2 (\nu_{j'}^t)^2 \right] \right. \\
&\quad \left. + \frac{4}{n_t^3 \beta_T} \sum_{j, j', j'' \in J_T^t} \left[\cos \left(2\pi \langle \mathcal{P}(r, \psi), x_j^t - x_{j'} \rangle \right) \cos \left(2\pi \langle \mathcal{P}(r', \psi'), x_{j''}^t - x_{j'} \rangle \right) \right. \right. \\
&\quad \left. \left. \cdot f^t(x_j^t) f^t(x_{j''}^t) (\nu_{j'}^t)^2 \right] \right\} d\psi d\psi' \frac{dr}{r} \frac{dr'}{r'}.
\end{aligned}$$

In particular, because of the triangle inequality and $|\cos(x)| \leq 1$ for all $x \in \mathbb{R}$,

$$\begin{aligned}
&\mathbb{E} \left[|\mathcal{A} \mathcal{F} \mathcal{M}_{\mathcal{W}^t}^T(u, v)|^2 \right] \\
&\leq \int_0^{r_T} \int_0^{r_T} \int_0^{2\pi} \int_0^{2\pi} r^\gamma (r')^\gamma \left\{ \frac{3\nu_{\max}^4}{n_t^2 \beta_T^2} + \frac{4}{n_t^3 \beta_T} \sum_{j, j'' \in J_T^t} f^t(x_j^t) f^t(x_{j''}^t) \nu_{\max}^2 \right\} d\psi d\psi' \frac{dr}{r} \frac{dr'}{r'} \\
&= 4\pi^2 \left\{ \frac{3\nu_{\max}^4}{n_t^2 \beta_T^2} + \frac{4}{n_t^3 \beta_T} \sum_{j, j'' \in J_T^t} f^t(x_j^t) f^t(x_{j''}^t) \nu_{\max}^2 \right\} \left(\int_0^{r_T} r^{\gamma-1} dr \right)^2 \\
&= \frac{4\pi^2 r_T^{2\gamma}}{\gamma^2} \left\{ \frac{3\nu_{\max}^4}{n_t^2 \beta_T^2} + \frac{4}{n_t^3 \beta_T} \sum_{j, j'' \in J_T^t} f^t(x_j^t) f^t(x_{j''}^t) \nu_{\max}^2 \right\} \\
&\leq \frac{4\pi^2 r_T^{2\gamma}}{\gamma^2} \left\{ \frac{3\nu_{\max}^4}{\beta_T^2} + \frac{4}{\beta_T} \|f\|_\infty \nu_{\max}^2 \right\} = \mathcal{O} \left(\frac{r_T^{2\gamma}}{\beta_T} \right).
\end{aligned}$$

where we used that $n_t \geq 1$ and that f is bounded. \square

Lemma 5.4 (Asymptotic order of $V_T^t(\omega)$). *Recall $V_T^t(\omega)$ from (4.6). Under Assumption 5.2, for all $T \in \mathbb{N}$, $t, t' \in \mathbb{T}$, and $\omega \in \mathbb{R}^2$, we have $\mathbb{E} |V_T^t(\omega) \overline{V_T^{t'}(\omega)}| = \mathcal{O}(1/\beta_T)$.*

Proof. Since the ϵ_j^t are independent standard normal random variables, we get

$$\begin{aligned}
\mathbb{E} \left| V_T^t(\omega) \overline{V_T^{t'}(\omega)} \right| &= \mathbb{E} \left| \frac{1}{\sqrt{n_t n_{t'} \beta_T (\sigma_t^{\hat{\alpha}_T})^2 (\sigma_{t'}^{\hat{\alpha}_T})^2}} \sum_{j \in J_T^t} \sum_{j' \in J_T^{t'}} \exp \left(-2\pi i \langle 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{\rho_t^{\hat{\phi}_T}} \omega, x_j \rangle \right) \right. \\
&\quad \left. \cdot \exp \left(2\pi i \langle 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega, x_{j'} \rangle \right) \nu_j^t \nu_{j'}^{t'} \epsilon_j^t \epsilon_{j'}^{t'} \right| \\
&\leq \frac{1}{\sqrt{n_t n_{t'} \beta_T \sigma_{\min}^4}} \sum_{j \in J_T^t} \sum_{j' \in J_T^{t'}} \nu_j^t \nu_{j'}^{t'} \mathbb{E} \left| \epsilon_j^t \epsilon_{j'}^{t'} \right| \leq \frac{\nu_{\max}^2}{\beta_T \sigma_{\min}^4} = \mathcal{O} \left(\frac{1}{\beta_T} \right).
\end{aligned}$$

\square

Lemma 5.5 (Properties of $d_{u,v}(\sigma, \rho)$). *Recall from (4.2) that $d_{u,v}(\sigma, \rho) := \sigma^{-iv} e^{2\pi i u \rho}$. We have*

$$|d_{u,v}(\sigma, \rho)| \equiv 1, \quad (5.2)$$

$$\overline{d_{u,v}(\sigma, \rho)} = 1/d_{u,v}(\sigma, \rho) = d_{u,v}(1/\sigma, -\rho), \quad (5.3)$$

$$d_{u,v}(1, 0) = 1, \quad (5.4)$$

$$d_{u,v}(\sigma, \rho) d_{u,v}(\sigma', \rho') = d_{u,v}(\sigma\sigma', \rho + \rho'). \quad (5.5)$$

Proof. Define $\lambda := \log(\sigma)$. Then, $d_{u,v}(\sigma, \rho) = e^{i(2\pi u \rho - v \lambda)}$. The fact that $|e^{ix}| = 1$ and $\overline{e^{ix}} = e^{-ix}$ for all $x \in \mathbb{R}$ implies (5.2) and (5.3). The properties (5.4) and (5.5) follow directly from the definition. \square

Assumption 5.6. *There is a convex open neighborhood $U \subseteq \Phi \times \mathbf{A}$ of (ϕ_0, α_0) and some $C > 0$ such that $\alpha \mapsto \sigma_t^\alpha$ and $\phi \mapsto \rho_t^\phi$ are twice differentiable on U for all $t \in [0, 1]$ and*

$$\left\| \text{grad}_\phi \rho_t^\phi \right\|, \left\| \text{grad}_\alpha \sigma_t^\alpha \right\| \leq C, \quad \left\| \text{Hess}_\phi \rho_t^\phi \right\|_1, \left\| \text{Hess}_\alpha \sigma_t^\alpha \right\|_1 \leq C^2 \quad \text{uniformly in } \phi, \alpha, t.$$

Lemma 5.7 (Derivatives of $d_{u,v}(\sigma_t^\alpha, \rho_t^\phi)$). *Under Assumption 5.6, for $u \in \mathbb{Z}$, $v \in \mathbb{R}$, $t, t' \in [0, 1]$, and $(\phi, \alpha) \in U$, we define*

$$d_{u,v}^{t,t'}(\phi, \alpha) := d_{u,v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) \in \mathbb{C}, \quad (5.6)$$

$$\mathbf{a}_{u,v}^{t,t'}(\phi, \alpha) := \left(2\pi u \text{grad}_\phi^\top(\rho_t^\phi - \rho_{t'}^\phi), -v \sigma_{t'}^\alpha \sigma_t^{\alpha_0} / \sigma_t^\alpha \text{grad}_\alpha^\top(\sigma_t^\alpha / \sigma_{t'}^\alpha) \right)^\top \in \mathbb{R}^{d_2+d_3}, \quad (5.7)$$

$$\mathbf{H}_{u,v}^{t,t'}(\phi, \alpha) := \begin{pmatrix} 2\pi u \text{Hess}_\phi(\rho_t^\phi - \rho_{t'}^\phi) & 0 \\ 0 & -v \sigma_{t'}^\alpha \sigma_t^{\alpha_0} / \sigma_t^\alpha \text{Hess}_\alpha(\sigma_t^\alpha / \sigma_{t'}^\alpha) \end{pmatrix} \in \mathbb{R}^{(d_2+d_3) \times (d_2+d_3)}. \quad (5.8)$$

Note, that $d_{u,v}^{t,0}(\phi_0, \alpha_0) = d_{u,v}(1, 0) = 1$. There is a constant $\tilde{C} > 0$ (independent from u, v, t, t', ϕ , and α) such that

$$\text{grad}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi, \alpha) = \mathbf{ia}_{u,v}^{t,t'}(\phi, \alpha) d_{u,v}^{t,t'}(\phi, \alpha), \quad (5.9)$$

$$\text{Hess}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi, \alpha) = \left(\mathbf{iH}_{u,v}^{t,t'}(\phi, \alpha) - \mathbf{a}_{u,v}^{t,t'}(\phi, \alpha) \mathbf{a}_{u,v}^{t,t'}(\phi, \alpha)^\top \right) d_{u,v}^{t,t'}(\phi, \alpha), \quad (5.10)$$

$$\left\| \mathbf{a}_{u,v}^{t,t'}(\phi, \alpha) \right\| \leq \tilde{C} \|(u, v)\|, \quad (5.11)$$

$$\left\| \text{grad}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi, \alpha) \right\| \leq \tilde{C} \|(u, v)\|, \quad (5.12)$$

$$\left\| \text{Hess}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi, \alpha) \right\|_1 \leq \tilde{C} \|(u, v)\| + \tilde{C}^2 \|(u, v)\|^2. \quad (5.13)$$

Proof. First of all,

$$\begin{aligned} & \text{grad}_{(\phi,\alpha)} d_{u,v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) \\ &= \text{grad}_{(\phi,\alpha)} \exp(2\pi i u (\rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) - i v \log(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0})) \\ &= d_{u,v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) \\ & \quad \cdot \mathbf{i} \left(2\pi u \text{grad}_\phi^\top(\rho_t^\phi - \rho_{t'}^\phi), -v \sigma_{t'}^\alpha \sigma_t^{\alpha_0} / \sigma_t^\alpha \text{grad}_\alpha^\top(\sigma_t^\alpha / \sigma_{t'}^\alpha) \right)^\top, \end{aligned}$$

which proves (5.9). It follows that

$$\begin{aligned}
& \text{Hess}_{(\phi, \alpha)} d_{u,v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) \\
&= \text{grad}_{(\phi, \alpha)}^\top \left[d_{u,v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) \right. \\
&\quad \cdot i \left(2\pi u \text{grad}_\phi^\top(\rho_t^\phi - \rho_{t'}^\phi), -v \sigma_{t'}^\alpha \sigma_t^{\alpha_0} / \sigma_t^\alpha \text{grad}_\alpha^\top(\sigma_t^\alpha / \sigma_{t'}^\alpha) \right)^\top \left. \right] \\
&= d_{u,v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) \\
&\quad \cdot \left[- \left(2\pi u \text{grad}_\phi^\top(\rho_t^\phi - \rho_{t'}^\phi), -v \sigma_{t'}^\alpha \sigma_t^{\alpha_0} / \sigma_t^\alpha \text{grad}_\alpha^\top(\sigma_t^\alpha / \sigma_{t'}^\alpha) \right)^\top \right. \\
&\quad \cdot \left(2\pi u \text{grad}_\phi^\top(\rho_t^\phi - \rho_{t'}^\phi), -v \sigma_{t'}^\alpha \sigma_t^{\alpha_0} / \sigma_t^\alpha \text{grad}_\alpha^\top(\sigma_t^\alpha / \sigma_{t'}^\alpha) \right) \\
&\quad \left. + i \begin{pmatrix} 2\pi u \text{Hess}_\phi(\rho_t^\phi - \rho_{t'}^\phi) & 0 \\ 0 & -v \sigma_{t'}^\alpha \sigma_t^{\alpha_0} / \sigma_t^\alpha \text{Hess}_\alpha(\sigma_t^\alpha / \sigma_{t'}^\alpha) \end{pmatrix} \right],
\end{aligned}$$

proving (5.10). Now, let $\tilde{C}_1 := 4\pi C \sigma_{\max}^2 / \sigma_{\min}^2$ with $C > 0$ from Assumption 5.6. Then,

$$\begin{aligned}
\left\| \mathbf{a}_{u,v}^{t,t'}(\phi, \alpha) \right\|^2 &\leq 4\pi^2 u^2 \left\| \text{grad}_\phi(\rho_t^\phi - \rho_{t'}^\phi) \right\|^2 + v^2 \left(\frac{\sigma_{t'}^\alpha \sigma_t^{\alpha_0}}{\sigma_t^\alpha} \right)^2 \left\| \text{grad}_\alpha(\sigma_t^\alpha / \sigma_{t'}^\alpha) \right\|^2 \\
&\leq 4\pi^2 u^2 \left(\left\| \text{grad}_\phi \rho_t^\phi \right\| + \left\| \text{grad}_\phi \rho_{t'}^\phi \right\| \right)^2 \\
&\quad + v^2 \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha \sigma_{t'}^\alpha} \right)^2 \left(\sigma_{t'}^\alpha \left\| \text{grad}_\alpha \sigma_t^\alpha \right\| + \sigma_t^\alpha \left\| \text{grad}_\alpha \sigma_{t'}^\alpha \right\| \right)^2 \\
&\leq 16\pi^2 C^2 u^2 + 4 \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^4 C^2 v^2 \leq \tilde{C}_1^2 \|(u, v)\|^2,
\end{aligned}$$

which implies (5.11). Hence, (5.12) holds because, by (5.9),

$$\left\| \text{grad}_{(\phi, \alpha)} d_{u,v}^{t,t'}(\phi, \alpha) \right\| = \left\| \mathbf{a}_{u,v}^{t,t'}(\phi, \alpha) \right\|.$$

Furthermore, by Assumption 5.6, and because of

$$\|xy^\top\|_1 = \|x\|_1 \|y\|_1 \leq d_3 \|x\| \|y\| \quad \text{for all } x, y \in \mathbb{R}^{d_3},$$

we have that

$$\begin{aligned}
& \left\| \text{Hess}_\alpha(\sigma_t^\alpha / \sigma_{t'}^\alpha) \right\|_1 \\
&= \left\| \text{grad}_\alpha^\top \left(\frac{\sigma_{t'}^\alpha \text{grad}_\alpha \sigma_t^\alpha - \sigma_t^\alpha \text{grad}_\alpha \sigma_{t'}^\alpha}{(\sigma_{t'}^\alpha)^2} \right) \right\|_1 \\
&= \left\| (\sigma_{t'}^\alpha \text{Hess}_\alpha \sigma_t^\alpha + \text{grad}_\alpha \sigma_t^\alpha \text{grad}_\alpha^\top \sigma_{t'}^\alpha - \sigma_t^\alpha \text{Hess}_\alpha \sigma_{t'}^\alpha - \text{grad}_\alpha \sigma_{t'}^\alpha \text{grad}_\alpha^\top \sigma_t^\alpha) (\sigma_{t'}^\alpha)^{-2} \right. \\
&\quad \left. - 2(\sigma_{t'}^\alpha)^{-3} (\sigma_{t'}^\alpha \text{grad}_\alpha \sigma_t^\alpha - \sigma_t^\alpha \text{grad}_\alpha \sigma_{t'}^\alpha) \text{grad}_\alpha^\top \sigma_{t'}^\alpha \right\|_1 \\
&\leq \frac{2C^2(\sigma_{\max} + 1)}{\sigma_{\min}^2} + \frac{4d_3 C^2 \sigma_{\max}}{\sigma_{\min}^3} \leq \frac{8d_3 C^2 \sigma_{\max}}{\sigma_{\min}^3}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\left\| \mathbf{H}_{u,v}^{t,t'}(\phi, \alpha) \right\|_1 &= 2\pi |u| \left\| \text{Hess}_\phi(\rho_t^\phi - \rho_{t'}^\phi) \right\|_1 + |v| \frac{\sigma_{t'}^\alpha \sigma_t^{\alpha_0}}{\sigma_t^\alpha} \left\| \text{Hess}_\alpha(\sigma_t^\alpha / \sigma_{t'}^\alpha) \right\|_1 \\
&\leq 4\pi C^2 |u| + 8d_3 C^2 \frac{\sigma_{\max}^3}{\sigma_{\min}^4} |v| \leq \tilde{C}_2 \|(u, v)\|,
\end{aligned}$$

where $\tilde{C}_2 := 8\pi d_3^{3/2} C^2 \sigma_{\max}^3 / \sigma_{\min}^4$, which implies

$$\begin{aligned} \left\| \text{Hess}_{(\phi, \alpha)} d_{u,v}^{t,t'}(\phi, \alpha) \right\|_1 &\leq \left\| \mathbf{H}_{u,v}^{t,t'}(\phi, \alpha) \right\|_1 + d_3 \left\| \mathbf{a}_{u,v}^{t,t'}(\phi, \alpha) \right\|^2 \\ &\leq \tilde{C}_2 \|(u, v)\| + d_3 \tilde{C}_1^2 \|(u, v)\|^2. \end{aligned}$$

Then, (5.13) holds with $\tilde{C} := \max\{\sqrt{d_3} \tilde{C}_1, \tilde{C}_2\}$. \square

Lemma 5.8 (Properties of $h_\omega(\delta)$). *Recall from (4.5) that $h_\omega(\delta) := e^{2\pi i(\omega, \delta)}$. We have*

$$|h_\omega(\delta)| \equiv 1, \quad (5.14)$$

$$\overline{h_\omega(\delta)} = 1/h_\omega(\delta) = h_\omega(-\delta), \quad (5.15)$$

$$h_\omega(0) = 1, \quad (5.16)$$

$$h_\omega(\delta)h_\omega(\delta') = h_\omega(\delta + \delta'). \quad (5.17)$$

Proof. The fact that $|e^{ix}| = 1$ and $\overline{e^{ix}} = e^{-ix}$ for all $x \in \mathbb{R}$ implies (5.14) and (5.15). The properties (5.16) and (5.17) follow because $e^0 = 1$ and $e^x e^y = e^{x+y}$ for all $x, y \in \mathbb{C}$. \square

Assumption 5.9. *There is a convex open neighborhood $U' \subseteq \Theta$ of θ_0 and some $C > 0$ such that $\theta \mapsto \delta_t^\theta$ is twice differentiable in U' for all $t \in [0, 1]$ and, uniformly in θ and t ,*

$$\left\| \text{grad}_\theta(\delta_t^\theta)_1 \right\|, \left\| \text{grad}_\theta(\delta_t^\theta)_2 \right\| \leq C, \quad \left\| \text{Hess}_\theta(\delta_t^\theta)_1 \right\|_1, \left\| \text{Hess}_\theta(\delta_t^\theta)_2 \right\|_1 \leq C^2.$$

Lemma 5.10 (Derivatives of $h_\omega(\delta_t^\theta)$). *Under Assumption 5.9, for $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$, $t, t' \in [0, 1]$, $\theta \in U'$, and with $\delta_t^\theta - \delta_{t'}^\theta = ((\delta_t^\theta - \delta_{t'}^\theta)_1, (\delta_t^\theta - \delta_{t'}^\theta)_2)$, we define*

$$\tau_T^t := 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\hat{\rho}_t^{\hat{\alpha}_T}} \in \mathbb{R}^{2 \times 2}, \quad (5.18)$$

$$h_{T,\omega}^{t,t'}(\theta) := h_\omega(\tau_T^t(\delta_t^\theta - \delta_{t'}^\theta) - \tau_T^{t'}\delta_{t'}^\theta) \in \mathbb{C}, \quad (5.19)$$

$$\mathbf{b}_{T,\omega}^{t,t'}(\theta) := 2\pi(\omega_1 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_1 + \omega_2 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_2) \in \mathbb{R}^{d_1}, \quad (5.20)$$

$$\mathbf{H}_{T,\omega}^{t,t'}(\theta) := 2\pi(\omega_1 \text{Hess}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_1 + \omega_2 \text{Hess}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_2) \in \mathbb{R}^{d_1 \times d_1}. \quad (5.21)$$

Note, that $h_{T,\omega}^{t,0}(\theta_0) = h_\omega(0) = 1$ and that $\mathbf{b}_{T,\omega}^{t,0}(\theta) - \mathbf{b}_{T,\omega}^{t',0}(\theta) = \mathbf{b}_{T,\omega}^{t,t'}(\theta)$. For later use, we also define the limits (given the consistency of the rotation and scaling parameter estimators which we will prove in Chapter 6)

$$\tau_\infty^t := 1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\alpha_0}} \in \mathbb{R}^{2 \times 2}, \quad (5.22)$$

$$\mathbf{b}_{\infty,\omega}^{t,t'}(\theta) := 2\pi(\omega_1 \text{grad}_\theta(\tau_\infty^t\delta_t^\theta - \tau_\infty^{t'}\delta_{t'}^\theta)_1 + \omega_2 \text{grad}_\theta(\tau_\infty^t\delta_t^\theta - \tau_\infty^{t'}\delta_{t'}^\theta)_2) \in \mathbb{R}^{d_1}, \quad (5.23)$$

$$\mathbf{H}_{\infty,\omega}^{t,t'}(\theta) := 2\pi(\omega_1 \text{Hess}_\theta(\tau_\infty^t\delta_t^\theta - \tau_\infty^{t'}\delta_{t'}^\theta)_1 + \omega_2 \text{Hess}_\theta(\tau_\infty^t\delta_t^\theta - \tau_\infty^{t'}\delta_{t'}^\theta)_2) \in \mathbb{R}^{d_1 \times d_1}. \quad (5.24)$$

There is a constant $\tilde{C} > 0$ (independent from T , ω , t , t' , and θ) such that

$$\text{grad}_\theta h_{T,\omega}^{t,t'}(\theta) = \mathbf{i} \mathbf{b}_{T,\omega}^{t,t'}(\theta) h_{T,\omega}^{t,t'}(\theta), \quad (5.25)$$

$$\text{Hess}_\theta h_{T,\omega}^{t,t'}(\theta) = (\mathbf{i} \mathbf{H}_{T,\omega}^{t,t'}(\theta) - \mathbf{b}_{T,\omega}^{t,t'}(\theta) \mathbf{b}_{T,\omega}^{t,t'}(\theta)^\top) h_{T,\omega}^{t,t'}(\theta), \quad (5.26)$$

$$\left\| \mathbf{b}_{T,\omega}^{t,t'}(\theta) \right\| \leq \tilde{C} \|\omega\|, \quad (5.27)$$

$$\left\| \text{grad}_\theta h_{T,\omega}^{t,t'}(\theta) \right\| \leq \tilde{C} \|\omega\|, \quad (5.28)$$

$$\left\| \mathbf{H}_{T,\omega}^{t,t'}(\theta) \right\|_1 \leq \tilde{C} \|\omega\|, \quad (5.29)$$

$$\left\| \text{Hess}_\theta h_{T,\omega}^{t,t'}(\theta) \right\|_1 \leq \tilde{C} \|\omega\| + \tilde{C}^2 \|\omega\|^2. \quad (5.30)$$

Proof. First, consider

$$\begin{aligned}
& \text{grad}_\theta h_\omega(\tau_T^t(\delta_t^\theta - \delta_t^{\theta_0}) - \tau_T^{t'}\delta_{t'}^\theta) \\
&= \text{grad}_\theta \exp\left(2\pi i \left\langle \omega, \tau_T^t(\delta_t^\theta - \delta_t^{\theta_0}) - \tau_T^{t'}\delta_{t'}^\theta \right\rangle\right) \\
&= 2\pi i(\omega_1 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_1 + \omega_2 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_2)h_\omega(\tau_T^t(\delta_t^\theta - \delta_t^{\theta_0}) - \tau_T^{t'}\delta_{t'}^\theta),
\end{aligned}$$

proving (5.25). Hence,

$$\begin{aligned}
& \text{Hess}_\theta h_\omega(\tau_T^t(\delta_t^\theta - \delta_t^{\theta_0}) - \tau_T^{t'}\delta_{t'}^\theta) \\
&= \text{grad}_\theta^\top \left(2\pi i(\omega_1 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_1 + \omega_2 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_2)\right. \\
&\quad \left.\cdot h_\omega(\tau_T^t(\delta_t^\theta - \delta_t^{\theta_0}) - \tau_T^{t'}\delta_{t'}^\theta)\right) \\
&= 2\pi h_\omega(\tau_T^t(\delta_t^\theta - \delta_t^{\theta_0}) - \tau_T^{t'}\delta_{t'}^\theta) \left[i(\omega_1 \text{Hess}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_1 + \omega_2 \text{Hess}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_2)\right. \\
&\quad \left.- 2\pi(\omega_1 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_1 + \omega_2 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_2)\right. \\
&\quad \left.\cdot (\omega_1 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_1 + \omega_2 \text{grad}_\theta(\tau_T^t\delta_t^\theta - \tau_T^{t'}\delta_{t'}^\theta)_2)^\top \right],
\end{aligned}$$

which proves (5.26). We denote the components of τ_T^t by $(\tau_T^t)_{i,j}$, $i, j \in \{1, 2\}$, for which we have $|(\tau_T^t)_{i,j}| \leq \sigma_{\min}^{-1}$. Let $\tilde{C}_1 := 8\pi\sqrt{d_1}C/\sigma_{\min}$ with $C > 0$ from Assumption 5.9. Then, with Lemma 2.7,

$$\begin{aligned}
\left\| \mathbf{b}_{T,\omega}^{t,t'}(\theta) \right\| &\leq 2\pi |\omega_1| \left(\left\| \text{grad}_\theta(\tau_T^t\delta_t^\theta)_1 \right\| + \left\| \text{grad}_\theta(\tau_T^{t'}\delta_{t'}^\theta)_1 \right\| \right) \\
&\quad + 2\pi |\omega_2| \left(\left\| \text{grad}_\theta(\tau_T^t\delta_t^\theta)_2 \right\| + \left\| \text{grad}_\theta(\tau_T^{t'}\delta_{t'}^\theta)_2 \right\| \right) \\
&\leq 2\pi |\omega_1| \left(|(\tau_T^t)_{1,1}| \left\| \text{grad}_\theta(\delta_t^\theta)_1 \right\| + |(\tau_T^t)_{1,2}| \left\| \text{grad}_\theta(\delta_t^\theta)_2 \right\| \right) \\
&\quad + \left| (\tau_T^{t'})_{1,1} \right| \left\| \text{grad}_\theta(\delta_{t'}^\theta)_1 \right\| + \left| (\tau_T^{t'})_{1,2} \right| \left\| \text{grad}_\theta(\delta_{t'}^\theta)_2 \right\| \\
&\quad + 2\pi |\omega_2| \left(|(\tau_T^t)_{2,1}| \left\| \text{grad}_\theta(\delta_t^\theta)_1 \right\| + |(\tau_T^t)_{2,2}| \left\| \text{grad}_\theta(\delta_t^\theta)_2 \right\| \right) \\
&\quad + \left| (\tau_T^{t'})_{2,1} \right| \left\| \text{grad}_\theta(\delta_{t'}^\theta)_1 \right\| + \left| (\tau_T^{t'})_{2,2} \right| \left\| \text{grad}_\theta(\delta_{t'}^\theta)_2 \right\| \\
&\leq \tilde{C}_1 \|\omega\|, \tag{5.31}
\end{aligned}$$

from which we get (5.27). By (5.25), we have

$$\left\| \text{grad}_\theta h_{T,\omega}^{t,t'}(\theta) \right\| = \left\| \mathbf{b}_{T,\omega}^{t,t'}(\theta) \right\|,$$

implying (5.28). Similarly to (5.31), we get that $\left\| \mathbf{H}_{T,\omega}^{t,t'}(\theta) \right\|_1 \leq \tilde{C}_1 \|\omega\|$, which yields (5.29). With $\tilde{C} := \sqrt{d_1}\tilde{C}_1$, it follows that

$$\left\| \text{Hess}_\theta h_\omega(\delta_t^\theta - \delta_t^{\theta_0}) \right\|_1 \leq \left\| \mathbf{H}_{T,\omega}^{t,t'}(\theta) \right\|_1 + d_1 \left\| \mathbf{b}_{T,\omega}^{t,t'}(\theta) \right\|^2 \leq \tilde{C} \|\omega\| + \tilde{C}^2 \|\omega\|^2.$$

Thus, (5.30) holds, completing the proof. \square

Chapter 6

Main Results

In this chapter, we prove the consistency of $(\hat{\phi}_T, \hat{\alpha}_T)$, $\hat{\theta}_T$, and \hat{f}_T , as well as the asymptotic normality of $\sqrt{T}(\hat{\phi}_T, \hat{\alpha}_T)$ and the uniform tightness of $(\hat{\theta}_T)_{T \in \mathbb{N}}$.

6.1 Estimation of rotation and scaling

Before we start, we need another definition (see e.g., Rudin (1990, page 265)).

Definition 6.1 (Total variation). *Let $g: [0, 1] \rightarrow \mathbb{C}$ and let*

$$P := \{\{t_0, \dots, t_k\} \mid k \in \mathbb{N}, 0 = t_0 < t_1 < \dots < t_k = 1\}$$

the set of all finite partitions of $[0, 1]$. We define the total variation of g as

$$\text{TV}(g) := \sup_{\{t_0, \dots, t_k\} \in P} \sum_{i=0}^{k-1} |g(t_{i+1}) - g(t_i)|.$$

Assumption 6.2. *Under Assumption 5.6, we assume that the second partial derivatives*

$$\phi \mapsto \frac{\partial^2 \rho_t^\phi}{\partial \phi_m \partial \phi_{m'}}, \quad \alpha \mapsto \frac{\partial^2 \sigma_t^\alpha}{\partial \alpha_{m''} \partial \alpha_{m'''}}$$

are continuous at the true parameters ϕ_0 and α_0 , respectively, for all $m, m' \in \{1, \dots, d_2\}$ and $m'', m''' \in \{1, \dots, d_3\}$. Furthermore, the first partial derivatives at ϕ_0 and α_0 , as functions in t , are of bounded total variation, that is, there is a $C' > 0$ such that

$$\text{TV} \left(t \mapsto \frac{\partial \rho_t^\phi}{\partial \phi_m} \right) < C', \quad \text{TV} \left(t \mapsto \frac{\partial \sigma_t^\alpha}{\partial \alpha_{m''}} \right) < C',$$

for all $m \in \{1, \dots, d_2\}$ and $m'' \in \{1, \dots, d_3\}$.

Definition 6.3 (Not rotation or scaling invariant). *A function $g: \mathbb{R}^2 \rightarrow \mathbb{C}$ is called not rotation invariant if there is no $\rho \in (0, 2\pi)$ such that $g(x) = g(R_{-\rho}x)$ for all $x \in \mathbb{R}^2$. Similarly, g is called not scaling invariant if there is no $\sigma \in (0, 1) \cup (1, \infty)$ such that $g(x) = g(1/\sigma \cdot x)$ for all $x \in \mathbb{R}^2$.*

Lemma 6.4. *Let $g \in L^2(\mathbb{R}^2)$ and $\gamma > 0$ such that $(\omega \mapsto \|\omega\|^\gamma |\mathcal{F}_g(\omega)|^2) \in L^1(\mathbb{R}^2)$. The function g is not rotation invariant if and only if there are coprime $u, u' \in \mathbb{Z} \setminus \{0\}$ and $v \in \mathbb{R}$ such that $\mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) \neq 0$ and $\mathcal{AFM}_{|\mathcal{F}_g|^2}(u', v) \neq 0$.*

Proof. Let $\rho \in [0, 2\pi)$ such that $g(x) = g(R_{-\rho}x)$ for all $x \in \mathbb{R}^2$. From Lemma 2.5, we get that

$$|\mathcal{F}_g(\omega)|^2 = |\mathcal{F}_g(R_{-\rho}\omega)|^2 \quad \text{for all } \omega \in \mathbb{R}^2.$$

Hence, by Lemma 2.18,

$$\mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) = e^{-2\pi i u \rho} \mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) \quad \text{for all } (u, v) \in \mathbb{Z} \times \mathbb{R}. \quad (6.1)$$

Let $u, u' \in \mathbb{Z} \setminus \{0\}$ coprime with $\mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) \neq 0$ and $\mathcal{AFM}_{|\mathcal{F}_g|^2}(u', v) \neq 0$. Then, (6.1) implies that $e^{-2\pi i u \rho} = 1$ and $e^{-2\pi i u' \rho} = 1$ and thus, $u\rho \in \mathbb{Z}$ and $u'\rho \in \mathbb{Z}$. In particular, $\rho \in [0, 2\pi) \cap \mathbb{Q}$. Let $\rho = z_1/z_2$, where $z_1 \in \mathbb{Z}$ and $z_2 \in \mathbb{N}$ are coprime.

Assume that $\rho \neq 0$. Then, $z_1 \neq 0$ and it follows that both uz_1 and $u'z_1$ are multiples of z_2 , which is a contradiction to the fact that u and u' are coprime.

Now, assume that $\mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) = 0$ for all coprime $u, u' \in \mathbb{Z} \setminus \{0\}$ and $v \in \mathbb{R}$. Since for every $u \in \mathbb{Z} \setminus \{0\}$ there is a $u' \in \mathbb{Z} \setminus \{0\}$ such that u and u' are coprime (e.g., $u' := pu$ with a prime number p that is not part of the prime decomposition of u), this means that $\mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) = 0$ for all $u \neq 0$. With the inverse analytical Fourier-Mellin transform, we get that

$$r^\gamma (|\mathcal{F}_g|^2 \circ \mathcal{P})(r, \psi) = \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} e^{2\pi i u \psi} r^{iv} \mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) dv = 0$$

for all $(r, \psi) \in [0, \infty) \times [0, 2\pi)$, implying that $\mathcal{F}_g = 0$. Hence,

$$g(x) = \int_{\mathbb{R}^2} e^{2\pi i \langle x, \omega \rangle} \mathcal{F}_g(\omega) d\omega = 0 \quad \text{for all } x \in \mathbb{R}^2,$$

which means that g is constant and, in particular, $g(x) = g(R_{-\rho}x)$ for all $x \in \mathbb{R}^2$ and $\rho \in [0, 2\pi)$ \square

Lemma 6.5. *Let $g \in L^2(\mathbb{R}^2)$ and $\gamma > 0$ such that $(\omega \mapsto \|\omega\|^\gamma |\mathcal{F}_g(\omega)|^2) \in L^1(\mathbb{R}^2)$. The function g is not scaling invariant if and only if there are $u \in \mathbb{Z}$ and an open Borel set $B \subseteq \mathbb{R}$ with positive Lebesgue measure such that $\mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) \neq 0$ for all $v \in B$.*

Proof. Let $\sigma \in (0, \infty)$ such that $g(x) = g(1/\sigma \cdot x)$ for all $x \in \mathbb{R}^2$. Because of Lemma 2.5, this implies that

$$|\mathcal{F}_g(\omega)|^2 = \sigma^4 |\mathcal{F}_g(\sigma\omega)|^2 \quad \text{for all } \omega \in \mathbb{R}^2.$$

Hence, by Lemma 2.18,

$$\mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) = \sigma^{4-\gamma+iv} \mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) \quad \text{for all } (u, v) \in \mathbb{Z} \times \mathbb{R}.$$

For $\mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) \neq 0$, this implies that $\sigma^{4-\gamma+iv} = 1$. In the case of $\gamma \neq 4$, taking the absolute value yields $\sigma = 1$. For $\gamma = 4$, it follows that

$$e^{iv \log(\sigma)} = 1. \quad (6.2)$$

Since g is not scaling invariant, this holds for all v in an open Borel set $B \subseteq \mathbb{R}$ with positive Lebesgue-measure. Because both \mathbb{Q} and $\mathbb{R} \setminus \mathbb{Q}$ are dense in \mathbb{R} , we can therefore fix $v \in B$ and choose $v_1, v_2 \in B \setminus \{v\}$ such that

$$v - v_1 \in \mathbb{Q}, \quad \text{while } v - v_2 \in \mathbb{R} \setminus \mathbb{Q}. \quad (6.3)$$

By (6.2), we have that

$$e^{i(v-v_1)\log(\sigma)} = \frac{e^{iv\log(\sigma)}}{e^{iv_1\log(\sigma)}} = 1,$$

which implies $(v - v_1)\log(\sigma)/2\pi \in \mathbb{Z}$. Similarly, we get $(v - v_2)\log(\sigma)/2\pi \in \mathbb{Z}$. Because of (6.3), this implies $\log(\sigma) = 0$ and thus, $\sigma = 1$ for all possible values of γ .

On the other hand, if there are no $u \in \mathbb{Z}$ and no open Borel sets $B \subseteq \mathbb{R}$ with positive Lebesgue-measure such that $\mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) \neq 0$ for all $v \in B$, then $\mathcal{AFM}_{|\mathcal{F}_g|^2} = 0$ a.e. With the inverse analytical Fourier-Mellin transform, it follows that

$$r^\gamma (|\mathcal{F}_g|^2 \circ \mathcal{P})(r, \psi) = \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} e^{2\pi i u \psi} r^{iv} \mathcal{AFM}_{|\mathcal{F}_g|^2}(u, v) dv = 0$$

for all $(r, \psi) \in [0, \infty) \times [0, 2\pi)$. Hence, $\mathcal{F}_g = 0$, and thus

$$g(x) = \int_{\mathbb{R}^2} e^{2\pi i \langle x, \omega \rangle} \mathcal{F}_g(\omega) d\omega = 0 \quad \text{for all } x \in \mathbb{R}^2,$$

implying that g is constant and, in particular, $g(x) = g(1/\sigma \cdot x)$ for all $x \in \mathbb{R}^2$ and $\sigma > 0$. \square

Assumption 6.6. *Let $f \in L^2(\mathbb{R}^2) \cap H^{1/2}(\mathbb{R}^2)$ not rotation or scaling invariant.*

Assumption 6.7. *The maps*

$$\begin{aligned} \Phi &\rightarrow L^1([0, 1], (-\pi/2, \pi/2]), & \phi &\mapsto \left(\rho_t^\phi : t \mapsto \rho_t^\phi \right), \\ \mathbb{A} &\rightarrow L^1([0, 1], [\sigma_{\min}, \sigma_{\max}]), & \alpha &\mapsto \left(\sigma_t^\alpha : t \mapsto \sigma_t^\alpha \right), \end{aligned}$$

are continuous with respect to the L^1 -norm. Moreover, for each $\phi \in \Phi$ and $\alpha \in \mathbb{A}$, the rotation and scaling functions $t \mapsto \rho_t^\phi$ and $t \mapsto \sigma_t^\alpha$ are continuous at $t = 0$.

Definition 6.8 (Identifiability). *For some index set I , let $G_I = \{g^i : [0, 1] \rightarrow \mathbb{R} \mid i \in I\}$ a set of functions. We call G_I identifiable if for all $i, j \in I$, the existence of a Borel set $B \subseteq [0, 1]$ of Lebesgue measure equal to 1 with $g^i(t) = g^j(t)$ for all $t \in B$ implies $i = j$.*

Assumption 6.9. *The set of scaling functions $\{t \mapsto \sigma_t^\alpha \mid \alpha \in \mathbb{A}\}$ and the set of rotation functions $\{t \mapsto \rho_t^\phi \mid \phi \in \Phi\}$ are identifiable.*

Assumption 6.10. *There is an open neighborhood $U_{\rho, \sigma} \subseteq \Phi \times \mathbb{A}$ of (ϕ_0, α_0) and there are constants $L_\rho, L_\sigma > 0$ such that the following local uniform Lipschitz conditions hold,*

$$\sup_{t \in [0, 1]} \left| \rho_t^\phi - \rho_t^{\phi_0} \right| \leq L_\rho \|\phi - \phi_0\|, \quad \sup_{t \in [0, 1]} |\sigma_t^\alpha - \sigma_t^{\alpha_0}| \leq L_\sigma \|\alpha - \alpha_0\| \quad \text{for all } (\phi, \alpha) \in U_{\rho, \sigma}.$$

Assumption 6.11. *There is a $C > 0$ such that, uniformly in ϕ and α ,*

$$\text{TV}(t \mapsto \rho_t^\phi) + \text{TV}(t \mapsto \sigma_t^\alpha) < C.$$

Assumption 6.12. *We have the following Sobolev-1/2 condition,*

$$\int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \|(u, v)\| \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv < \infty.$$

Theorem 6.13 (Consistency of rotation and scaling parameters). *Suppose that the Assumptions 5.2, 6.6, 6.7, 6.9, 6.10, 6.11, and 6.12 hold. If we have $r_T, u_T, v_T, \beta_T \xrightarrow{T \rightarrow \infty} \infty$, $\beta_T = o(T)$, and $r_T^{2\gamma} u_T v_T \beta_T^{-1} \xrightarrow{T \rightarrow \infty} 0$, then the scaling and rotation estimator $(\hat{\phi}_T, \hat{\alpha}_T)$ from Definition 4.3 is consistent, that is,*

$$(\hat{\phi}_T, \hat{\alpha}_T) \xrightarrow{T \rightarrow \infty} (\phi_0, \alpha_0) \text{ in probability.} \quad (6.4)$$

Proof. The proof follows a standard three step argument in M-estimation (e.g., (van der Vaart, 2000; Gamboa et al., 2007; Bigot et al., 2009; Hartmann et al., 2016)). The three steps are:

1. Show the uniqueness of the population contrast minimizer (ϕ_0, α_0) which, together with the compactness of $\Phi \times \mathbb{A}$, yields (B.2) in the Appendix.
2. Show the continuity of the population contrast functional.
3. Show that $\widetilde{M}_T \xrightarrow{T \rightarrow \infty} \widetilde{M}$ in probability uniformly over (ϕ, α) , implying (B.1) in the Appendix.

Note, that (B.3) in the Appendix holds because $(\hat{\phi}_T, \hat{\alpha}_T)$ is defined as a minimizer of \widetilde{M}_T . Then, Theorem B.6 in the Appendix yields the desired consistency of $(\hat{\phi}_T, \hat{\alpha}_T)$.

Step 1: uniqueness of the contrast minimizer (ϕ_0, α_0) . First, note that because $F^t(u, v) = (\sigma_t^{\alpha_0})^{4-\gamma} \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v)$, we have

$$\begin{aligned} \widetilde{M}(\phi, \alpha) &= - \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \int_0^1 d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) (\sigma_t^{\alpha_0})^{4-\gamma} \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) dt \right|^2 dv \\ &= - \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left| \int_0^1 d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) (\sigma_t^{\alpha_0})^{4-\gamma} dt \right|^2 dv \\ &\geq - \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left(\int_0^1 \left| d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) (\sigma_t^{\alpha_0})^{4-\gamma} \right| dt \right)^2 dv \\ &= - \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left(\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} dt \right)^2 dv \end{aligned}$$

for all (ϕ, α) with equality if $(\phi, \alpha) = (\phi_0, \alpha_0)$. Let $(\phi, \alpha) \in \Phi \times \mathbb{A}$ such that

$$\widetilde{M}(\phi, \alpha) = - \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left(\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} dt \right)^2 dv. \quad (6.5)$$

Since f is not rotation or scaling invariant by Assumption 6.6, by Lemma 6.4, there is are coprime $u, u' \in \mathbb{Z} \setminus \{0\}$ and an open Borel set $B \subseteq \mathbb{R}$ with positive Lebesgue-measure such that $\mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \neq 0$ and $\mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u', v) \neq 0$ for all $v \in B$. Then, for (6.5) to hold, we must have

$$\left| \int_0^1 d_{u'',v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) (\sigma_t^{\alpha_0})^{4-\gamma} dt \right| = \int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} dt$$

for all $(u'', v) \in \{u, u'\} \times B$. By Lemma A.11 in the Appendix, $t \mapsto d_{u'',v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0})$ is constant a.e. on $[0, 1]$. Because of the identifiability constraint $(\sigma_0^\alpha, \rho_0^\phi) = (1, 0)$ for all

(ϕ, α) and the continuity of σ^α and ρ^ϕ at $t = 0$ (Assumption 6.7), this constant has to be 1 which implies that

$$1 = d_{u'',v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) = \exp\left(-iv(\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0})) + 2\pi i u''(\rho_t^\phi - \rho_t^{\phi_0})\right) \quad (6.6)$$

a.e. on $[0, 1]$ for all $(u'', v) \in \{u, u'\} \times B$. Choose $v_1, v_2 \in B \setminus \{v\}$ such that

$$v - v_1 \in \mathbb{Q}, \quad v - v_2 \in \mathbb{R} \setminus \mathbb{Q}. \quad (6.7)$$

From (6.6),

$$\exp\left(-i(v - v_1)(\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0}))\right) = \frac{d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0})}{d_{u,v_1}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0})} = 1,$$

and similarly, $\exp\left(-i(v - v_2)(\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0}))\right) = 1$, which implies that

$$\begin{aligned} (v - v_1)(\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0})) / (2\pi) &\in \mathbb{Z}, \\ (v - v_2)(\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0})) / (2\pi) &\in \mathbb{Z}. \end{aligned}$$

Because of (6.7), this means that $\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0}) = 0$ a.e. on $[0, 1]$. Since the logarithm is bijective and the scaling functions are identifiable (Assumption 6.9), we get $\alpha = \alpha_0$. Because of (6.6), it follows that

$$\exp\left(2\pi i u(\rho_t^\phi - \rho_t^{\phi_0})\right) = 1, \quad \exp\left(2\pi i u'(\rho_t^\phi - \rho_t^{\phi_0})\right) = 1, \quad \text{a.e. on } [0, 1],$$

which means that $u(\rho_t^\phi - \rho_t^{\phi_0}) \in \mathbb{Z}$ and $u'(\rho_t^\phi - \rho_t^{\phi_0}) \in \mathbb{Z}$ a.e. on $[0, 1]$. In particular, $\rho_t^\phi - \rho_t^{\phi_0} \in \mathbb{Q}$. Since $t \mapsto \rho_t^\phi - \rho_t^{\phi_0}$ is continuous at $t = 0$ with value 0 (Assumption 6.7), we can choose $t > 0$ small such that $|\rho_t^\phi - \rho_t^{\phi_0}| < 1$. Let $z_1 \in \mathbb{Z}$ and $z_2 \in \mathbb{N}_{>1}$ coprime such that $\rho_t^\phi - \rho_t^{\phi_0} = z_1/z_2$. Then, uz_1 and $u'z_1$ are both multiples of z_2 , which contradicts the fact that u and u' are coprime unless $z_1 = 0$. Thus, $\rho_t^\phi = \rho_t^{\phi_0}$ and the identifiability of the rotation functions (Assumption 6.9) yields $\phi = \phi_0$.

Step 2: Continuity of \widetilde{M} . By Assumption 6.7, the functions $t \mapsto \sigma_t^\alpha$ and $t \mapsto \rho_t^\alpha$ are measurable for all $\alpha \in \mathbb{A}$, $\phi \in \Phi$. Therefore, the functions

$$t \mapsto d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) = d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) (\sigma_t^{\alpha_0})^{4-\gamma} \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v)$$

are measurable, too, for all $u \in \mathbb{Z}$, $v \in \mathbb{R}$, $\alpha \in \mathbb{A}$, $\phi \in \Phi$, as they are concatenations of measurable functions. By the same Assumption, the functions $\alpha \mapsto \sigma_t^\alpha$ and $\phi \mapsto \rho_t^\phi$ are continuous for all $t \in [0, 1]$. As a concatenation of continuous functions, $(\phi, \alpha) \mapsto d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v)$ is also continuous for all $u \in \mathbb{Z}$, $v \in \mathbb{R}$, $t \in [0, 1]$. Furthermore, the constant function

$$t \mapsto \tilde{g}_{u,v} := \max\{\sigma_{\max}^{4-\gamma}, \sigma_{\min}^{4-\gamma}\} \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v)$$

is an integrable majorant for $t \mapsto d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v)$. Thus, we can apply Theorem B.1 in the Appendix to get that

$$(\phi, \alpha) \mapsto \int_0^1 d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) dt$$

is continuous for all $u \in \mathbb{Z}$, $v \in \mathbb{R}$. Because $x \mapsto |x|^2$ is continuous, so is

$$(\phi, \alpha) \mapsto g_{u,v}^{\alpha,\phi} := \left| \int_0^1 d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) dt \right|^2.$$

By Assumption 6.6 and Lemma A.13 in the Appendix, $\mathcal{AFM}_{|\mathcal{F}_f|^2}: \mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{C}$ is continuous. Since $(u, v) \mapsto d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0})$ is continuous, too, we get at once that $(u, v) \mapsto d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v)$ is continuous. Hence, by the same chain of arguments as above, the function $(u, v) \mapsto g_{u,v}^{\alpha,\phi}$ is continuous and therefore Lebesgue-measurable for all $\alpha \in \mathbb{A}$, $\phi \in \Phi$. Since $g_{u,v}^{\alpha,\phi} \leq |\tilde{g}_{u,v}|^2$ for all $\alpha \in \mathbb{A}$, $\phi \in \Phi$, and, by Assumption 6.12,

$$\int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} |\tilde{g}_{u,v}|^2 dv = \max \{ \sigma_{\max}^{8-2\gamma}, \sigma_{\min}^{8-2\gamma} \} \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv < \infty,$$

$(u, v) \mapsto |\tilde{g}_{u,v}|^2$ is an integrable majorant for $(u, v) \mapsto g_{u,v}^t(\phi, \alpha)$. Applying Theorem B.1 in the Appendix again yields the continuity of \widetilde{M} .

Step 3: $\widetilde{M}_T \xrightarrow{T \rightarrow \infty} \widetilde{M}$ in probability uniformly in (ϕ, α) . From (3.11) and (4.1), we get

$$\mathcal{AFM}_{|Y^t|^2}(u, v) = d_{u,v}(1/\sigma_t^{\alpha_0}, -\rho_t^{\phi_0}) F^t(u, v) + \mathcal{AFM}_{\mathcal{W}^t}(u, v). \quad (6.8)$$

Hence,

$$\begin{aligned} \widetilde{M}_T(\phi, \alpha) &= - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{AFM}_{|Y^t|^2}(u, v) \right|^2 dv \\ &= - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) \right. \\ &\quad \left. + d_{u,v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{AFM}_{\mathcal{W}^t}(u, v) \right|^2 dv \\ &= A_T(\phi, \alpha) + B_T(\phi, \alpha) + C_T(\phi, \alpha), \end{aligned} \quad (6.9)$$

with

$$\begin{aligned} A_T(\phi, \alpha) &:= - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) \right|^2 dv, \\ B_T(\phi, \alpha) &:= - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} 2\Re \left[\left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) \right) \right. \\ &\quad \left. \cdot \left(\frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} \overline{d_{u,v}(\sigma_{t'}^\alpha, \rho_{t'}^\phi) \mathcal{AFM}_{\mathcal{W}^{t'}}(u, v)} \right) \right] dv, \\ C_T(\phi, \alpha) &:= - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{AFM}_{\mathcal{W}^t}(u, v) \right|^2 dv, \end{aligned}$$

where we used Lemma A.1 in the Appendix. To derive the desired uniform convergence we will show that the deterministic part $A_T \xrightarrow{T \rightarrow \infty} \widetilde{M}$ uniformly in (ϕ, α) while the random part $B_T + C_T \xrightarrow{T \rightarrow \infty} 0$ in probability uniformly in (ϕ, α) . To keep the notation short, in the following, we will write

$$c_{u,v}^t(\phi, \alpha) := d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0})(\sigma_t^{\alpha_0})^{4-\gamma}. \quad (6.10)$$

Consider

$$\begin{aligned} & \left| A_T(\phi, \alpha) - \widetilde{M}(\phi, \alpha) \right| \\ &= \left| \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v} \left(\frac{\sigma_t^\alpha}{\sigma_t^{\alpha_0}}, \rho_t^\phi - \rho_t^{\phi_0} \right) F^t(u, v) \right|^2 \right. \\ & \quad \left. - \left| \int_0^1 d_{u,v} \left(\frac{\sigma_t^\alpha}{\sigma_t^{\alpha_0}}, \rho_t^\phi - \rho_t^{\phi_0} \right) F^t(u, v) dt \right|^2 dv \right. \\ & \quad \left. - \left(\int_{-\infty}^{-v_T} \sum_{|u| \leq u_T} + \int_{v_T}^{\infty} \sum_{|u| \leq u_T} + \int_{\mathbb{R}} \sum_{|u| > u_T} \right) \left| \int_0^1 d_{u,v} \left(\frac{\sigma_t^\alpha}{\sigma_t^{\alpha_0}}, \rho_t^\phi - \rho_t^{\phi_0} \right) F^t(u, v) dt \right|^2 dv \right| \\ &= \left| \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left(\left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} c_{u,v}^t(\phi, \alpha) \right|^2 - \left| \int_0^1 c_{u,v}^t(\phi, \alpha) dt \right|^2 \right) dv \right. \\ & \quad \left. - \left(\int_{-\infty}^{-v_T} \sum_{|u| \leq u_T} + \int_{v_T}^{\infty} \sum_{|u| \leq u_T} + \int_{\mathbb{R}} \sum_{|u| > u_T} \right) \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left| \int_0^1 c_{u,v}^t(\phi, \alpha) dt \right|^2 dv \right| \\ &\leq \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left| \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} c_{u,v}^t(\phi, \alpha) \right|^2 - \left| \int_0^1 c_{u,v}^t(\phi, \alpha) dt \right|^2 \right| dv \\ & \quad + \left(\int_{-\infty}^{-v_T} \sum_{|u| \leq u_T} + \int_{v_T}^{\infty} \sum_{|u| \leq u_T} + \int_{\mathbb{R}} \sum_{|u| > u_T} \right) \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \int_0^1 |c_{u,v}^t(\phi, \alpha)|^2 dt dv \\ &\leq 2C_\gamma \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} c_{u,v}^t(\phi, \alpha) - \int_0^1 c_{u,v}^t(\phi, \alpha) dt \right| dv \\ & \quad + C_\gamma \left(\int_{-\infty}^{-v_T} \sum_{|u| \leq u_T} + \int_{v_T}^{\infty} \sum_{|u| \leq u_T} + \int_{\mathbb{R}} \sum_{|u| > u_T} \right) \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv \\ &\leq \frac{2C_\gamma \beta_T}{T} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \text{TV}(t \mapsto c_{u,v}^t(\phi, \alpha)) \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv \\ & \quad + C_\gamma \left(\int_{-\infty}^{-v_T} \sum_{|u| \leq u_T} + \int_{v_T}^{\infty} \sum_{|u| \leq u_T} + \int_{\mathbb{R}} \sum_{|u| > u_T} \right) \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv, \quad (6.11) \end{aligned}$$

where we used Lemma A.4 in the Appendix for the second inequality with

$$C_\gamma := \max \{ \sigma_{\min}^{-8+2\gamma}, \sigma_{\max}^{8-2\gamma} \} \quad (6.12)$$

and Lemma A.6 in the Appendix for the last inequality. Because

$$\int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv < \infty \quad (6.13)$$

by Assumption 6.12, and because $u_T, v_T \xrightarrow{T \rightarrow \infty} \infty$, the second term in (6.11) vanishes for $T \rightarrow \infty$. To tackle the first term, we show that

$$\sup_{(\phi, \alpha) \in \Phi \times \mathbf{A}} \text{TV}(t \mapsto c_{u,v}^t(\phi, \alpha)) \leq C_4 \|(u, v)\| + C_3, \quad (6.14)$$

with some constants $C_3, C_4 > 0$. First of all, $\sigma^\alpha: t \mapsto \sigma_t^\alpha$ and $\rho^\phi: t \mapsto \rho_t^\phi$ are of bounded variation uniformly in $\alpha \in \mathbf{A}$, $\phi \in \Phi$, by Assumption 6.11. By Lemma A.10 in the Appendix, $x \mapsto x^{4-\gamma}$ is Lipschitz-continuous on $[\sigma_{\min}, \sigma_{\max}]$ with Lipschitz-constant in the set $\{(4-\gamma)\sigma_{\max}^{3-\gamma}, (4-\gamma)\sigma_{\min}^{-3+\gamma}, 0, (-4+\gamma)\sigma_{\min}^{-3+\gamma}\}$ (for $\gamma < 3$, $3 < \gamma < 4$, $\gamma = 4$, and $4 < \gamma$, respectively). Furthermore, the logarithm, restricted to the interval $[\sigma_{\min}, \sigma_{\max}]$, is differentiable with derivative bounded by $\log'(\sigma_{\min}) = 1/\sigma_{\min}$. Hence, by Lemma A.10 in the Appendix, $\log_{|[\sigma_{\min}, \sigma_{\max}]}$ is Lipschitz-continuous with Lipschitz-constant $1/\sigma_{\min}$. Since $t \mapsto \sigma_t^\alpha$ is of bounded variation uniformly in $\alpha \in \mathbf{A}$, so is $t \mapsto \log(\sigma_t^\alpha)$. This and Lemma A.9 in the Appendix imply that there are $C_1, C_2, C_3 > 0$ such that

$$\text{TV}(\log(\sigma^\alpha) - \log(\sigma^{\alpha_0})) \leq C_1, \quad \text{TV}(\rho^\phi - \rho^{\phi_0}) \leq C_2, \quad \text{TV}((\sigma^{\alpha_0})^{4-\gamma}) \leq C_3, \quad (6.15)$$

uniformly in (ϕ, α) . Since $\|x\| \leq \|x\|_1$ for all $x \in \mathbb{R}^2$, we have for all $t, t' \in [0, 1]$, that

$$\begin{aligned} & \left| \exp\left(i \left\langle (-v, 2\pi u), (\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_t^{\phi_0}) \right\rangle\right) \right. \\ & \quad \left. - \exp\left(i \left\langle (-v, 2\pi u), (\log(\sigma_{t'}^\alpha) - \log(\sigma_{t'}^{\alpha_0}), \rho_{t'}^\phi - \rho_{t'}^{\phi_0}) \right\rangle\right) \right| \\ & \leq \sqrt{2} \|(2\pi u, -v)\| \left\| (\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_t^{\phi_0}) - (\log(\sigma_{t'}^\alpha) - \log(\sigma_{t'}^{\alpha_0}), \rho_{t'}^\phi - \rho_{t'}^{\phi_0}) \right\|_1 \\ & \leq 2\pi\sqrt{2} \|(u, v)\| \left(\left| (\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0})) - (\log(\sigma_{t'}^\alpha) - \log(\sigma_{t'}^{\alpha_0})) \right| \right. \\ & \quad \left. + \left| (\rho_t^\phi - \rho_t^{\phi_0}) - (\rho_{t'}^\phi - \rho_{t'}^{\phi_0}) \right| \right), \end{aligned}$$

where we applied Lemma A.8 in the Appendix for the first inequality. Hence,

$$\begin{aligned} & \text{TV}\left(t \mapsto d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0})\right) \\ & = \text{TV}\left(t \mapsto \exp\left(i \left\langle (-v, 2\pi u), (\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_t^{\phi_0}) \right\rangle\right)\right) \\ & \leq 2\pi\sqrt{2}(C_1^2 + C_2^2) \|(u, v)\|. \end{aligned}$$

Now, the second part of Lemma A.5 in the Appendix yields

$$\begin{aligned} \text{TV}(t \mapsto c_{u,v}^t(\phi, \alpha)) & \leq \left\| t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma} \right\|_\infty \text{TV}(t \mapsto d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0})) \\ & \quad + \left\| t \mapsto d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) \right\|_\infty \text{TV}(t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma}) \\ & \leq C_4 \|(u, v)\| + C_3, \end{aligned}$$

uniformly in (ϕ, α) , where $C_4 := 2\pi\sqrt{2}C_\gamma(C_1^2 + C_2^2)$ with C_γ from (6.12), proving (6.14). From (6.11) and (6.14), we get

$$\begin{aligned} & \left| A_T(\phi, \alpha) - \widetilde{M}(\phi, \alpha) \right| \\ & \leq \frac{2C_\gamma\beta_T}{T} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} (C_4 \|(u, v)\| + C_3) \left| \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv + o(1) \\ & \leq \frac{2C_\gamma C_5 \beta_T}{T} + o(1) = o(1), \end{aligned}$$

where the integral is bounded by some constant $C_5 > 0$ because of Assumption 6.12. Since C_γ and C_5 do not depend on (ϕ, α) ,

$$A_T \xrightarrow{T \rightarrow \infty} \widetilde{M} \quad \text{uniformly in } (\phi, \alpha). \quad (6.16)$$

Next, we show that $\mathbb{E}C_T \xrightarrow{T \rightarrow \infty} 0$ uniformly in (ϕ, α) which implies uniform convergence $C_T \xrightarrow{T \rightarrow \infty} 0$ in probability due to the Markov inequality (Theorem B.10 in the Appendix). With the Cauchy-Schwarz inequality and Lemma 5.3, we get that

$$\begin{aligned} 0 &\geq \inf_{(\phi, \alpha) \in \Phi \times \Lambda} \mathbb{E}C_T(\phi, \alpha) \\ &= \inf_{(\phi, \alpha) \in \Phi \times \Lambda} \mathbb{E} \left(- \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{AFM}_{\mathcal{W}^t}^T(u, v) \right|^2 dv \right) \\ &\geq \inf_{(\phi, \alpha) \in \Phi \times \Lambda} \mathbb{E} \left(- \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left(\frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} |d_{u,v}(\sigma_{t'}^\alpha, \rho_{t'}^\phi)|^2 \right) \left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} |\mathcal{AFM}_{\mathcal{W}^t}^T(u, v)|^2 \right) dv \right) \\ &= - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \mathbb{E} \left(|\mathcal{AFM}_{\mathcal{W}^t}^T(u, v)|^2 \right) dv \\ &\geq - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \mathcal{O} \left(\frac{r_T^{2\gamma}}{\beta_T} \right) dv = \mathcal{O} \left(\frac{r_T^{2\gamma} u_T v_T}{\beta_T} \right). \end{aligned}$$

Since $r_T^{2\gamma} u_T v_T \beta_T^{-1} \xrightarrow{T \rightarrow \infty} 0$, $\mathbb{E}C_T(\phi, \alpha) \xrightarrow{T \rightarrow \infty} 0$ uniformly in (ϕ, α) , and thus

$$C_T(\phi, \alpha) \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability uniformly in } (\phi, \alpha). \quad (6.17)$$

Finally, the Cauchy-Schwarz inequality implies that

$$(B_T(\phi, \alpha))^2 \leq 4A_T(\phi, \alpha)C_T(\phi, \alpha) \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability uniformly in } (\phi, \alpha). \quad (6.18)$$

From (6.16), (6.17), (6.18), and Slutsky's Lemma (Theorem B.9 in the Appendix), we conclude that

$$\widetilde{M}_T(\phi, \alpha) = A_T(\phi, \alpha) + o_{\mathbb{P}}(1) \xrightarrow{T \rightarrow \infty} \widetilde{M}(\phi, \alpha)$$

in probability uniformly in (ϕ, α) . \square

Assumption 6.14. For each $j \in \{1, \dots, n\}$, there is a function $\bar{\nu}_j: [0, 1] \rightarrow (0, \nu_{\max}]$ which is integrable such that

$$(\nu_j^t)^2 = \frac{1}{n_t \beta_T} \sum_{i=1}^{\beta_T-1} \mathbb{I}_{\bar{J}_{t+i/T}}(j) (\bar{\nu}_j^{t+i/T})^2 \xrightarrow{T \rightarrow \infty} (\bar{\nu}_j(t))^2 \quad \text{uniformly in } t \in [0, 1].$$

Definition 6.15. Let

$$G_j^{t'} := \Re \left[\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \mathcal{AFM}_{|\mathcal{F}_t|}^2(u, v) \overline{Q_j^t(u, v)} \text{grad}_{(\phi, \alpha)} d_{u,v}(1/\sigma_{t'}^{\alpha_0}, -\rho_{t'}^{\phi_0}) dv dt \right]$$

where $Q_j^t(u, v) := \mathcal{AFM}_{q_j^t}(u, v)$ with $q_j^t(\omega) := \Re(e^{2\pi i \langle \omega, x_j \rangle} \mathcal{F}_{f_t}(\omega))$. We define

$$\Sigma := 16 \int_0^1 \frac{1}{n_{t'}} \sum_{j \in J_{t'}} (\nu_j^{t'})^2 G_j^{t'} (G_j^{t'})^\top dt'. \quad (6.19)$$

Remark 6.16. Note, that $Q_j^t(u, v)$ is well defined because $\omega \mapsto e^{2\pi i \langle \omega, x_j \rangle}$ is continuous for all x_j and thus Lebesgue-measurable, $|q_j^t(\omega)| \leq |\mathcal{F}_{f^t}(\omega)|$, and \mathcal{F}_{f^t} is integrable.

Theorem 6.17 (Central limit theorem for $\text{grad}_{(\phi, \alpha)} M_T(\phi_0, \alpha_0)$). Under the Assumptions 5.6, 6.12, and 6.14, if $T, u_T, v_T, r_T, \beta_T \rightarrow \infty$ such that $\sqrt{T} r_T^{2+\gamma} \beta_T^{-1} \xrightarrow{T \rightarrow \infty} 0$ and $\sqrt{T} u_T v_T \|(u_T, v_T)\| r_T^{2\gamma} \beta_T^{-1} \xrightarrow{T \rightarrow \infty} 0$, we have that

$$\sqrt{T} \text{grad}_{(\phi, \alpha)} M_T(\phi_0, \alpha_0) \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, \Sigma) \quad \text{in distribution,}$$

with covariance matrix Σ from (6.19).

Proof. We will use the decomposition $\widetilde{M}_T = A_T + B_T + C_T$ from (6.9) and show that $\sqrt{T} \text{grad}_{(\phi, \alpha)} A_T(\phi_0, \alpha_0) = 0$, $\sqrt{T} \text{grad}_{(\phi, \alpha)} C_T(\phi_0, \alpha_0) \xrightarrow{T \rightarrow \infty} 0$ in probability, and

$$\sqrt{T} \text{grad}_{(\phi, \alpha)} B_T(\phi_0, \alpha_0) \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, \Sigma) \quad \text{in distribution.}$$

First, consider the gradient of the integrand of A_T . By Lemma A.14 in the Appendix and Lemma 5.7, we get

$$\begin{aligned} & \text{grad}_{(\phi, \alpha)} \left(- \sum_{t, t' \in \mathbb{T}} d_{u, v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) \overline{d_{u, v}(\sigma_{t'}^\alpha / \sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0}) F^{t'}(u, v)} \right) \\ &= -2 \sum_{t, t' \in \mathbb{T}} \Re \left(\text{grad}_{(\phi, \alpha)} d_{u, v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) \overline{d_{u, v}(\sigma_{t'}^\alpha / \sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0}) F^{t'}(u, v)} \right) \\ &= 2 \sum_{t, t' \in \mathbb{T}} \left(2\pi u (\text{grad}_\phi \rho_t^\phi)^\top, -v \sigma_t^{\alpha_0} / \sigma_t^\alpha (\text{grad}_\alpha \sigma_t^\alpha)^\top \right)^\top \\ & \quad \cdot \Im \left(d_{u, v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) \overline{d_{u, v}(\sigma_{t'}^\alpha / \sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0}) F^t(u, v) F^{t'}(u, v)} \right). \end{aligned} \quad (6.20)$$

Because of $d_{u, v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) = d_{u, v}(1, 0) = 1$ and

$$\Im \left(F^t(u, v) \overline{F^{t'}(u, v)} \right) = (\sigma_t^{\alpha_0})^{4-\gamma} (\sigma_{t'}^{\alpha_0})^{4-\gamma} \Im \left(\left| \mathcal{AFM}_{|\mathcal{F}_f|^2}^T(u, v) \right|^2 \right) = 0,$$

(6.20) vanishes for $(\phi, \alpha) = (\phi_0, \alpha_0)$, implying that

$$\sqrt{T} \text{grad}_{(\phi, \alpha)} A_T(\phi_0, \alpha_0) = 0. \quad (6.21)$$

Next, we consider the asymptotic behaviour of $\sqrt{T} \text{grad}_{(\phi, \alpha)} C_T(\phi_0, \alpha_0)$. By Lemma A.14 in the Appendix and (5.9), we get

$$\begin{aligned} & \text{grad}_{(\phi, \alpha)} \left(- \sum_{t, t' \in \mathbb{T}} d_{u, v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{AFM}_{\mathcal{W}^t}(u, v) \overline{d_{u, v}(\sigma_{t'}^\alpha, \rho_{t'}^\phi) \mathcal{AFM}_{\mathcal{W}^{t'}}(u, v)} \right) \\ &= -2 \sum_{t, t' \in \mathbb{T}} \Re \left(\text{grad}_{(\phi, \alpha)} d_{u, v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{AFM}_{\mathcal{W}^t}(u, v) \overline{d_{u, v}(\sigma_{t'}^\alpha, \rho_{t'}^\phi) \mathcal{AFM}_{\mathcal{W}^{t'}}(u, v)} \right) \\ &= 2 \sum_{t \in \mathbb{T}} \sum_{t' \in \mathbb{T} \setminus \{t\}} \left(2\pi u (\text{grad}_\phi \rho_t^\phi)^\top, -v / \sigma_t^\alpha \cdot (\text{grad}_\alpha \sigma_t^\alpha)^\top \right)^\top \\ & \quad \cdot \Im \left(d_{u, v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{AFM}_{\mathcal{W}^t}(u, v) \overline{d_{u, v}(\sigma_{t'}^\alpha, \rho_{t'}^\phi) \mathcal{AFM}_{\mathcal{W}^{t'}}(u, v)} \right), \end{aligned}$$

where the terms with $t' = t$ vanish due to

$$\Im \left(\left| d_{u,v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{A} \mathcal{F} \mathcal{M}_{\mathcal{W}^t}(u, v) \right|^2 \right) = 0.$$

With Lemma 5.7, it follows that

$$\begin{aligned} & \mathbb{E} \left\| \sqrt{T} \operatorname{grad}_{(\phi, \alpha)} C_T(\phi_0, \alpha_0) \right\| \\ &= \mathbb{E} \left\| \sqrt{T} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{2\beta_T^2}{T^2} \sum_{t \in \mathbb{T}} \sum_{t' \in \mathbb{T} \setminus \{t\}} \left(2\pi u (\operatorname{grad}_\phi \rho_t^\phi)^\top, -v/\sigma_t^\alpha \cdot (\operatorname{grad}_\alpha \sigma_t^\alpha)^\top \right)^\top \right. \\ & \quad \cdot \Im \left(d_{u,v}(\sigma_t^\alpha, \rho_t^\phi) \mathcal{A} \mathcal{F} \mathcal{M}_{\mathcal{W}^t}(u, v) \overline{d_{u,v}(\sigma_{t'}^\alpha, \rho_{t'}^\phi) \mathcal{A} \mathcal{F} \mathcal{M}_{\mathcal{W}^{t'}}(u, v)} \right) dv \left. \right\| \\ &\leq 2\tilde{C}\sqrt{T} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \|(u, v)\| \frac{\beta_T^2}{T^2} \sum_{t \in \mathbb{T}} \sum_{t' \in \mathbb{T} \setminus \{t\}} \mathbb{E} |\mathcal{A} \mathcal{F} \mathcal{M}_{\mathcal{W}^t}(u, v)| \mathbb{E} |\mathcal{A} \mathcal{F} \mathcal{M}_{\mathcal{W}^{t'}}(u, v)| dv \\ &\leq 2\tilde{C}\sqrt{T} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \|(u_T, v_T)\| \frac{\beta_T^2}{T^2} \sum_{t \in \mathbb{T}} \sum_{t' \in \mathbb{T} \setminus \{t\}} \mathcal{O} \left(\frac{r_T^{2\gamma}}{\beta_T} \right) dv \\ &= \mathcal{O} \left(\frac{\sqrt{T} u_T v_T \|(u_T, v_T)\| r_T^{2\gamma}}{\beta_T} \right), \end{aligned}$$

where we used that, due to Lemma 5.3,

$$\mathbb{E} |\mathcal{A} \mathcal{F} \mathcal{M}_{\mathcal{W}^t}(u, v)| \leq \sqrt{\mathbb{E} (|\mathcal{A} \mathcal{F} \mathcal{M}_{\mathcal{W}^t}(u, v)|^2)} = \mathcal{O} \left(\frac{r_T^\gamma}{\sqrt{\beta_T}} \right).$$

Since $\sqrt{T} u_T v_T \|(u_T, v_T)\| r_T^{2\gamma} / \beta_T \xrightarrow{T \rightarrow \infty} 0$, we have

$$\sqrt{T} \operatorname{grad}_{(\phi, \alpha)} C_T(\phi_0, \alpha_0) \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \quad (6.22)$$

Finally, we tackle $\sqrt{T} \operatorname{grad}_{(\phi, \alpha)} B_T(\phi_0, \alpha_0)$. We write $B_T = B_T^{(1)} + B_T^{(2)}$ with

$$\begin{aligned} B_T^{(1)}(\phi, \alpha) &:= - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} 2\Re \left[\left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) \right) \right. \\ & \quad \cdot \left. \left(\frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} d_{u,v}(1/\sigma_{t'}^\alpha, -\rho_{t'}^\phi) \overline{\mathcal{A} \mathcal{F} \mathcal{M}_{|\mathcal{W}^{t'}|^2}(u, v)} \right) \right] dv, \\ B_T^{(2)}(\phi, \alpha) &:= - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} 2\Re \left[\left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) F^t(u, v) \right) \right. \\ & \quad \cdot \left. \left(\frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} d_{u,v}(1/\sigma_{t'}^\alpha, -\rho_{t'}^\phi) \overline{\mathcal{A} \mathcal{F} \mathcal{M}_{2\Re(\mathcal{F}_{f^{t'}} \overline{\mathcal{W}^{t'}})}(u, v)} \right) \right] dv. \end{aligned}$$

With Lemma 5.1, we have

$$\begin{aligned} \mathbb{E} \left| \mathcal{A} \mathcal{F} \mathcal{M}_{|\mathcal{W}^t|^2}^T(u, v) \right| &\leq \int_{\Omega_T} \|\omega\|^\gamma \mathbb{E} (|\mathcal{W}^t(\omega)|^2) d\omega \\ &= \int_{\Omega_T} \|\omega\|^\gamma \frac{1}{n_t \beta_T} \sum_{j \in J_T^t} (\nu_j^t)^2 d\omega \leq \frac{\pi \nu_{\max}^2 r_T^{2+\gamma}}{\beta_T}, \end{aligned}$$

because $1/n_t \leq 1$. Hence, with Lemma 5.7,

$$\begin{aligned}
& \mathbb{E} \left\| \sqrt{T} \text{grad}_{\phi, \alpha} B_T^{(1)}(\phi_0, \alpha_0) \right\| \\
&= \mathbb{E} \left\| 2\sqrt{T} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} \right. \\
&\quad \left. \cdot \Re \left[\text{grad}_{(\phi, \alpha)} d_{u, v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) F^t(u, v) \overline{\mathcal{AFM}_{|W^{t'}|^2}(u, v)} \right] dv \right\| \\
&\leq 2\tilde{C}\sqrt{T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \|(u, v)\| |F^t(u, v)| \mathbb{E} \left| \mathcal{AFM}_{|W^{t'}|^2}(u, v) \right| dv \\
&\leq 2\pi\tilde{C}\nu_{\max}^2 \max\{\sigma_{\max}^{4-\gamma}, \sigma_{\min}^{4-\gamma}\} \frac{\sqrt{T}r_T^{2+\gamma}}{\beta_T} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \|(u, v)\| \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}^T(u, v) \right| dv \\
&= \mathcal{O} \left(\frac{\sqrt{T}r_T^{2+\gamma}}{\beta_T} \right),
\end{aligned}$$

due to Assumption 6.12. Since $\sqrt{T}r_T^{2+\gamma}/\beta_T \xrightarrow{T \rightarrow \infty} 0$, we get

$$\sqrt{T} \text{grad}_{\phi, \alpha} B_T^{(1)}(\phi_0, \alpha_0) \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \quad (6.23)$$

It remains to show the asymptotic normality of $\sqrt{T} \text{grad}_{\phi, \alpha} B_T^{(2)}(\phi_0, \alpha_0)$. We have

$$\begin{aligned}
& \mathcal{AFM}_{2\Re(\mathcal{F}_{ft} \overline{W^t})}^T(u, v) \\
&= 2 \int_0^{r_T} \int_0^{2\pi} e^{-2\pi i u \psi} r^{-iv} r^\gamma \left(\Re(\mathcal{F}_{ft} \overline{W^t}) \circ \mathcal{P} \right)(r, \psi) d\psi \frac{dr}{r} \\
&= 2 \int_0^{r_T} \int_0^{2\pi} e^{-2\pi i u \psi} r^{-iv} r^\gamma \Re \left[(\mathcal{F}_{ft} \circ \mathcal{P})(r, \psi) \left(\frac{1}{\sqrt{n_t} \beta_T} \sum_{j \in J_T^t} e^{2\pi i \langle \mathcal{P}(r, \psi), x_j \rangle} \nu_j^t \epsilon_j^t \right) \right] d\psi \frac{dr}{r} \\
&= \frac{2}{\sqrt{n_t} \beta_T} \sum_{j \in J_T^t} \nu_j^t \epsilon_j^t \int_0^{r_T} \int_0^{2\pi} e^{-2\pi i u \psi} r^{-iv} r^\gamma \Re \left[e^{2\pi i \langle \mathcal{P}(r, \psi), x_j \rangle} (\mathcal{F}_{ft} \circ \mathcal{P})(r, \psi) \right] d\psi \frac{dr}{r} \\
&= \frac{2}{\sqrt{n_t} \beta_T} \sum_{j \in J_T^t} \mathcal{AFM}_{q_j^t}^T(u, v) \nu_j^t \epsilon_j^t,
\end{aligned}$$

with $q_j^t(\omega) := \Re(e^{2\pi i \langle \omega, x_j \rangle} \mathcal{F}_{ft}(\omega))$ as in Definition 6.15. Hence,

$$\begin{aligned}
B_T^{(2)}(\phi, \alpha) &= -\frac{2\beta_T}{T} \sum_{t' \in \mathbb{T}} \Re \left[\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} d_{u, v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) \right. \\
&\quad \left. \cdot F^t(u, v) \overline{\mathcal{AFM}_{2\Re(\mathcal{F}_{f t'} \overline{W^{t'}})}(u, v)} \right] dv \\
&= -\frac{2\beta_T}{T} \sum_{t' \in \mathbb{T}} \Re \left[\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} d_{u, v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) \right. \\
&\quad \left. \cdot F^t(u, v) \frac{2}{\sqrt{n_{t'} \beta_T}} \sum_{j \in J_T^{t'}} \overline{\mathcal{AFM}_{q_j^{t'}}^T(u, v)} \nu_j^{t'} \epsilon_j^{t'} \right] dv.
\end{aligned}$$

It follows that

$$B_T^{(2)}(\phi, \alpha) = -\frac{4\sqrt{\beta_T}}{T} \sum_{t' \in \mathbb{T}} \frac{1}{\sqrt{n_{t'}}} \sum_{j \in J_{t'}'} \nu_j^{t'} \epsilon_j^{t'} \Re \left[\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \cdot d_{u,v}(\sigma_t^\alpha / \sigma_{t'}^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) F^t(u, v) \overline{\mathcal{AFM}_{q_j^{t'}}^T(u, v)} dv \right]. \quad (6.24)$$

Now, let $\xi \in \mathbb{R}^{d_2+d_3}$. From (5.9) and (6.24), we get for $(\phi, \alpha) = (\phi_0, \alpha_0)$ that

$$\begin{aligned} & \left\langle \xi, \sqrt{T} \text{grad}_{(\phi, \alpha)} B_T^{(2)}(\phi_0, \alpha_0) \right\rangle \\ &= -4\sqrt{\frac{\beta_T}{T}} \sum_{t' \in \mathbb{T}} \frac{1}{\sqrt{n_{t'}}} \sum_{j \in J_{t'}'} \nu_j^{t'} \epsilon_j^{t'} \Re \left[\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \right. \\ & \quad \left. \cdot \left\langle \xi, \text{grad}_{(\phi, \alpha)} d_{u,v}(1/\sigma_{t'}^{\alpha_0}, -\rho_{t'}^{\phi_0}) \right\rangle F^t(u, v) \overline{\mathcal{AFM}_{q_j^{t'}}^T(u, v)} dv \right] \end{aligned}$$

is a linear combination of independent standard-normal random variables $\epsilon_j^{t'}$ and therefore a centred Gaussian random variable with variance

$$\begin{aligned} & \frac{16\beta_T}{T} \sum_{t' \in \mathbb{T}} \frac{1}{n_{t'}} \sum_{j \in J_{t'}'} (\nu_j^{t'})^2 \Re \left[\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} F^t(u, v) \overline{\mathcal{AFM}_{q_j^{t'}}^T(u, v)} \right. \\ & \quad \left. \cdot \left\langle \xi, \text{grad}_{(\phi, \alpha)} d_{u,v}(1/\sigma_{t'}^{\alpha_0}, -\rho_{t'}^{\phi_0}) \right\rangle dv \right]^2 \\ & \xrightarrow{T \rightarrow \infty} 16 \int_0^1 \frac{1}{n_{t'}} \sum_{j \in J_{t'}'} (\nu_j^{t'})^2 \Re \left[\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \overline{\mathcal{AFM}_{q_j^{t'}}^T(u, v)} \right. \\ & \quad \left. \cdot \left\langle \xi, \text{grad}_{(\phi, \alpha)} d_{u,v}(1/\sigma_{t'}^{\alpha_0}, -\rho_{t'}^{\phi_0}) \right\rangle dv dt \right]^2 dt = \xi^\top \Sigma \xi, \quad (6.25) \end{aligned}$$

with Σ from (6.19). By Corollary B.5 in the Appendix and Lemma A.16 in the Appendix,

$$\sqrt{T} \text{grad}_{(\phi, \alpha)} B_T^{(2)}(\phi_0, \alpha_0) \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, \Sigma) \quad \text{in distribution.} \quad (6.26)$$

From (6.21), (6.22), (6.23), (6.26), and Slutsky's Lemma (Theorem B.9 in the Appendix), we deduce

$$\sqrt{T} \text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\phi_0, \alpha_0) = \sqrt{T} \text{grad}_{(\phi, \alpha)} B_T^{(2)}(\phi_0, \alpha_0) + o_{\mathbb{P}}(1) \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, \Sigma)$$

in distribution, completing the proof. \square

Assumption 6.18. For all $m \in \{1, \dots, d_2\}$ and $m' \in \{1, \dots, d_3\}$, there are Borel sets $B_m, B_{m'} \subseteq [0, 1]$, each of positive Lebesgue-measure, such that the functions

$$B_m \rightarrow \mathbb{R}, \quad t \mapsto \frac{\partial \rho_t^\phi}{\partial \phi_m} \Big|_{\phi=\phi_0}, \quad \text{and} \quad B_{m'} \rightarrow \mathbb{R}, \quad t \mapsto \frac{\partial \sigma_t^\alpha}{\partial \alpha_{m'}} \Big|_{\alpha=\alpha_0},$$

are not constant.

Example 6.19 (Polynomial rotation and scaling). *Suppose that we have polynomial rotation and scaling models, that is,*

$$\rho_t^{\tilde{\phi}} = \sum_{m=0}^{d_2} \phi_m t^m, \quad \sigma_t^{\tilde{\alpha}} = 1 + \sum_{m'=0}^{d_3} \alpha_{m'} t^{m'},$$

where $\tilde{\phi} = (\phi_0, \dots, \phi_{d_2})$, $\tilde{\alpha} = (\alpha_0, \dots, \alpha_{d_3})$. In this case, Assumption 6.18 is an implication of the identifiability constraints $\rho_0^{\tilde{\phi}} = 0$ for all $\tilde{\phi} \in \Phi$ and $\sigma_0^{\tilde{\alpha}} = 1$ for all $\tilde{\alpha} \in \mathbf{A}$ from Assumption 3.3, since those mean that the intercepts ϕ_0 and α_0 are zero. Hence, we have

$$\rho_t^{\phi} = \sum_{m=1}^{d_2} \phi_m t^m, \quad \sigma_t^{\alpha} = 1 + \sum_{m'=1}^{d_3} \alpha_{m'} t^{m'},$$

with $\phi = (\phi_1, \dots, \phi_{d_2})$, $\alpha = (\alpha_1, \dots, \alpha_{d_3})$. Since $\partial \rho_t^{\phi} / \partial \phi_m = t^m$ and $\partial \sigma_t^{\alpha} / \partial \alpha_{m'} = t^{m'}$ are not constant in t for all $m, m' \geq 1$, Assumption 6.18 holds.

Lemma 6.20. *Under the Assumptions 5.6 and 6.12, $\text{Hess}_{(\phi, \alpha)} \widetilde{M}(\phi, \alpha)$ has finite operator norm for all $(\phi, \alpha) \in U$ with $U \subseteq \Phi \times \mathbf{A}$ from Assumption 5.6. Furthermore, the matrix*

$$H_M := \text{Hess}_{(\phi, \alpha)} \widetilde{M}(\phi_0, \alpha_0) \tag{6.27}$$

is symmetric. If the Assumptions 6.6 and 6.18 hold, H_M is also positive definite and hence invertible.

Proof. By Lemma 5.7, Lemma A.15, and Theorem B.2 in the Appendix,

$$\begin{aligned} & \text{Hess}_{(\phi, \alpha)} \widetilde{M}(\phi, \alpha) \\ &= - \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \text{Hess}_{(\phi, \alpha)} \left(\int_0^1 \int_0^1 d_{u,v}(\sigma_t^{\alpha} / \sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0}) F^t(u, v) \right. \\ & \quad \left. \cdot \overline{d_{u,v}(\sigma_{t'}^{\alpha} / \sigma_{t'}^{\alpha_0}, \rho_{t'}^{\phi} - \rho_{t'}^{\phi_0}) F^{t'}(u, v)} dt dt' \right) dv \\ &= -2 \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \int_0^1 \int_0^1 (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \\ & \quad \cdot \Re \left(\overline{d_{u,v}(\sigma_{t'}^{\alpha} / \sigma_{t'}^{\alpha_0}, \rho_{t'}^{\phi} - \rho_{t'}^{\phi_0})} \text{Hess}_{(\phi, \alpha)} d_{u,v}(\sigma_t^{\alpha} / \sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0}) \right. \\ & \quad \left. + \text{grad}_{(\phi, \alpha)} d_{u,v}(\sigma_t^{\alpha} / \sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0}) \overline{\text{grad}_{(\phi, \alpha)}^{\top} d_{u,v}(\sigma_{t'}^{\alpha} / \sigma_{t'}^{\alpha_0}, \rho_{t'}^{\phi} - \rho_{t'}^{\phi_0})} \right) dt dt' dv \\ &= -2 \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \int_0^1 \int_0^1 (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \\ & \quad \cdot \Re \left(d_{u,v}(\sigma_t^{\alpha} / \sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0}) \overline{d_{u,v}(\sigma_{t'}^{\alpha} / \sigma_{t'}^{\alpha_0}, \rho_{t'}^{\phi} - \rho_{t'}^{\phi_0})} \right. \\ & \quad \left. \cdot \left[\mathbf{iH}_{u,v}^{t,0}(\phi, \alpha) - \mathbf{a}_{u,v}^{t,0}(\phi, \alpha) (\mathbf{a}_{u,v}^{t,0}(\phi, \alpha) - \mathbf{a}_{u,v}^{t',0}(\phi, \alpha))^{\top} \right] \right) dt dt' dv. \tag{6.28} \end{aligned}$$

Let $\xi \in \mathbb{R}^{d_2+d_3}$ with $\|\xi\| = 1$. By Assumption 6.12, we have

$$\begin{aligned}
& \left\| \text{Hess}_{(\phi,\alpha)} \widetilde{M}(\phi, \alpha) \xi \right\| \\
& \leq 2 \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \int_0^1 \int_0^1 (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \\
& \quad \cdot \left[\left\| \mathbf{H}_{u,v}^{t,0}(\phi, \alpha) \xi \right\| + \left\| \mathbf{a}_{u,v}^{t,0}(\phi, \alpha) (\mathbf{a}_{u,v}^{t,0}(\phi, \alpha) - \mathbf{a}_{u,v}^{t',0}(\phi, \alpha))^\top \xi \right\| \right] dv \\
& \leq 2 \max\{\sigma_{\max}, \sigma_{\min}\}^{4-\gamma} \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left[\widetilde{C} \|(u, v)\| + 2\widetilde{C}^2 \|(u, v)\|^2 \right] dv \\
& < \infty,
\end{aligned}$$

with $\widetilde{C} > 0$ from Lemma 5.7. Hence, $\text{Hess}_{(\phi,\alpha)} \widetilde{M}(\phi, \alpha)$ has finite operator norm. From (6.28), we get for $(\phi, \alpha) = (\phi_0, \alpha_0)$ that

$$\begin{aligned}
H_M &= \text{Hess}_{(\phi,\alpha)} \widetilde{M}(\phi_0, \alpha_0) \\
&= 2 \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left[\int_0^1 \int_0^1 (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \mathbf{a}_{u,v}^{t',0}(\phi_0, \alpha_0)^\top dt dt' \right. \\
& \quad \left. - \left(\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) dt \right) \left(\int_0^1 (\sigma_{t'}^{\alpha_0})^{4-\gamma} \mathbf{a}_{u,v}^{t',0}(\phi_0, \alpha_0) dt \right)^\top \right] dv. \quad (6.29)
\end{aligned}$$

Since matrices of the form xx^\top with $x \in \mathbb{R}^{d_2+d_3}$ are always symmetric, it follows that H_M is symmetric.

Now, let $\xi^{(2)} \in \mathbb{R}^{d_2}$ and $\xi^{(3)} \in \mathbb{R}^{d_3}$ such that $\xi := ((\xi^{(2)})^\top, (\xi^{(3)})^\top)^\top \neq 0$. By Assumption 6.6 and Lemma 6.4, there are $u \in \mathbb{Z} \setminus \{0\}$ and an open Borel set $B \subseteq \mathbb{R}$ with positive Lebesgue-measure such that $\mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \neq 0$ for all $v \in B$. If Assumption 6.18 holds, there is another Borel set $B' \subseteq [0, 1]$ with positive Lebesgue-measure such that

$$B' \rightarrow \mathbb{R}, \quad t \mapsto \langle \xi, \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \rangle = 2\pi u \sum_{m=1}^{d_2} \xi_m^{(2)} \frac{\partial \rho_t^\phi}{\partial \phi_m} \Big|_{\phi=\phi_0} - v \sum_{m'=1}^{d_3} \xi_{m'}^{(3)} \frac{\partial \sigma_{t'}^\alpha}{\partial \alpha_{m'}} \Big|_{\alpha=\alpha_0}$$

is not constant for all $v \in \mathbb{R} \setminus V^0$ with some Lebesgue null-set $V^0 \subseteq \mathbb{R}$. From the Cauchy-Schwarz inequality, we have that

$$\left(\int_0^1 g_1(t) g_2(t) dt \right)^2 \leq \int_0^1 g_1(t)^2 dt \int_0^1 g_2(t')^2 dt'$$

for all integrable functions $g_1, g_2: [0, 1] \rightarrow \mathbb{R}$, with equality if and only if g_1 and g_2 are linearly dependent a.e. Let

$$g_1^{u,v}(t) := (\sigma_t^{\alpha_0})^{2-\gamma/2} \langle \xi, \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \rangle, \quad g_2^{u,v}(t) := (\sigma_t^{\alpha_0})^{2-\gamma/2}.$$

For all $v \in \mathbb{R} \setminus V^0$, these are not linearly dependent, since $t \mapsto \langle \xi, \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \rangle$ is not constant. Hence,

$$\left(\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \langle \xi, \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \rangle dt \right)^2 < \int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \langle \xi, \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \rangle^2 dt \int_0^1 (\sigma_{t'}^{\alpha_0})^{4-\gamma} dt'$$

for all $v \in \mathbb{R} \setminus V^0$. It follows that

$$\begin{aligned} \xi^\top H_M \xi &\geq 2 \int_{B \setminus V^0} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left[\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \langle \xi, \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \rangle^2 dt \int_0^1 (\sigma_{t'}^{\alpha_0})^{4-\gamma} dt' \right. \\ &\quad \left. - \left(\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \langle \xi, \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \rangle dt \right)^2 \right] dv > 0, \end{aligned} \quad (6.30)$$

as the integrand (as a function in v) is strictly positive. We conclude that H_M is symmetric and positive definite and thus invertible. \square

Assumption 6.21. *We have the following Sobolev-2 condition,*

$$\int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \|(u, v)\|^4 \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv < \infty.$$

Theorem 6.22. *Under the Assumption 5.6, let $(\hat{\phi}_T^*, \hat{\alpha}_T^*)_{T \in \mathbb{N}}$ a sequence of random vectors with values in U , such that $(\hat{\phi}_T^*, \hat{\alpha}_T^*) \xrightarrow{T \rightarrow \infty} (\phi_0, \alpha_0)$ in probability. Suppose that the Assumptions 6.2, 6.7, 6.14, and 6.21 hold. Then, with H_M from (6.27),*

$$\left\| \text{Hess}_{(\phi, \alpha)} \widetilde{M}_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - H_M \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability,}$$

Proof. We will use the decomposition $\widetilde{M}_T = A_T + B_T + C_T$ from (6.9) and show that

$$\begin{aligned} &\left\| \text{Hess}_{(\phi, \alpha)} A_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - H_M \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability,} \\ \sup_{(\phi, \alpha) \in U} &\left\| \text{Hess}_{(\phi, \alpha)} B_T(\phi, \alpha) + \text{Hess}_{(\phi, \alpha)} C_T(\phi, \alpha) \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \end{aligned}$$

By Lemma 5.7 and Lemma A.14 in the Appendix, we get

$$\begin{aligned} &\text{Hess}_{(\phi, \alpha)} A_T(\phi, \alpha) \\ = & - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{2\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \Re \left[F^t(u, v) \overline{F^{t'}(u, v)} \right. \\ &\cdot \left(\overline{d_{u,v}(\sigma_{t'}^\alpha / \sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0})} \text{Hess}_{(\phi, \alpha)} d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) \right. \\ &\left. \left. + \text{grad}_{(\phi, \alpha)} d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) \overline{\text{grad}_{(\phi, \alpha)} d_{u,v}(\sigma_{t'}^\alpha / \sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0})} \right) \right] dv \\ = & - \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{2\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \Re \left[F^t(u, v) \overline{F^{t'}(u, v)} \right. \\ &\cdot \left(\overline{d_{u,v}(\sigma_{t'}^\alpha / \sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0})} \text{Hess}_{(\phi, \alpha)} d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) \right. \\ &\left. \left. + \text{grad}_{(\phi, \alpha)} d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) \overline{\text{grad}_{(\phi, \alpha)} d_{u,v}(\sigma_{t'}^\alpha / \sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0})} \right) \right] dv \\ = & -2 \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \\ &\cdot \Re \left[d_{u,v}(\sigma_t^\alpha / \sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) \overline{d_{u,v}(\sigma_{t'}^\alpha / \sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0})} \right. \\ &\left. \cdot \left(i \mathbf{H}_{u,v}^{t,0}(\phi, \alpha) - \mathbf{a}_{u,v}^{t,0}(\phi, \alpha) (\mathbf{a}_{u,v}^{t,0}(\phi, \alpha) - \mathbf{a}_{u,v}^{t',0}(\phi, \alpha))^\top \right) \right] dv. \end{aligned} \quad (6.31)$$

Since $\phi \mapsto \rho_t^\phi$ and $\alpha \mapsto \sigma_t^\alpha$ are continuous (Assumption 6.7) and $d_{u,v}$ is Lipschitz-continuous (Lemma A.8 in the Appendix), the continuous mapping theorem (Theorem B.8 in the Appendix) yields that

$$\begin{aligned} \left| d_{u,v}(\sigma_t^{\hat{\alpha}_T^*}/\sigma_t^{\alpha_0}, \rho_t^{\hat{\phi}_T^*} - \rho_t^{\phi_0}) - 1 \right| &\leq \sqrt{2} \|(2\pi u, -v)\| \left\| \begin{pmatrix} \rho_t^{\hat{\phi}_T^*} - \rho_t^{\phi_0} \\ \log(\sigma_t^{\hat{\alpha}_T^*}) - \log(\sigma_t^{\alpha_0}) \end{pmatrix} \right\| \\ &= o_{\mathbb{P}}(\|(u, v)\|), \end{aligned}$$

where the logarithm is Lipschitz-continuous on $[\sigma_{\min}, \sigma_{\max}]$ by Lemma A.10 in the Appendix. Hence,

$$d_{u,v}(\sigma_t^{\hat{\alpha}_T^*}/\sigma_t^{\alpha_0}, \rho_t^{\hat{\phi}_T^*} - \rho_t^{\phi_0}) \overline{d_{u,v}(\sigma_{t'}^{\hat{\alpha}_T^*}/\sigma_{t'}^{\alpha_0}, \rho_{t'}^{\hat{\phi}_T^*} - \rho_{t'}^{\phi_0})} = 1 + o_{\mathbb{P}}(\|(u, v)\|^2).$$

In particular, the imaginary part vanishes asymptotically. With Lemma 5.7 and Assumption 6.12, it follows that

$$\begin{aligned} &\left\| \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \right. \\ &\quad \cdot \Re \left[d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) \overline{d_{u,v}(\sigma_{t'}^\alpha/\sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0})} \mathbf{i}\mathbf{H}_{u,v}^{t,0}(\phi, \alpha) \right] dv \Big\|_1 \\ &\leq \max \{ \sigma_{\min}^{8-2\gamma}, \sigma_{\max}^{8-2\gamma} \} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \\ &\quad \cdot \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \left| \Im \left[d_{u,v}(\sigma_t^\alpha/\sigma_t^{\alpha_0}, \rho_t^\phi - \rho_t^{\phi_0}) \overline{d_{u,v}(\sigma_{t'}^\alpha/\sigma_{t'}^{\alpha_0}, \rho_{t'}^\phi - \rho_{t'}^{\phi_0})} \right] \right| \|\mathbf{H}_{u,v}^{t,0}(\phi, \alpha)\|_1 dv \\ &= o_{\mathbb{P}} \left(\int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \|(u, v)\|^3 \left| \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv \right) = o_{\mathbb{P}}(1). \end{aligned} \quad (6.32)$$

Moreover,

$$\int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{A}\mathcal{F}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \Re \left[\mathbf{i}\mathbf{H}_{u,v}^{t,0}(\phi, \alpha) \right] dv = 0. \quad (6.33)$$

From Assumption 5.6, Lemma 5.7, and the continuous mapping theorem (Theorem B.8 in the Appendix), we get that

$$\begin{aligned} &\left\| \mathbf{a}_{u,v}^{t,0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) (\mathbf{a}_{u,v}^{t,0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) - \mathbf{a}_{u,v}^{t',0}(\hat{\phi}_T^*, \hat{\alpha}_T^*))^\top \right. \\ &\quad \left. - \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) - \mathbf{a}_{u,v}^{t',0}(\phi_0, \alpha_0))^\top \right\|_1 \\ &\leq \left\| \mathbf{a}_{u,v}^{t,0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) \left[(\mathbf{a}_{u,v}^{t,0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) - \mathbf{a}_{u,v}^{t',0}(\hat{\phi}_T^*, \hat{\alpha}_T^*))^\top - (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) - \mathbf{a}_{u,v}^{t',0}(\phi_0, \alpha_0))^\top \right] \right\|_1 \\ &\quad + \left\| (\mathbf{a}_{u,v}^{t,0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) - \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0)) (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) - \mathbf{a}_{u,v}^{t',0}(\phi_0, \alpha_0))^\top \right\|_1 \\ &\leq \left\| \mathbf{a}_{u,v}^{t,0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) \right\|_1 \left(\left\| \mathbf{a}_{u,v}^{t,0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) - \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \right\|_1 + \left\| \mathbf{a}_{u,v}^{t',0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) - \mathbf{a}_{u,v}^{t',0}(\phi_0, \alpha_0) \right\|_1 \right) \\ &\quad + \left\| \mathbf{a}_{u,v}^{t,0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) - \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \right\|_1 \left(\left\| \mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0) \right\|_1 + \left\| \mathbf{a}_{u,v}^{t',0}(\phi_0, \alpha_0) \right\|_1 \right) \end{aligned}$$

$$\begin{aligned}
&\leq (d_2 + d_3)\tilde{C} \|(u, v)\| \\
&\quad \cdot \left[3 \left(2\pi |u| \left\| \text{grad}_\phi(\rho_t^{\hat{\phi}_T^*}) - \text{grad}_\phi(\rho_t^{\phi_0}) \right\| + |v| \left\| \sigma_t^{\alpha_0} / \sigma_t^{\hat{\alpha}_T^*} \text{grad}_\alpha(\sigma_t^{\hat{\alpha}_T^*}) - \text{grad}_\alpha(\sigma_t^{\alpha_0}) \right\| \right) \right. \\
&\quad \left. + \left(2\pi |u| \left\| \text{grad}_\phi(\rho_{t'}^{\hat{\phi}_T^*}) - \text{grad}_\phi(\rho_{t'}^{\phi_0}) \right\| + |v| \left\| \sigma_{t'}^{\alpha_0} / \sigma_{t'}^{\hat{\alpha}_T^*} \text{grad}_\alpha(\sigma_{t'}^{\hat{\alpha}_T^*}) - \text{grad}_\alpha(\sigma_{t'}^{\alpha_0}) \right\| \right) \right] \\
&= o_{\mathbb{P}}(\|(u, v)\|^2).
\end{aligned}$$

This, together with (6.31), (6.32), (6.33), and Assumption 5.6, gives

$$\begin{aligned}
&\left\| \text{Hess}_{(\phi, \alpha)} A_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - \text{Hess}_{(\phi, \alpha)} A_T(\phi_0, \alpha_0) \right\|_1 \\
&= 2 \left\| \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \right. \\
&\quad \cdot \left\{ \Re \left[d_{u, v}(\sigma_t^{\hat{\alpha}_T^*} / \sigma_t^{\alpha_0}, \rho_t^{\hat{\phi}_T^*} - \rho_t^{\phi_0}) \overline{d_{u, v}(\sigma_{t'}^{\hat{\alpha}_T^*} / \sigma_{t'}^{\alpha_0}, \rho_{t'}^{\hat{\phi}_T^*} - \rho_{t'}^{\phi_0})} \right] \right. \\
&\quad \cdot \mathbf{a}_{u, v}^{t, 0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) (\mathbf{a}_{u, v}^{t, 0}(\hat{\phi}_T^*, \hat{\alpha}_T^*) - \mathbf{a}_{u, v}^{t', 0}(\hat{\phi}_T^*, \hat{\alpha}_T^*))^\top \\
&\quad \left. \left. - \mathbf{a}_{u, v}^{t, 0}(\phi_0, \alpha_0) (\mathbf{a}_{u, v}^{t, 0}(\phi_0, \alpha_0) - \mathbf{a}_{u, v}^{t', 0}(\phi_0, \alpha_0))^\top \right\} dv \right\|_1 + o_{\mathbb{P}}(1) \\
&\leq o_{\mathbb{P}} \left(2 \max \{ \sigma_{\min}^{8-2\gamma}, \sigma_{\max}^{8-2\gamma} \} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \|(u, v)\|^4 \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv \right) + o_{\mathbb{P}}(1) \\
&= o_{\mathbb{P}}(1). \tag{6.34}
\end{aligned}$$

By Lemma 6.20, H_M has finite operator norm. In particular, the components of H_M are finite. Hence,

$$\begin{aligned}
&\left\| \int_0^1 \int_0^1 (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \mathbf{a}_{u, v}^{t, 0}(\phi_0, \alpha_0) (\mathbf{a}_{u, v}^{t, 0}(\phi_0, \alpha_0) - \mathbf{a}_{u, v}^{t', 0}(\phi_0, \alpha_0))^\top dt dt' dv \right\|_1 \\
&\quad \cdot 2 \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 = o(1),
\end{aligned}$$

as $T \rightarrow \infty$. With (6.31) and (6.34), this implies that

$$\begin{aligned}
&\left\| \text{Hess}_{(\phi, \alpha)} A_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - H_M \right\|_1 \\
&= \left\| \text{Hess}_{(\phi, \alpha)} A_T(\phi_0, \alpha_0) - H_M \right\|_1 + o_{\mathbb{P}}(1) \\
&\leq 2 \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 \left(\left\| \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \mathbf{A}_{u, v}^{t, t'} - \int_0^1 \int_0^1 \mathbf{A}_{u, v}^{t, t'} dt dt' \right\|_1 \right. \\
&\quad \left. + \left\| \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \mathbf{B}_{u, v}^{t, t'} - \int_0^1 \int_0^1 \mathbf{B}_{u, v}^{t, t'} dt dt' \right\|_1 \right) dv + o_{\mathbb{P}}(1), \tag{6.35}
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{A}_{u, v}^{t, t'} &:= (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \mathbf{a}_{u, v}^{t, 0}(\phi_0, \alpha_0) \mathbf{a}_{u, v}^{t', 0}(\phi_0, \alpha_0)^\top, \\
\mathbf{B}_{u, v}^{t, t'} &:= (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \mathbf{a}_{u, v}^{t, 0}(\phi_0, \alpha_0) \mathbf{a}_{u, v}^{t', 0}(\phi_0, \alpha_0)^\top.
\end{aligned}$$

By Lemma A.5, Lemma A.7, Lemma A.9, and Lemma A.10 (all in the Appendix), and Assumption 6.2, we have that

$$\begin{aligned}
& \left\| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \mathbf{A}_{u,v}^{t,t'} - \int_0^1 \int_0^1 \mathbf{A}_{u,v}^{t,t'} dt dt' \right\|_1 \\
&= \sum_{m,m'=1}^{d_2+d_3} \left| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\mathbf{A}_{u,v}^{t,t'} \right)_{m,m'} - \int_0^1 \int_0^1 \left(\mathbf{A}_{u,v}^{t,t'} \right)_{m,m'} dt dt' \right| \\
&\leq C_\gamma \tilde{C}^2 \|(u, v)\|^2 \frac{\beta_T}{T} \\
&\quad \cdot \sum_{m,m'=1}^{d_2+d_3} \left\{ \text{TV} \left[t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma} (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0))_m (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0))_{m'} \right] + \text{TV} \left[t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma} \right] \right\} \\
&\leq C_\gamma \tilde{C}^2 \|(u, v)\|^2 \frac{\beta_T}{T} \left\{ (d_2 + d_3) C_\gamma \tilde{C} \|(u, v)\| \sum_{m=1}^{d_2+d_3} \text{TV} \left[t \mapsto (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0))_m \right] \right. \\
&\quad \left. + (d_2 + d_3) C_\gamma \tilde{C} \|(u, v)\| \sum_{m'=1}^{d_2+d_3} \text{TV} \left[t \mapsto (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0))_{m'} \right] \right. \\
&\quad \left. + (d_2 + d_3)^2 C_3 \tilde{C}^2 \|(u, v)\|^2 + (d_2 + d_3)^2 C_3 \right\} \\
&\leq (d_2 + d_3) C_\gamma \tilde{C}^2 \|(u, v)\|^2 \frac{\beta_T}{T} \left\{ 2C_\gamma \tilde{C} \|(u, v)\| \left(2\pi |u| \sum_{m=1}^{d_2} \text{TV} \left[t \mapsto \frac{\partial \rho_t^\phi}{\partial \phi_m} \Big|_{\phi=\phi_0} \right] \right. \right. \\
&\quad \left. \left. + |v| \sum_{m'=1}^{d_3} \text{TV} \left[t \mapsto \frac{\partial \sigma_t^\alpha}{\partial \alpha_{m'}} \Big|_{\alpha=\alpha_0} \right] \right) + (d_2 + d_3) C_3 (\tilde{C}^2 \|(u, v)\|^2 + 1) \right\} \\
&= \mathcal{O}_{\mathbb{P}} \left(\|(u, v)\|^4 \frac{\beta_T}{T} \right), \tag{6.36}
\end{aligned}$$

with $\tilde{C} > 0$ from Lemma 5.7, $C_\gamma := \max \{ \sigma_{\min}^{4-\gamma}, \sigma_{\max}^{4-\gamma} \}$, and $C_3 > 0$ from (6.15). Similarly,

$$\begin{aligned}
& \left\| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \mathbf{B}_{u,v}^{t,t'} - \int_0^1 \int_0^1 \mathbf{B}_{u,v}^{t,t'} dt dt' \right\|_1 \\
&\leq C_\gamma \tilde{C} \|(u, v)\| \frac{\beta_T}{T} \\
&\quad \cdot \sum_{m,m'=1}^{d_2+d_3} \left\{ \text{TV} \left[t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma} (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0))_m \right] + \text{TV} \left[t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma} (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0))_{m'} \right] \right\} \\
&\leq 2(d_2 + d_3) C_\gamma \tilde{C} \|(u, v)\| \frac{\beta_T}{T} \\
&\quad \cdot \left\{ C_\gamma \sum_{m=1}^{d_2+d_3} \text{TV} \left[t \mapsto (\mathbf{a}_{u,v}^{t,0}(\phi_0, \alpha_0))_m \right] + (d_2 + d_3) C_3 \tilde{C} \|(u, v)\| \right\} \\
&= \mathcal{O}_{\mathbb{P}} \left(\|(u, v)\|^2 \frac{\beta_T}{T} \right). \tag{6.37}
\end{aligned}$$

From (6.35), (6.36), and (6.37), we get that

$$\left\| \text{Hess}_{(\phi, \alpha)} A_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - H_M \right\|_1 = \mathcal{O}_{\mathbb{P}} \left(\frac{\beta_T}{T} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \|(u, v)\|^4 \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right|^2 dv \right), \quad (6.38)$$

which converges to 0 due to Assumption 6.21.

Next, we show that $\sup_{(\phi, \alpha) \in U} \left\| \text{Hess}_{(\phi, \alpha)} B_T(\phi, \alpha) \right\|_1 \xrightarrow{T \rightarrow \infty} 0$ in probability. From Lemma 5.7, we have

$$\sup_{(\phi, \alpha) \in U} \left\| \text{Hess}_{(\phi, \alpha)} d_{u, v}(\sigma_t^\alpha / \sigma_t^{\alpha_0} / \sigma_{t'}^\alpha, \rho_t^\phi - \rho_t^{\phi_0} - \rho_{t'}^\phi) \right\|_1 = \mathcal{O}(\|(u, v)\|^2). \quad (6.39)$$

By Lemma 5.3 and (6.39), we get that

$$\begin{aligned} & \sup_{(\phi, \alpha) \in U} \left\| \text{Hess}_{(\phi, \alpha)} B_T(\phi, \alpha) \right\|_1 \\ &= \sup_{(\phi, \alpha) \in U} \left\| \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} 2\Re \left[\frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} F^t(u, v) \overline{\mathcal{AFM}_{\mathcal{W}^{t'}}(u, v)} \right. \right. \\ & \quad \left. \left. \cdot \text{Hess}_{(\phi, \alpha)} d_{u, v}(\sigma_t^\alpha / \sigma_t^{\alpha_0} / \sigma_{t'}^\alpha, \rho_t^\phi - \rho_t^{\phi_0} - \rho_{t'}^\phi) \right] dv \right\|_1 \\ &\leq 2C_\gamma \sup_{(\phi, \alpha) \in U} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right| \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} |\mathcal{AFM}_{\mathcal{W}^{t'}}(u, v)| \\ & \quad \cdot \left\| \text{Hess}_{(\phi, \alpha)} d_{u, v}(\sigma_t^\alpha / \sigma_t^{\alpha_0} / \sigma_{t'}^\alpha, \rho_t^\phi - \rho_t^{\phi_0} - \rho_{t'}^\phi) \right\|_1 dv \\ &= \mathcal{O}_{\mathbb{P}} \left(\frac{r_T^\gamma}{\sqrt{\beta_T}} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \|(u, v)\|^2 \left| \mathcal{AFM}_{|\mathcal{F}_f|^2}(u, v) \right| dv \right). \end{aligned}$$

Because of Assumption 6.21 and $r_T^{2\gamma} / \beta_T \xrightarrow{T \rightarrow \infty} 0$, the above converges to 0 and we conclude that

$$\sup_{(\phi, \alpha) \in U} \left\| \text{Hess}_{(\phi, \alpha)} B_T(\phi, \alpha) \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \quad (6.40)$$

Finally, we consider $\sup_{(\phi, \alpha) \in U} \left\| \text{Hess}_{(\phi, \alpha)} C_T(\phi, \alpha) \right\|_1$. By Lemma 5.3 and (6.39),

$$\begin{aligned} & \sup_{(\phi, \alpha) \in U} \left\| \text{Hess}_{(\phi, \alpha)} C_T(\phi, \alpha) \right\|_1 \\ &= \sup_{(\phi, \alpha) \in U} \left\| \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \text{Hess}_{(\phi, \alpha)} d_{u, v}(\sigma_t^\alpha / \sigma_t^\alpha, \rho_t^\phi - \rho_{t'}^\phi) \right. \\ & \quad \left. \cdot \mathcal{AFM}_{\mathcal{W}^t}(u, v) \overline{\mathcal{AFM}_{\mathcal{W}^{t'}}(u, v)} dv \right\|_1 \\ &\leq \sup_{(\phi, \alpha) \in U} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \left\| \text{Hess}_{(\phi, \alpha)} d_{u, v}(\sigma_t^\alpha / \sigma_t^\alpha, \rho_t^\phi - \rho_{t'}^\phi) \right\|_1 \\ & \quad \cdot |\mathcal{AFM}_{\mathcal{W}^t}(u, v)| |\mathcal{AFM}_{\mathcal{W}^{t'}}(u, v)| dv \\ &= \mathcal{O} \left(\frac{r_T^{2\gamma}}{\beta_T} \int_{-v_T}^{v_T} \sum_{|u| \leq u_T} \|(u, v)\|^2 dv \right) = \mathcal{O} \left(\frac{\|(u_T, v_T)\|^2 u_T v_T r_T^{2\gamma}}{\beta_T} \right). \end{aligned}$$

Since $\|(u_T, v_T)\|^2 u_T v_T r_T^{2\gamma} / \beta_T \xrightarrow{T \rightarrow \infty} 0$, we get

$$\sup_{(\phi, \alpha) \in U} \left\| \text{Hess}_{(\phi, \alpha)} C_T(\phi, \alpha) \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \quad (6.41)$$

From (6.38), (6.40), and (6.41), we conclude that

$$\left\| \text{Hess}_{(\phi, \alpha)} \widetilde{M}_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - H_M \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.}$$

□

Theorem 6.23 (Central limit theorem for rotation and scaling parameters). *Suppose that Assumption 6.18 as well as the Assumptions of the Theorems 6.17 and 6.22 hold. Then,*

$$\sqrt{T} \begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, H_M^{-1} \Sigma H_M^{-1}) \quad \text{in distribution,}$$

with Σ from Definition 6.15 and H_M from (6.27).

Proof. By Assumption 6.2, \widetilde{M}_T is twice continuously differentiable in a convex open neighborhood $U \subseteq \Phi \times \mathbf{A}$ of (ϕ_0, α_0) . In particular, if \widetilde{M}_T has a minimum at some $(\phi, \alpha) \in U$, then $\text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\phi, \alpha) = 0$. Let

$$\widetilde{G}_T(\phi, \alpha) := \begin{cases} \text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\phi, \alpha), & \text{if } (\phi, \alpha) \in U, \\ \text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\phi_0, \alpha_0), & \text{if } (\phi, \alpha) \in (\Phi \times \mathbf{A}) \setminus U. \end{cases}$$

As $(\hat{\phi}_T, \hat{\alpha}_T)$ is defined as a minimizer of \widetilde{M}_T and $(\hat{\phi}_T, \hat{\alpha}_T) \xrightarrow{T \rightarrow \infty} (\phi_0, \alpha_0) \in U$ in probability, we have for all $\epsilon > 0$ that

$$\begin{aligned} \mathbb{P} \left(\sqrt{T} \widetilde{G}_T(\hat{\phi}_T, \hat{\alpha}_T) > \epsilon \right) &= \mathbb{P} \left(\sqrt{T} \text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\hat{\phi}_T, \hat{\alpha}_T) > \epsilon, (\hat{\phi}_T, \hat{\alpha}_T) \in U \right) \\ &\quad + \mathbb{P} \left(\sqrt{T} \text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\phi_0, \alpha_0) > \epsilon, (\hat{\phi}_T, \hat{\alpha}_T) \notin U \right) \\ &\leq \mathbb{P} \left((\hat{\phi}_T, \hat{\alpha}_T) \notin U \right) \xrightarrow{T \rightarrow \infty} 0, \end{aligned}$$

which means that

$$\sqrt{T} \widetilde{G}_T(\hat{\phi}_T, \hat{\alpha}_T) = o_{\mathbb{P}}(1). \quad (6.42)$$

For $(\phi, \alpha) \in U$, we can apply the mean value theorem for real functions of multiple variables to each component of $\text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\phi, \alpha)$ to get that

$$\text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\phi, \alpha) = \text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\phi_0, \alpha_0) + \text{Hess}_{(\phi, \alpha)} \widetilde{M}_T(\phi^\dagger, \alpha^\dagger) \begin{pmatrix} \phi - \phi_0 \\ \alpha - \alpha_0 \end{pmatrix}, \quad (6.43)$$

where $(\phi^\dagger, \alpha^\dagger) \in U$ such that its components are convex combinations of the respective components of (ϕ, α) and (ϕ_0, α_0) . By (6.43), on the event $\{(\hat{\phi}_T, \hat{\alpha}_T) \in U\}$, we can find $(\hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger) \in U$ between $(\hat{\phi}_T, \hat{\alpha}_T)$ and (ϕ_0, α_0) such that

$$\text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\hat{\phi}_T, \hat{\alpha}_T) = \text{grad}_{(\phi, \alpha)} \widetilde{M}_T(\phi_0, \alpha_0) + \text{Hess}_{(\phi, \alpha)} \widetilde{M}_T(\hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger) \begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix}. \quad (6.44)$$

With the definitions

$$(\hat{\phi}_T^*, \hat{\alpha}_T^*) := \begin{cases} (\hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger), & \text{if } (\hat{\phi}_T, \hat{\alpha}_T) \in U, \\ (\phi_0, \alpha_0), & \text{if } (\hat{\phi}_T, \hat{\alpha}_T) \in (\Phi \times \mathbf{A}) \setminus U, \end{cases}$$

and

$$\tilde{H}_T := \begin{cases} \text{Hess}_{(\phi, \alpha)} \tilde{M}_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) = \text{Hess}_{(\phi, \alpha)} \tilde{M}_T(\hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger), & \text{if } (\hat{\phi}_T, \hat{\alpha}_T) \in U, \\ 0, & \text{if } (\hat{\phi}_T, \hat{\alpha}_T) \in (\Phi \times \mathbb{A}) \setminus U, \end{cases}$$

as well as (6.42), we get that

$$\sqrt{T} \text{grad}_{(\phi, \alpha)} \tilde{M}_T(\phi_0, \alpha_0) + \tilde{H}_T \sqrt{T} \begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} = \sqrt{T} \tilde{G}_T(\hat{\phi}_T, \hat{\alpha}_T) = o_{\mathbb{P}}(1), \quad (6.45)$$

which holds on $\{(\hat{\phi}_T, \hat{\alpha}_T) \notin U\}$ by design of \tilde{G}_T and \tilde{H}_T and on $\{(\hat{\phi}_T, \hat{\alpha}_T) \in U\}$ due to (6.44). Equation (6.45) and Theorem 6.17 yield that

$$\tilde{H}_T \sqrt{T} \begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} = -\sqrt{T} \text{grad}_{(\phi, \alpha)} \tilde{M}_T(\phi_0, \alpha_0) + o_{\mathbb{P}}(1) = \mathcal{O}_{\mathbb{P}}(1). \quad (6.46)$$

Since $(\hat{\phi}_T^*, \hat{\alpha}_T^*) \in U$ is between $(\hat{\phi}_T, \hat{\alpha}_T)$ and (ϕ_0, α_0) and $(\hat{\phi}_T, \hat{\alpha}_T)$ is a consistent estimator, we have that $(\hat{\phi}_T^*, \hat{\alpha}_T^*) \xrightarrow{T \rightarrow \infty} (\phi_0, \alpha_0)$ in probability. Because of Assumption 6.18 and Lemma 6.20, H_M is invertible, and by Theorem 6.22, $\tilde{H}_T \xrightarrow{T \rightarrow \infty} H_M$ in probability. Together with (6.46) and Lemma A.17 in the Appendix, this implies that

$$\sqrt{T} \begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} = \mathcal{O}_{\mathbb{P}}(1).$$

Hence, again with Theorem 6.22,

$$\left(\tilde{H}_T - H_M \right) \sqrt{T} \begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} = o_{\mathbb{P}}(1) \mathcal{O}_{\mathbb{P}}(1) = o_{\mathbb{P}}(1).$$

From this, (6.46), and Theorem 6.17, it follows that

$$H_M \sqrt{T} \begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} = -\sqrt{T} \text{grad}_{(\phi, \alpha)} \tilde{M}_T(\phi_0, \alpha_0) + o_{\mathbb{P}}(1) \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, \Sigma) \quad \text{in distribution,}$$

where we used that, for all centred normal random vectors X , X and $-X$ have the same distribution. Finally, multiplication with H_M^{-1} yields the assertion. \square

Corollary 6.24. *Under the Assumption 6.2, if $\sqrt{T}(\hat{\phi}_T - \phi_0, \hat{\alpha}_T - \alpha_0)$ is asymptotically centred normal, we have for all $\omega \in \mathbb{R}^2 \setminus \{0\}$ and $t \in [0, 1]$, that*

$$\sqrt{T} \begin{pmatrix} \frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T - \rho_t^{\phi_0}} \omega - \omega} \end{pmatrix}$$

is asymptotically centred normal.

Proof. For $\omega \in \mathbb{R}^2 \setminus \{0\}$ and $t \in [0, 1]$, let

$$g_\omega^t : \mathbb{R}^{d_2 + d_3} \rightarrow \mathbb{R}^2, \quad (\phi, \alpha) \mapsto \frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha} R_{\rho_t^\phi - \rho_t^{\phi_0}} \omega.$$

By Assumption 6.2, there is an open neighborhood $U \subseteq \Phi \times \mathbb{A}$ of (ϕ_0, α_0) such that $\phi \mapsto \rho_t^\phi$ and $\alpha \mapsto \sigma_t^\alpha$ are continuously differentiable on U . Hence, g_ω^t is continuously differentiable on U . Because $g_\omega^t(\phi_0, \alpha_0) = \omega$, applying the Delta method (Theorem B.7 in the Appendix) yields the assertion. \square

6.2 Drift Estimation

Assumption 6.25. *Suppose that Assumption 5.9 holds. We assume that the second partial derivatives*

$$\theta \mapsto \frac{\partial^2(\delta_t^\theta)_1}{\partial\theta_m\partial\theta_{m'}}, \quad \theta \mapsto \frac{\partial^2(\delta_t^\theta)_2}{\partial\theta_m\partial\theta_{m'}}$$

are continuous at the true parameter θ_0 for all $m, m' \in \{1, \dots, d_1\}$. Moreover, the first partial derivatives at θ_0 , as functions in t , are of bounded variation, that is, there is a $C' > 0$ such that

$$\text{TV} \left(t \mapsto \frac{\partial(\delta_t^\theta)_1}{\partial\theta_m} \right) < C', \quad \text{TV} \left(t \mapsto \frac{\partial(\delta_t^\theta)_2}{\partial\theta_m} \right) < C',$$

for all $m \in \{1, \dots, d_1\}$ and all $\theta \in \Theta$.

Definition 6.26 (Not translation invariant). *A function $g: \mathbb{R}^2 \rightarrow \mathbb{C}$ is called not translation invariant if there is no $\delta \in \mathbb{R}^2 \setminus \{0\}$ such that $g(x) = g(x - \delta)$ for all $x \in \mathbb{R}^2$.*

Lemma 6.27. *A function $g \in L^2(\mathbb{R}^2)$ is not translation invariant if and only if there is an open Borel set $B \subseteq \mathbb{R}^2$ with positive Lebesgue-measure such that $\mathcal{F}_g(\omega) \neq 0$ for all $\omega \in B$.*

Proof. Let $\delta \in \mathbb{R}^2$ such that $g(x) = g(x - \delta)$ for all $x \in \mathbb{R}^2$. Then,

$$\begin{aligned} \mathcal{F}_g(\omega) &= \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, x \rangle} g(x) dx = \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, x \rangle} g(x - \delta) dx \\ &= \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, y + \delta \rangle} g(y) dy = e^{-2\pi i \langle \omega, \delta \rangle} \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, y \rangle} g(y) dy = e^{-2\pi i \langle \omega, \delta \rangle} \mathcal{F}_g(\omega), \end{aligned}$$

with the substitution $y := x - \delta$. For $\mathcal{F}_g(\omega) \neq 0$, this implies that $e^{-2\pi i \langle \omega, \delta \rangle} = 1$ and thus $\langle \omega, \delta \rangle \in \mathbb{Z}$. If this holds for all ω in an open Borel set $B \subseteq \mathbb{R}^2$, we can fix $\omega = (\omega_1, \omega_2) \in B$ and choose $\omega_1^{(1)}, \omega_1^{(2)} \in \mathbb{R} \setminus \{\omega_1\}$ and $\omega_2^{(1)}, \omega_2^{(2)} \in \mathbb{R} \setminus \{\omega_2\}$ with $(\omega_1^{(1)}, \omega_2), (\omega_1^{(2)}, \omega_2), (\omega_1, \omega_2^{(1)}), (\omega_1, \omega_2^{(2)}) \in B$ and

$$\omega_1 - \omega_1^{(1)}, \omega_2 - \omega_2^{(1)} \in \mathbb{Q}, \quad \omega_1 - \omega_1^{(2)}, \omega_2 - \omega_2^{(2)} \in \mathbb{R} \setminus \mathbb{Q}. \quad (6.47)$$

Hence,

$$\begin{aligned} (\omega_1 - \omega_1^{(1)})\delta_1 &= \langle \omega, \delta \rangle - \langle (\omega_1^{(1)}, \omega_2), \delta \rangle \in \mathbb{Z}, \\ (\omega_1 - \omega_1^{(2)})\delta_1 &= \langle \omega, \delta \rangle - \langle (\omega_1^{(2)}, \omega_2), \delta \rangle \in \mathbb{Z}, \\ (\omega_2 - \omega_2^{(1)})\delta_2 &= \langle \omega, \delta \rangle - \langle (\omega_1, \omega_2^{(1)}), \delta \rangle \in \mathbb{Z}, \\ (\omega_2 - \omega_2^{(2)})\delta_2 &= \langle \omega, \delta \rangle - \langle (\omega_1, \omega_2^{(2)}), \delta \rangle \in \mathbb{Z}. \end{aligned}$$

Because of (6.47), this means that $\delta = 0$, implying that g is not translation invariant.

On the other hand, if there is no open Borel set $B \subseteq \mathbb{R}^2$ with positive Lebesgue-measure such that $\mathcal{F}_g(\omega) \neq 0$ for all $\omega \in B$, then $\mathcal{F}_g = 0$ a.e. Applying the inverse Fourier transform, we get

$$g(x) = \int_{\mathbb{R}^2} e^{2\pi i \langle x, \omega \rangle} \mathcal{F}_g(\omega) d\omega = 0 \quad \text{for all } x \in \mathbb{R}^2.$$

Hence, g is constant. In particular, $g(x) = g(x - \delta)$ for all $x, \delta \in \mathbb{R}^2$. □

Assumption 6.28. Let $f \in L^2(\mathbb{R}^2) \cap H^{1/2}(\mathbb{R}^2)$ and f not translation invariant.

Assumption 6.29. The map

$$\Theta \rightarrow L^1([0, 1], \mathbb{R}^2), \quad \theta \mapsto \left(\delta^\theta : t \mapsto \delta_t^\theta = ((\delta_t^\theta)_1, (\delta_t^\theta)_2) \right),$$

is continuous with respect to the component-wise L^1 -norm

$$\|\theta \mapsto \delta^\theta\|_{L^1} := \|\theta \mapsto (\delta^\theta)_1\|_{L^1} + \|\theta \mapsto (\delta^\theta)_2\|_{L^1}.$$

Furthermore, for each $\theta \in \Theta$, the drift function $t \mapsto \delta_t^\theta$ is continuous at $t = 0$.

Assumption 6.30. The set of drift functions $\{t \mapsto \delta_t^\theta | \theta \in \Theta\}$ is identifiable.

Assumption 6.31. There is an open neighborhood $U_\delta \subseteq \Theta$ of θ_0 and a constant $L_\delta > 0$ such that the following local uniform Lipschitz condition holds,

$$\sup_{t \in [0, 1]} \|\delta_t^\theta - \delta_t^{\theta_0}\| \leq L_\delta \|\theta - \theta_0\| \quad \text{for all } \theta \in U_\delta.$$

Assumption 6.32. There is a $C > 0$ such that, uniformly in θ ,

$$\text{TV}(t \mapsto (\delta_t^\theta)_1) + \text{TV}(t \mapsto (\delta_t^\theta)_2) < C.$$

Theorem 6.33 (Consistency of drift parameter). Suppose that the Assumptions 6.2, 6.7, 6.26, 6.28, 6.29, 6.30, 6.31, and 6.32 hold. If $r_T, \beta_T \xrightarrow{T \rightarrow \infty} \infty$, $\beta_T = o(T)$, $r_T^4 T^{-1} \xrightarrow{T \rightarrow \infty} 0$, $r_T^2 \beta_T^{-1} \xrightarrow{T \rightarrow \infty} 0$, and $\sqrt{T}(\hat{\phi}_T - \phi_0, \hat{\alpha}_T - \alpha_0)$ is asymptotically centred normal, then the drift estimator $\hat{\theta}_T$ from Definition 4.8 is consistent, that is,

$$\hat{\theta}_T \xrightarrow{T \rightarrow \infty} \theta_0 \text{ in probability.} \quad (6.48)$$

Additionally, the image estimator \hat{f}_T from Definition 4.9 is consistent, that is,

$$\|\hat{f}_T - f\|_{L^2} \xrightarrow{T \rightarrow \infty} 0 \text{ in probability.} \quad (6.49)$$

Proof. The proof follows the standard three step argument which we used in the proof of Theorem 6.13. The three steps are:

1. Show the uniqueness of the population contrast minimizer θ_0 which, together with the compactness of Θ , yields (B.2) in the Appendix.
2. Show the continuity of the population contrast functional.
3. Show that $\tilde{N}_T \xrightarrow{T \rightarrow \infty} \tilde{N}$ in probability uniformly over θ , implying (B.1) in the Appendix.

Note, that (B.3) in the Appendix holds because $\hat{\theta}_T$ is defined as a minimizer of \tilde{N}_T . Then, Theorem B.6 in the Appendix yields the desired consistency of $\hat{\theta}_T$.

Step 1: uniqueness of population contrast minimizer θ_0 . We have

$$\begin{aligned}\tilde{N}(\theta) &= - \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \left| \int_0^1 h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right) dt \right|^2 d\omega \\ &\geq - \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \left(\int_0^1 \left| h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right) \right| dt \right)^2 d\omega = - \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 d\omega\end{aligned}$$

for all $\theta \in \Theta$ with equality if $\theta = \theta_0$. Let $\theta \in \Theta$ such that

$$\tilde{N}(\theta) = - \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 d\omega. \quad (6.50)$$

Since f is not translation invariant by Assumption 6.28, by Lemma 6.27, there is an open Borel set $B \subseteq \mathbb{R}^2$ with positive Lebesgue measure such that $\mathcal{F}_f(\omega) \neq 0$ for all $\omega \in B$. Since (6.50) holds, we have

$$\left| \int_0^1 h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right) dt \right| = 1 \quad \text{for all } \omega \in B.$$

By Lemma A.11 in the Appendix, $t \mapsto h_\omega(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}))$ is constant almost everywhere on $[0, 1]$. Because of the identifiability constraint $\delta_0^\theta = 0$ for all $\theta \in \Theta$ and the continuity of δ^θ at $t = 0$ (Assumption 6.29), this constant has to be 1 which implies that

$$\begin{aligned}1 &= h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right) \\ &= \exp \left(2\pi i \left\langle \omega, 1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right\rangle \right) \quad \text{for all } \omega \in B,\end{aligned} \quad (6.51)$$

almost everywhere on $[0, 1]$. Fix $\omega = (\omega_1, \omega_2) \in B$ and choose $\omega_1^{(1)}, \omega_1^{(2)} \in \mathbb{R} \setminus \{\omega_1\}$ and $\omega_2^{(1)}, \omega_2^{(2)} \in \mathbb{R} \setminus \{\omega_2\}$ such that $(\omega_1^{(1)}, \omega_2), (\omega_1^{(2)}, \omega_2), (\omega_1, \omega_2^{(1)}), (\omega_1, \omega_2^{(2)}) \in B$ and

$$\omega_1 - \omega_1^{(1)}, \omega_2 - \omega_2^{(1)} \in \mathbb{Q}, \quad \omega_1 - \omega_1^{(2)}, \omega_2 - \omega_2^{(2)} \in \mathbb{R} \setminus \mathbb{Q}. \quad (6.52)$$

From (6.51),

$$\begin{aligned}1 &= \frac{\exp \left(2\pi i \left\langle \omega, 1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right\rangle \right)}{\exp \left(2\pi i \left\langle (\omega_1^{(1)}, \omega_2)^\top, 1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right\rangle \right)} \\ &= \exp \left[2\pi i (\omega_1 - \omega_1^{(1)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right)_1 \right]\end{aligned}$$

and similarly,

$$\begin{aligned}1 &= \exp \left[2\pi i (\omega_1 - \omega_1^{(2)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right)_1 \right], \\ 1 &= \exp \left[2\pi i (\omega_2 - \omega_2^{(1)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right)_2 \right], \\ 1 &= \exp \left[2\pi i (\omega_2 - \omega_2^{(2)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right)_2 \right],\end{aligned}$$

which implies that

$$\begin{aligned}(\omega_1 - \omega_1^{(1)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right)_1 &\in \mathbb{Z}, \\ (\omega_1 - \omega_1^{(2)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right)_1 &\in \mathbb{Z}, \\ (\omega_2 - \omega_2^{(1)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right)_2 &\in \mathbb{Z}, \\ (\omega_2 - \omega_2^{(2)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right)_2 &\in \mathbb{Z}.\end{aligned}$$

Because of (6.52), we get that

$$1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) = 0 \quad \text{a.e. on } [0, 1]. \quad (6.53)$$

Since $R_{-\rho_t^{\phi_0}}$ is a rotation matrix,

$$\det \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}} \right) = (\sigma_t^{\alpha_0})^{-2} \geq \sigma_{\max}^{-2} > 0 \quad \text{for all } t \in [0, 1],$$

which, together with (6.53), implies that $\delta_t^\theta - \delta_t^{\theta_0} = 0$ a.e. on $[0, 1]$. Since the drift function is identifiable (Assumption 6.30), we conclude that $\theta = \theta_0$.

Step 2: Continuity of \tilde{N} . By Assumptions 6.7 and 6.29, the functions $t \mapsto \rho_t^{\phi_0}$, $t \mapsto \sigma_t^{\alpha_0}$ and $t \mapsto \delta_t^\theta$ are measurable for all $\theta \in \Theta$. Hence, $t \mapsto h_\omega(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}))$ is measurable for all $\theta \in \Theta$, $\omega \in \mathbb{R}^2$, as a concatenation of measurable functions. Also by Assumption 6.29, the functions $\theta \mapsto \delta_t^\theta$ are continuous for all $t \in [0, 1]$, giving that $\theta \mapsto h_\omega(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}))$ is continuous for all $t \in [0, 1]$, $\omega \in \mathbb{R}^2$, as a concatenation of continuous functions. Furthermore, $t \mapsto 1$ is an integrable majorant for $t \mapsto h_\omega(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}))$. Consequently, Theorem B.1 in the Appendix yields that

$$\theta \mapsto \int_0^1 h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right) dt$$

is continuous for all $\omega \in \mathbb{R}^2$. Since $x \mapsto |x|^2$ is continuous, we get that

$$\theta \mapsto g_\omega^\theta := \left| \int_0^1 h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right) dt \right|^2$$

is continuous for all $\omega \in \mathbb{R}^2$. By the same argument, the continuity of the functions $\omega \mapsto h_\omega(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}))$ for all $t \in [0, 1]$ and $\theta \in \Theta$ implies that $\omega \mapsto g_\omega^\theta$ is continuous, too, and hence Lebesgue-measurable. Since $|\mathcal{F}_f(\omega)|^2$ is constant in θ and continuous in ω (Lemma A.12 in the Appendix), $|\mathcal{F}_f(\omega)|^2 g_\omega^\theta$ is also continuous in ω as well as in θ as a concatenation of continuous functions. In particular, $\omega \mapsto |\mathcal{F}_f(\omega)|^2 g_\omega^\theta$ is Lebesgue-measurable. Furthermore, the function $\omega \mapsto |\mathcal{F}_f(\omega)|^2$ is an integrable majorant for $\omega \mapsto |\mathcal{F}_f(\omega)|^2 g_\omega^\theta$ because of Assumption 6.28 and $g_\omega^\theta \leq 1$. Hence, we can apply Theorem B.1 in the Appendix again to get the continuity of \tilde{N} .

Step 3: $\tilde{N}_T \xrightarrow{T \rightarrow \infty} \tilde{N}$ in probability uniformly in θ . From (4.4), we get

$$\begin{aligned} \tilde{N}_T(\theta) &= - \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^\theta \right) Z_T^t(\omega) \right|^2 d\omega \\ &= - \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right. \\ &\quad \left. + h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^\theta \right) V_T^t(\omega) \right|^2 d\omega \\ &= A_T(\theta) + B_T(\theta) + C_T(\theta), \end{aligned} \quad (6.54)$$

with

$$\begin{aligned}
A_T(\theta) &:= - \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \right. \\
&\quad \cdot \left. \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right|^2 d\omega, \\
B_T(\theta) &:= - \int_{\Omega_T} 2\Re \left[\left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \right) \right. \\
&\quad \cdot \left. \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right) \\
&\quad \cdot \left. \left(\frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} h_\omega \left(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^\theta \right) V_T^{t'}(\omega) \right) \right] d\omega, \\
C_T(\theta) &:= - \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^\theta \right) V_T^t(\omega) \right|^2 d\omega,
\end{aligned}$$

where we used Lemma A.1 in the Appendix. To derive the desired uniform convergence, we will show that $A_T \xrightarrow{T \rightarrow \infty} \tilde{N}$ in probability uniformly in θ while $B_T + C_T \xrightarrow{T \rightarrow \infty} 0$ in probability uniformly in θ . Because of Assumption 6.28, we have that

$$\int_{\mathbb{R}^2 \setminus \Omega_T} |\mathcal{F}_f(\omega)|^2 \left| \int_0^1 h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) dt \right|^2 d\omega \leq \int_{\mathbb{R}^2 \setminus \Omega_T} |\mathcal{F}_f(\omega)|^2 d\omega \xrightarrow{T \rightarrow \infty} 0.$$

Hence,

$$\begin{aligned}
& \left| A_T(\theta) - \tilde{N}(\theta) \right| \\
&= \left| \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right|^2 \right. \\
&\quad \left. - \left| \int_0^1 h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \mathcal{F}_f(\omega) dt \right|^2 d\omega \right. \\
&\quad \left. - \int_{\mathbb{R}^2 \setminus \Omega_T} |\mathcal{F}_f(\omega)|^2 \left| \int_0^1 h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) dt \right|^2 d\omega \right| \\
&= \left| \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right|^2 \right. \\
&\quad \left. - \left| \int_0^1 h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \mathcal{F}_f(\omega) dt \right|^2 d\omega \right| + o(1).
\end{aligned}$$

With Lemma A.4 in the Appendix, it follows that

$$\begin{aligned}
& \left| A_T(\theta) - \tilde{N}(\theta) \right| \\
&\leq \int_{\Omega_T} 2 \|\mathcal{F}_f\|_\infty \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^2 \left| \int_0^1 h_\omega \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \mathcal{F}_f(\omega) dt \right. \\
&\quad \left. - \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right| d\omega + o(1)
\end{aligned}$$

$$\begin{aligned}
&= 2 \|\mathcal{F}_f\|_\infty \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^2 \int_{\Omega_T} \left| h_\omega \left(1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}) \right) \mathcal{F}_f(\omega) \right| dt' \\
&\quad - \sum_{t \in \mathbb{T}} \int_0^{\beta_T/T} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \Big| d\omega \\
&\quad + o(1) \\
&\leq 2 \|\mathcal{F}_f\|_\infty \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^2 \int_{\Omega_T} \sum_{t \in \mathbb{T}} \int_0^{\beta_T/T} \left| h_\omega \left(1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}) \right) \mathcal{F}_f(\omega) \right. \\
&\quad \left. - h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \mathcal{F}_f(\omega) \right| + \left| h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \mathcal{F}_f(\omega) \right. \\
&\quad \left. - h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right| dt' d\omega + o(1) \\
&= 2 \|\mathcal{F}_f\|_\infty \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^2 \int_{\Omega_T} \sum_{t \in \mathbb{T}} \int_0^{\beta_T/T} \left| h_\omega \left(1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}) \right) \right. \\
&\quad \left. - h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \right| |\mathcal{F}_f(\omega)| \\
&\quad + \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) - \mathcal{F}_f(\omega) \right| dt' d\omega + o(1). \tag{6.55}
\end{aligned}$$

First, we consider the first part of (6.55). With Lemma A.8 in the Appendix, we get

$$\begin{aligned}
&\int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \int_0^{\beta_T/T} \left| h_\omega \left(1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}) \right) \right. \\
&\quad \left. - h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \right| |\mathcal{F}_f(\omega)| dt' d\omega \\
&\leq 2^{3/2} \pi \int_{\Omega_T} \|\omega\| |\mathcal{F}_f(\omega)| d\omega \\
&\quad \cdot \sum_{t \in \mathbb{T}} \int_0^{\beta_T/T} \left\| 1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}) - 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right\| dt' \\
&\leq 2^{3/2} \pi \int_{\Omega_T} \|\omega\| |\mathcal{F}_f(\omega)| d\omega \sum_{t \in \mathbb{T}} \int_0^{\beta_T/T} 1/\sigma_t^{\hat{\alpha}_T} \left\| (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}) - (\delta_t^\theta - \delta_t^{\theta_0}) \right\| \\
&\quad + \left\| \left(1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} - 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \right) (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}) \right\| dt', \tag{6.56}
\end{aligned}$$

where we used the fact that $\|R\delta\| = \|\delta\|$ for any rotation matrix $R \in \mathbb{R}^{2 \times 2}$ and any $\delta \in \mathbb{R}^2$. Since the total variation of the drift function is (uniformly over θ) bounded by $C > 0$ from Assumption 6.32, we have that

$$\begin{aligned}
&\sum_{t \in \mathbb{T}} \int_0^{\beta_T/T} 1/\sigma_t^{\hat{\alpha}_T} \left\| (\delta_t^\theta - \delta_t^{\theta_0}) - (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}) \right\| dt' \\
&\leq \sigma_{\min}^{-1} \int_0^{\beta_T/T} \sum_{t \in \mathbb{T}} \left\| \delta_t^\theta - \delta_{t+t'}^\theta \right\| + \left\| \delta_t^{\theta_0} - \delta_{t+t'}^{\theta_0} \right\| dt' \\
&\leq \sigma_{\min}^{-1} \int_0^{\beta_T/T} 2C dt' = \frac{2C\beta_T}{\sigma_{\min}T} = o(1), \tag{6.57}
\end{aligned}$$

because $\beta_T = o(T)$. Together with Assumption 6.28, it follows that the first part of (6.55) converges to zero in probability. For the second part of (6.55), we have

$$\begin{aligned}
& \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) - \mathcal{F}_f(\omega) \right| \\
& \leq \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left(\left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) - \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f(\omega) \right| \right. \\
& \quad \left. + \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f(\omega) - \mathcal{F}_f(\omega) \right| \right) \\
& = \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left(\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \left| \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) - \mathcal{F}_f(\omega) \right| + \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 - 1 \right| \mathcal{F}_f(\omega) \right) \\
& \leq \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^2 \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left(L_f \left\| \frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega - \omega \right\| + \left| (\sigma_t^{\alpha_0})^2 - (\sigma_t^{\hat{\alpha}_T})^2 \right| \|\mathcal{F}_f\|_{\infty} \right) \\
& = \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} \right), \tag{6.58}
\end{aligned}$$

where we used Lemma A.12 in the Appendix for the second inequality, Corollary 6.24, and the fact that $\sqrt{T}((\sigma_t^{\hat{\alpha}_T})^2 - (\sigma_t^{\alpha_0})^2)$ is asymptotically centred normal for $T \rightarrow \infty$ by the Delta-method (Theorem B.7 in the Appendix). Hence,

$$\int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) - \mathcal{F}_f(\omega) \right| d\omega = \mathcal{O}_{\mathbb{P}} \left(\frac{r_T^2}{\sqrt{T}} \right) = o_{\mathbb{P}}(1). \tag{6.59}$$

From (6.55), (6.57), and (6.59), it follows that

$$\left| A_T(\theta) - \tilde{N}(\theta) \right| \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability uniformly in } \theta. \tag{6.60}$$

Next, we show that $C_T \xrightarrow{T \rightarrow \infty} 0$ in probability uniformly in θ . With the Cauchy-Schwarz inequality and Lemma 5.4, we get

$$\begin{aligned}
0 & \geq \mathbb{E} C_T(\theta) \\
& = -\mathbb{E} \left(\int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_{\omega} \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^{\theta} \right) V_T^t(\omega) \right|^2 d\omega \right) \\
& \geq -\mathbb{E} \left(\int_{\Omega_T} \left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| h_{\omega} \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^{\theta} \right) \right|^2 \right) \left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} |V_T^t(\omega)|^2 \right) d\omega \right) \\
& = -\int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \mathbb{E} \left(|V_T^t(\omega)|^2 \right) d\omega = -\int_{\Omega_T} \mathcal{O} \left(\frac{1}{\beta_T} \right) d\omega = \mathcal{O} \left(\frac{r_T^2}{\beta_T} \right) = o(1).
\end{aligned}$$

Since $r_T^2/\beta_T \xrightarrow{T \rightarrow \infty} 0$, we get $\mathbb{E} C_T(\theta) \xrightarrow{T \rightarrow \infty} 0$ uniformly in θ , and thus

$$C_T \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability uniformly in } \theta. \tag{6.61}$$

Finally, the Cauchy-Schwarz inequality implies

$$(B_T(\theta))^2 \leq 4A_T(\theta)C_T(\theta) \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability uniformly in } \theta. \tag{6.62}$$

From (6.60), (6.61), and (6.62), we get $\tilde{N}_T \xrightarrow{T \rightarrow \infty} \tilde{N}$ in probability uniformly in θ , proving (6.48).

Now, we prove (6.49). Using the Plancherel equality (Theorem 2.12 in the Appendix) and (4.4), we have

$$\begin{aligned}
& \left\| \hat{f}_T - f \right\|_{L^2}^2 \\
&= \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^{\hat{\theta}_T} \right) Z_T(\omega) - \mathcal{F}_f(\omega) \right|^2 d\omega + \int_{\mathbb{R}^2 \setminus \Omega_T} |\mathcal{F}_f(\omega)|^2 d\omega \\
&= \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) - \mathcal{F}_f(\omega) \right. \\
&\quad \left. + \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^{\hat{\theta}_T} \right) V_T^t(\omega) \right|^2 d\omega + o(1) \\
&= D_T + E_T + C_T(\hat{\theta}_T) + o(1),
\end{aligned}$$

with

$$\begin{aligned}
D_T &:= \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right. \\
&\quad \left. - \mathcal{F}_f(\omega) \right|^2 d\omega, \\
E_T &:= \int_{\Omega_T} 2\Re \left[\left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right. \right. \\
&\quad \left. \left. - \mathcal{F}_f(\omega) \right) \overline{\left(\frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} h_\omega \left(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^{\hat{\theta}_T} \right) V_T^{t'}(\omega) \right)} \right] d\omega,
\end{aligned}$$

and $C_T(\hat{\theta}_T) \leq \sup_{\theta \in \Theta} C_T(\theta) \xrightarrow{T \rightarrow \infty} 0$ in probability as shown before. Because of (6.58), we have

$$\begin{aligned}
D_T &\leq \int_{\Omega_T} \left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right. \right. \\
&\quad \left. \left. - \mathcal{F}_f(\omega) \right| \right)^2 d\omega \\
&\leq \int_{\Omega_T} \left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega \right) \right. \right. \\
&\quad \left. \left. - h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0}) \right) \mathcal{F}_f(\omega) \right| \right. \\
&\quad \left. + \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0}) \right) \mathcal{F}_f(\omega) - \mathcal{F}_f(\omega) \right| \right)^2 d\omega
\end{aligned}$$

$$\begin{aligned}
&= \int_{\Omega_T} \left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T - \rho_t^{\phi_0}} \omega} \right) - \mathcal{F}_f(\omega) \right| \right. \\
&\quad \left. + |\mathcal{F}_f(\omega)| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left| h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}}(\delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0}) \right) - 1 \right| \right)^2 d\omega \\
&\leq \int_{\Omega_T} \left(\mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} \right) + 2\sqrt{2\pi}\sigma_{\min}^{-1} \|\omega\| |\mathcal{F}_f(\omega)| \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left\| \delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0} \right\| \right)^2 d\omega \\
&\leq \int_{\Omega_T} \left(\mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} \right) + 2\sqrt{2\pi}\sigma_{\min}^{-1} \|\omega\| |\mathcal{F}_f(\omega)| \mathcal{O}_{\mathbb{P}} \left(L_\delta \left\| \hat{\theta}_T - \theta_0 \right\| \right) \right)^2 d\omega \\
&= \mathcal{O}_{\mathbb{P}} \left(\frac{r_T^2}{T} + \frac{1}{\sqrt{T}} \left\| \hat{\theta}_T - \theta_0 \right\| \int_{\Omega_T} \|\omega\| |\mathcal{F}_f(\omega)| d\omega \right. \\
&\quad \left. + \left\| \hat{\theta}_T - \theta_0 \right\|^2 \int_{\Omega_T} \|\omega\|^2 |\mathcal{F}_f(\omega)|^2 d\omega \right),
\end{aligned}$$

where we used Lemma A.8 in the Appendix and the fact that rotations do not change the length of a vector for the third inequality and Assumption 6.31 for the fourth inequality. Because of $r_T^2/T \xrightarrow{T \rightarrow \infty} 0$, (6.48), and Assumption 6.28, it follows that $D_T \xrightarrow{T \rightarrow \infty} 0$ in probability.

Finally, by the Cauchy-Schwarz inequality,

$$E_T^2 \leq 4D_T C_T(\hat{\theta}_T) \xrightarrow{T \rightarrow \infty} 0$$

in probability, completing the proof of (6.49). \square

Remark 6.34 (Asymptotic orders of r_T , u_T , v_T , and β_T). *To get consistent estimators $(\hat{\phi}_T, \hat{\alpha}_T)$ and $\hat{\theta}_T$ for the rotation, scaling, and drift parameters, the Theorems 6.13, 6.23, and 6.33 suggest that we choose the cutoffs r_T , u_T , and v_T , and the bin size β_T , such that $r_T, u_T, v_T, \beta_T \xrightarrow{T \rightarrow \infty} \infty$ and*

$$\begin{aligned}
\beta_T &= o(T), \\
r_T^4 &= o(T), \\
\sqrt{T} r_T^{2+\gamma} &= o(\beta_T), \\
\sqrt{T} u_T v_T \|(u_T, v_T)\| r_T^{2\gamma} &= o(\beta_T).
\end{aligned}$$

In particular, β_T has to go to infinity slower than T , but faster than \sqrt{T} , much faster, in fact, if we do not want to throw away too much information by choosing small cutoffs r_T , u_T , and v_T . However, we will see in Chapter 10 that the estimation method still works fine if, for $T = 35,000$, we choose $\beta_T = 70$ and $r_T = u_T = v_T = 64$.

Theorem 6.35 (Uniform tightness of $(\text{grad}_\theta \tilde{N}_T(\theta_0))_{T \in \mathbb{N}}$). *Under the Assumptions 5.6, 6.2, and 6.28, if $r_T = o(T^{1/6})$, $\sqrt{T} r_T^3 / \beta_T \xrightarrow{T \rightarrow \infty} 0$, and $\sqrt{T}(\hat{\phi}_T - \phi_0, \hat{\alpha}_T - \alpha_0)$ is asymptotically centred normal, then $(\sqrt{T} \text{grad}_\theta \tilde{N}_T(\theta_0))_{T \in \mathbb{N}}$ is uniformly tight.*

Proof. We will use the decomposition $\tilde{N}_T = A_T + B_T + C_T$ from (6.54) and show that $\left\| \sqrt{T} \text{grad}_\theta A_T(\theta_0) \right\| = \mathcal{O}_{\mathbb{P}}(1)$, $\sqrt{T} \text{grad}_\theta C_T(\theta_0) \xrightarrow{T \rightarrow \infty} 0$ in probability, and $\sqrt{T} \text{grad}_\theta B_T(\theta_0)$

converges in distribution to a centred Gaussian random vector. First, consider

$$\omega_T^t := \frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega. \quad (6.63)$$

By Lemma A.12 in the Appendix and Corollary 6.24, we get

$$|\mathcal{F}_f(\omega_T^t) - \mathcal{F}_f(\omega)| \leq L_f \left\| \frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_t^{\hat{\phi}_T} - \rho_t^{\phi_0}} \omega - \omega \right\| = \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} \right), \quad (6.64)$$

with some Lipschitz-constant $L_f > 0$. Since $\Im[\mathcal{F}_f(\omega)\overline{\mathcal{F}_f(\omega)}] = \Im[|\mathcal{F}_f(\omega)|^2] = 0$, it follows that

$$\begin{aligned} \left| \Im \left[\mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right] \right| &= \left| \Im \left[(\mathcal{F}_f(\omega_T^t) - \mathcal{F}_f(\omega) + \mathcal{F}_f(\omega)) \overline{(\mathcal{F}_f(\omega_T^{t'}) - \mathcal{F}_f(\omega) + \mathcal{F}_f(\omega))} \right] \right| \\ &= \left| \Im \left[(\mathcal{F}_f(\omega_T^t) - \mathcal{F}_f(\omega)) \overline{(\mathcal{F}_f(\omega_T^{t'}) - \mathcal{F}_f(\omega))} \right. \right. \\ &\quad \left. \left. + (\mathcal{F}_f(\omega_T^t) - \mathcal{F}_f(\omega)) \overline{\mathcal{F}_f(\omega)} + \mathcal{F}_f(\omega) \overline{(\mathcal{F}_f(\omega_T^{t'}) - \mathcal{F}_f(\omega))} \right] \right| \\ &\leq \mathcal{O}_{\mathbb{P}} \left(\frac{1}{T} \right) + 2 |\mathcal{F}_f(\omega)| \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} \right). \end{aligned} \quad (6.65)$$

With Lemma A.14 in the Appendix, (5.25), and (6.65), we get for the gradient of the integrand of A_T , that

$$\begin{aligned} &\left\| \text{grad}_{\theta} \left[- \sum_{t,t' \in \mathbb{T}} h_{\omega} \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\theta} - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f(\omega_T^t) \right. \right. \\ &\quad \left. \left. \cdot h_{\omega} \left(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} (\delta_{t'}^{\theta} - \delta_{t'}^{\theta_0}) \right) \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f(\omega_T^{t'}) \right] \right\|_{\theta=\theta_0} \\ &= 2 \left\| \sum_{t,t' \in \mathbb{T}} \Re \left[\text{grad}_{\theta} h_{\omega} \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\theta} - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f(\omega_T^t) \right. \right. \\ &\quad \left. \left. \cdot h_{\omega} \left(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} (\delta_{t'}^{\theta} - \delta_{t'}^{\theta_0}) \right) \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f(\omega_T^{t'}) \right] \right\|_{\theta=\theta_0} \\ &= 2 \left\| \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \left(\mathbf{b}_{T,\omega}^{t,0}(\theta) \Im \left[\mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right. \right. \right. \\ &\quad \left. \left. \cdot h_{\omega} \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^{\theta} - \delta_t^{\theta_0}) \right) h_{\omega} \left(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} (\delta_{t'}^{\theta} - \delta_{t'}^{\theta_0}) \right) \right] \right) \right\|_{\theta=\theta_0} \\ &= 2 \left\| \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \Im \left[\mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right] \right\| \\ &\leq 2 \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \|\mathbf{b}_{T,\omega}^{t,0}(\theta_0)\| \left| \Im \left[\mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right] \right| \\ &\leq 2\tilde{C} \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^4 \|\omega\| \frac{T^2}{\beta_T^2} \left(\mathcal{O}_{\mathbb{P}} \left(\frac{1}{T} \right) + 2 |\mathcal{F}_f(\omega)| \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} \right) \right) \\ &= \mathcal{O}_{\mathbb{P}} \left(\|\omega\| \frac{T}{\beta_T^2} + \|\omega\| |\mathcal{F}_f(\omega)| \frac{T^{3/2}}{\beta_T^2} \right), \end{aligned} \quad (6.66)$$

with $\mathbf{b}_{T,\omega}^{t,0}(\theta)$ and $\tilde{C} > 0$ from Lemma 5.10. It follows that

$$\begin{aligned} \left\| \sqrt{T} \text{grad}_\theta A_T(\theta_0) \right\| &\leq \sqrt{T} \int_{\Omega_T} \frac{\beta_T^2}{T^2} \mathcal{O}_{\mathbb{P}} \left(\|\omega\| \frac{T}{\beta_T^2} + \|\omega\| |\mathcal{F}_f(\omega)| \frac{T^{3/2}}{\beta_T^2} \right) d\omega \\ &= \mathcal{O}_{\mathbb{P}} \left(\frac{r_T^3}{\sqrt{T}} + \int_{\Omega_T} \|\omega\| |\mathcal{F}_f(\omega)| d\omega \right). \end{aligned}$$

Since $r_T = o(T^{1/6})$ and Assumption 6.28 holds, we get that

$$\left\| \sqrt{T} \text{grad}_\theta A_T(\theta_0) \right\| = \mathcal{O}_{\mathbb{P}}(1). \quad (6.67)$$

Next, we consider the asymptotic behaviour of $\sqrt{T} \text{grad}_\theta C_T(\theta_0)$. By Lemma A.14 in the Appendix, Lemma 5.4, and (5.25), we get

$$\begin{aligned} &\mathbb{E} \left\| \text{grad}_\theta \left(- \sum_{t,t' \in \mathbb{T}} h_\omega(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^\theta) V_T^t(\omega) \overline{h_\omega(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^\theta) V_T^{t'}(\omega)} \right) \right\| \\ &= \mathbb{E} \left\| - \sum_{t,t' \in \mathbb{T}} 2\Re \left(\text{grad}_\theta h_\omega(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^\theta) V_T^t(\omega) \overline{h_\omega(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^\theta) V_T^{t'}(\omega)} \right) \right\| \\ &= 2\mathbb{E} \left\| \sum_{t,t' \in \mathbb{T}} \mathbf{b}_{T,\omega}^{t,0}(\theta) \Im \left(h_\omega(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^\theta - 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^\theta) V_T^t(\omega) \overline{V_T^{t'}(\omega)} \right) \right\| \\ &\leq 2\tilde{C} \|\omega\| \sum_{t,t' \in \mathbb{T}} \mathbb{E} \left| V_T^t(\omega) \overline{V_T^{t'}(\omega)} \right| = 2\tilde{C} \|\omega\| \sum_{t,t' \in \mathbb{T}} \mathcal{O} \left(\frac{1}{\beta_T} \right) = \mathcal{O} \left(\|\omega\| \frac{T^2}{\beta_T^3} \right), \quad (6.68) \end{aligned}$$

with $\mathbf{b}_{T,\omega}^{t,0}(\theta)$ and $\tilde{C} > 0$ from Lemma 5.10. It follows that

$$\begin{aligned} \mathbb{E} \left\| \sqrt{T} \text{grad}_\theta C_T(\theta_0) \right\| &= \mathbb{E} \left\| \sqrt{T} \int_{\Omega_T} \frac{\beta_T^2}{T^2} \text{grad}_\theta \left(- \sum_{t,t' \in \mathbb{T}} h_\omega(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} \delta_t^\theta) V_T^t(\omega) \right. \right. \\ &\quad \left. \left. \overline{h_\omega(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^\theta) V_T^{t'}(\omega)} \right) \Big|_{\theta=\theta_0} d\omega \right\| \\ &\leq \sqrt{T} \int_{\Omega_T} \frac{\beta_T^2}{T^2} \mathcal{O} \left(\|\omega\| \frac{T^2}{\beta_T^3} \right) d\omega = \mathcal{O} \left(\frac{\sqrt{T} r_T^3}{\beta_T} \right). \end{aligned}$$

Since $\sqrt{T} r_T^3 / \beta_T \xrightarrow{T \rightarrow \infty} 0$, we have

$$\sqrt{T} \text{grad}_\theta C_T(\theta_0) \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \quad (6.69)$$

It remains to show the asymptotic normality of $\sqrt{T} \text{grad}_\theta B_T(\theta_0)$. From Lemma 5.10, we get

$$\begin{aligned} \text{grad}_\theta B_T(\theta) &= \text{grad}_\theta \left\{ - \int_{\Omega_T} 2\Re \left[\left(\frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} h_\omega(-1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^\theta) \overline{V_T^{t'}(\omega)} \right) \right. \right. \\ &\quad \left. \left. \cdot \left(\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} (\delta_t^\theta - \delta_t^{\theta_0}) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f(\omega_T^t) \right) \right] d\omega \right\} \end{aligned}$$

from which it follows that

$$\begin{aligned}
\text{grad}_\theta B_T(\theta) &= \int_{\Omega_T} \frac{2\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \mathbf{b}_{T,\omega}^{t,t'}(\theta) \Im \left[\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f(\omega_T^t) \overline{V^{t'}(\omega)} \right. \\
&\quad \left. \cdot h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}}(\delta_t^\theta - \delta_t^{\theta_0}) - 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^\theta \right) \right] d\omega \\
&= \frac{2\sqrt{\beta_T}}{T} \sum_{t' \in \mathbb{T}} (\sigma_{t'}^{\hat{\alpha}_T})^{-2} \frac{1}{\sqrt{n_{t'}}} \sum_{j \in J_T^{t'}} \nu_j^{t'} \epsilon_j^{t'} \Im \left[\int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathbf{b}_{T,\omega}^{t,t'}(\theta) \right. \\
&\quad \left. \cdot h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}}(\delta_t^\theta - \delta_t^{\theta_0}) - 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^\theta \right) \mathcal{F}_f(\omega_T^t) \right. \\
&\quad \left. \cdot \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega, x_j \right\rangle \right) \right] d\omega. \tag{6.70}
\end{aligned}$$

Now, let

$$\begin{aligned}
G_j^{t'}(\omega) &:= (\sigma_{t'}^{\alpha_0})^{-2} \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\alpha_0} \cdot R_{\rho_{t'}^{\phi_0}} \omega, x_j \right\rangle \right) \mathcal{F}_f(\omega), \\
\widehat{G_j^{t'}(\omega)}_T^t &:= \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega, x_j \right\rangle \right) \mathcal{F}_f(\omega_T^t),
\end{aligned}$$

and $\xi \in \mathbb{R}^{d_1}$. From (6.70), we get that

$$\begin{aligned}
\left\langle \xi, \sqrt{T} \text{grad}_\theta B_T(\theta_0) \right\rangle &= \frac{2\sqrt{\beta_T}}{\sqrt{T}} \sum_{t' \in \mathbb{T}} \frac{1}{\sqrt{n_{t'}}} \sum_{j \in J_T^{t'}} \nu_j^{t'} \epsilon_j^{t'} \Im \left[\int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \widehat{G_j^{t'}(\omega)}_T^t \right. \\
&\quad \left. \cdot \left\langle \xi, \mathbf{b}_{T,\omega}^{t,t'}(\theta_0) \right\rangle h_\omega(-1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^{\theta_0}) d\omega \right] \\
&= D_T + E_T, \tag{6.71}
\end{aligned}$$

with

$$\begin{aligned}
D_T &:= \frac{2\sqrt{\beta_T}}{\sqrt{T}} \sum_{t' \in \mathbb{T}} \frac{1}{\sqrt{n_{t'}}} \sum_{j \in J_T^{t'}} \nu_j^{t'} \epsilon_j^{t'} \Im \left[\int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \left(\widehat{G_j^{t'}(\omega)}_T^t - G_j^{t'}(\omega) \right) \right. \\
&\quad \left. \cdot \left\langle \xi, \mathbf{b}_{T,\omega}^{t,t'}(\theta_0) \right\rangle h_\omega(-1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^{\theta_0}) d\omega \right], \\
E_T &:= \frac{2\sqrt{\beta_T}}{\sqrt{T}} \sum_{t' \in \mathbb{T}} \frac{1}{\sqrt{n_{t'}}} \sum_{j \in J_T^{t'}} \nu_j^{t'} \epsilon_j^{t'} \Im \left[\int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} G_j^{t'}(\omega) \right. \\
&\quad \left. \cdot \left\langle \xi, \mathbf{b}_{T,\omega}^{t,t'}(\theta_0) \right\rangle h_\omega(-1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^{\theta_0}) d\omega \right].
\end{aligned}$$

We will show that $D_T \xrightarrow{T \rightarrow \infty} 0$ in probability and E_T converges in distribution to a centred

Gaussian random variable. Consider

$$\begin{aligned}
\left| \widehat{G_j^{t'}}(\omega)_T^t - G_j^{t'}(\omega) \right| &= \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega, x_j \right\rangle \right) \mathcal{F}_f(\omega_{T^t}^t) \right. \\
&\quad \left. - (\sigma_{t'}^{\alpha_0})^{-2} \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\alpha_0} \cdot R_{\rho_{t'}^{\phi_0}} \omega, x_j \right\rangle \right) \mathcal{F}_f(\omega) \right| \\
&\leq \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega, x_j \right\rangle \right) \mathcal{F}_f(\omega_{T^t}^t) \right. \\
&\quad \left. - (\sigma_{t'}^{\alpha_0})^{-2} \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\alpha_0} \cdot R_{\rho_{t'}^{\phi_0}} \omega, x_j \right\rangle \right) \mathcal{F}_f(\omega_{T^t}^t) \right| \\
&\quad + \left| (\sigma_{t'}^{\alpha_0})^{-2} \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\alpha_0} \cdot R_{\rho_{t'}^{\phi_0}} \omega, x_j \right\rangle \right) \mathcal{F}_f(\omega_{T^t}^t) \right. \\
&\quad \left. - (\sigma_{t'}^{\alpha_0})^{-2} \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\alpha_0} \cdot R_{\rho_{t'}^{\phi_0}} \omega, x_j \right\rangle \right) \mathcal{F}_f(\omega) \right| \\
&= \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega, x_j \right\rangle \right) \right. \\
&\quad \left. - (\sigma_{t'}^{\alpha_0})^{-2} \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\alpha_0} \cdot R_{\rho_{t'}^{\phi_0}} \omega, x_j \right\rangle \right) \right| \|\mathcal{F}_f\|_\infty \\
&\quad + \sigma_{\min}^{-2} \left| \mathcal{F}_f \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} R_{\rho_{t'}^{\hat{\phi}_T} - \rho_{t'}^{\phi_0}} \omega \right) - \mathcal{F}_f(\omega) \right|. \tag{6.72}
\end{aligned}$$

It is

$$\begin{aligned}
&\left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega, x_j \right\rangle \right) \right. \\
&\quad \left. - (\sigma_{t'}^{\alpha_0})^{-2} \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\alpha_0} \cdot R_{\rho_{t'}^{\phi_0}} \omega, x_j \right\rangle \right) \right| \\
&\leq \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 - \left(\frac{1}{\sigma_{t'}^{\alpha_0}} \right)^2 \right| \left| \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega, x_j \right\rangle \right) \right| \\
&\quad + (\sigma_{t'}^{\alpha_0})^{-2} \left| \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega, x_j \right\rangle \right) - \exp \left(2\pi i \left\langle 1/\sigma_{t'}^{\alpha_0} \cdot R_{\rho_{t'}^{\phi_0}} \omega, x_j \right\rangle \right) \right| \\
&\leq \frac{|(\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T})^2 - (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^2|}{\sigma_{\min}^6} + \frac{2^{3/2} \pi \|x_j\|}{\sigma_{\min}^2} \left| \frac{1}{\sigma_{t'}^{\hat{\alpha}_T}} R_{\rho_{t'}^{\hat{\phi}_T}} \omega - \frac{1}{\sigma_{t'}^{\alpha_0}} R_{\rho_{t'}^{\phi_0}} \omega \right| \\
&= \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} \right), \tag{6.73}
\end{aligned}$$

where we used the Delta-Method (Theorem B.7 in the Appendix) to see that the random variables $\sqrt{T}((\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T})^2 - (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^2)$ and $\sqrt{T}(1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{\rho_{t'}^{\hat{\phi}_T}} \omega - 1/\sigma_{t'}^{\alpha_0} \cdot R_{\rho_{t'}^{\phi_0}} \omega)$ are asymptotically centred normal. From (6.65), (6.72), and (6.73), it follows that

$$\left| \widehat{G_j^{t'}}(\omega)_T^t - G_j^{t'}(\omega) \right| = \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} \right).$$

Since $|\epsilon_j^{t'}| = \mathcal{O}_{\mathbb{P}}(1)$ for all $t' \in \mathbb{T}$, $j \in J_T^{t'}$, we get

$$\begin{aligned} |D_T| &\leq 4\tilde{C} \|\xi\| \frac{\sqrt{\beta_T}}{\sqrt{T}} \sum_{t' \in \mathbb{T}} \frac{1}{\sqrt{n_{t'}}} \sum_{j \in J_T^{t'}} \nu_j^{t'} \mathcal{O}_{\mathbb{P}}(1) \int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \mathcal{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{T}}\right) \|\omega\| \, d\omega \\ &\leq \frac{\sqrt{\beta_T}}{\sqrt{T}} \sum_{t' \in \mathbb{T}} \mathcal{O}_{\mathbb{P}}\left(\frac{r_T^3}{\sqrt{T}}\right) = \mathcal{O}_{\mathbb{P}}\left(\frac{r_T^3}{\sqrt{\beta_T}}\right). \end{aligned}$$

with $\tilde{C} > 0$ from Lemma 5.10. Since $r_T^6/\beta_T = o(\sqrt{T}r_T^3/\beta_T) \xrightarrow{T \rightarrow \infty} 0$, we get

$$D_T \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \quad (6.74)$$

Finally, E_T is a linear combination of independent standard-normal random variables $\epsilon_j^{t'}$ and therefore a centred Gaussian random variable with variance

$$\frac{4\beta_T}{T} \sum_{t' \in \mathbb{T}} \frac{1}{n_{t'}} \sum_{j \in J_T^{t'}} (\nu_j^{t'})^2 \mathfrak{S} \left[\int_{\Omega_T} \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} G_j^{t'}(\omega) \langle \xi, \mathbf{b}_{T,\omega}^{t,t'}(\theta_0) \rangle h_{\omega}(-1/\sigma_{t'}^{\hat{\alpha}_T} \cdot R_{-\rho_{t'}^{\hat{\phi}_T}} \delta_{t'}^{\theta_0}) \, d\omega \right]^2$$

converging to $\xi^\top \tilde{\Sigma} \xi$, where

$$\tilde{\Sigma} := \int_0^1 \frac{1}{n} \sum_{j=1}^n S_j^{t'} (S_j^{t'})^\top \, dt',$$

with $\bar{\nu}_j$ from Assumption 6.14, $\mathbf{b}_{\infty,\omega}^{t,t'}(\theta_0)$ from (5.23) and

$$S_j^{t'} := 2\bar{\nu}_j(t') \mathfrak{S} \left[\int_{\mathbb{R}^2} \int_0^1 G_j^{t'}(\omega) \mathbf{b}_{\infty,\omega}^{t,t'}(\theta_0) h_{\omega}(-1/\sigma_{t'}^{\alpha_0} \cdot R_{-\rho_{t'}^{\phi_0}} \delta_{t'}^{\theta_0}) \, dt \, d\omega \right].$$

By Lemma A.16 and Corollary B.5 (both in the Appendix),

$$\sqrt{T} \text{grad}_{\theta} B_T(\theta_0) \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, \tilde{\Sigma}) \quad \text{in distribution.} \quad (6.75)$$

From (6.67), (6.69), (6.75), and Slutsky's Lemma (Theorem B.9 in the Appendix), we deduce $\sqrt{T} \text{grad}_{\theta} \tilde{N}_T(\theta_0) = \mathcal{O}_{\mathbb{P}}(1)$, completing the proof. \square

Assumption 6.36. For all $m \in \{1, \dots, d_1\}$, there is a Borel set $B_m \subseteq [0, 1]$ of positive Lebesgue-measure, such that at least one of the functions

$$B_m \rightarrow \mathbb{R}, \quad t \mapsto \frac{\partial(\delta_t^\theta)_1}{\partial \theta_m} \Big|_{\theta=\theta_0}, \quad \text{and} \quad B_m \rightarrow \mathbb{R}, \quad t \mapsto \frac{\partial(\delta_t^\theta)_2}{\partial \theta_m} \Big|_{\theta=\theta_0},$$

is not constant, where $\delta_t^\theta = ((\delta_t^\theta)_1, (\delta_t^\theta)_2)$.

Example 6.37 (Polynomial drift). Suppose that we have a polynomial drift model, that is,

$$(\delta_t^\theta)_1 = \sum_{m=0}^{d_1} \theta_m t^m, \quad (\delta_t^\theta)_2 = \sum_{m=0}^{d_1'} \theta_{d_1'+1+m} t^m,$$

where $\tilde{\theta} = (\theta_0, \dots, \theta_{d_1})$ and $d'_1, d''_1 \in \mathbb{N}$ with $d'_1 + d''_1 = d_1$. In this case, Assumption 6.36 is an implication of the identifiability constraint $\delta_0^\theta = 0$ for all $\tilde{\theta} \in \Theta$ from Assumption 3.3, since it means that the intercepts θ_0 and θ_{d_1+1} are zero. Hence,

$$(\delta_t^\theta)_1 = \sum_{m=1}^{d'_1} \theta_m t^m, \quad (\delta_t^\theta)_2 = \sum_{m=1}^{d''_1} \theta_{d'_1+1+m} t^m,$$

with $\theta = (\theta_1, \dots, \theta_{d_1})$. Then, we have $\partial(\delta_t^\theta)_1/\partial\theta_m = t^m$ for all $m \in \{1, \dots, d'_1\}$ and $\partial(\delta_t^\theta)_2/\partial\theta_m = t^{m-d'_1}$ for all $m \in \{d'_1+1, \dots, d_1\}$. Since those functions are not constant in t , Assumption 6.36 holds.

Assumption 6.38. We assume that $f \in H^1(\mathbb{R}^2)$.

Lemma 6.39. Under the Assumptions 5.9 and 6.38, $\text{Hess}_\theta \tilde{N}(\theta)$ has finite operator norm for all $\theta \in U'$ with $U' \subseteq \Theta$ from Assumption 5.9. Moreover, the matrix

$$H_N := \text{Hess}_\theta \tilde{N}(\theta_0) \tag{6.76}$$

is symmetric. If the Assumptions 6.28 and 6.36 hold, H_M is also positive definite and thus invertible.

Proof. From Lemma 5.10, Lemma A.15, and Theorem B.2 in the Appendix, with

$$h_{\omega, T, t}^\theta := h_\omega \left(1/\sigma_t^{\hat{\alpha}T} \cdot R_{-\rho_t^{\hat{\phi}T}}(\delta_t^\theta - \delta_t^{\theta_0}) \right), \tag{6.77}$$

we get

$$\begin{aligned} \text{Hess}_\theta \tilde{N}(\theta) &= -\text{Hess}_\theta \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \int_0^1 \int_0^1 h_{\omega, T, t}^\theta \overline{h_{\omega, T, t'}^\theta} dt dt' d\omega \\ &= -2 \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \int_0^1 \int_0^1 \Re \left(\overline{h_{\omega, T, t'}^\theta} \text{Hess}_\theta h_{\omega, T, t}^\theta \right. \\ &\quad \left. + \text{grad}_\theta h_{\omega, T, t}^\theta \overline{\text{grad}_\theta^\top h_{\omega, T, t'}^\theta} \right) dt dt' d\omega \\ &= -2 \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \int_0^1 \int_0^1 \Re \left(h_{\omega, T, t}^\theta \overline{h_{\omega, T, t'}^\theta} \right. \\ &\quad \left. \cdot \left[\mathbf{i} \mathbf{H}_{T, \omega}^{t, 0}(\theta) - \mathbf{b}_{T, \omega}^{t, 0}(\theta) (\mathbf{b}_{T, \omega}^{t, 0}(\theta) - \mathbf{b}_{T, \omega}^{t', 0}(\theta))^\top \right] \right) dt dt' d\omega. \end{aligned} \tag{6.78}$$

Let $\xi \in \mathbb{R}^{d_1}$ with $\|\xi\| = 1$. Then, because $f \in H^1(\mathbb{R}^2)$ by Assumption 6.38,

$$\begin{aligned} &\left\| \text{Hess}_\theta \tilde{N}(\theta) \xi \right\| \\ &\leq 2 \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \int_0^1 \int_0^1 \left\| \mathbf{H}_{T, \omega}^{t, 0}(\theta) \xi \right\| + \left\| \mathbf{b}_{T, \omega}^{t, 0}(\theta) (\mathbf{b}_{T, \omega}^{t, 0}(\theta) - \mathbf{b}_{T, \omega}^{t', 0}(\theta))^\top \xi \right\| dt dt' d\omega \\ &\leq 2 \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \int_0^1 \int_0^1 \left\| \mathbf{H}_{T, \omega}^{t, 0}(\theta) \right\|_1 + \left\| \mathbf{b}_{T, \omega}^{t, 0}(\theta) \right\| \left(\left\| \mathbf{b}_{T, \omega}^{t, 0}(\theta) \right\| + \left\| \mathbf{b}_{T, \omega}^{t', 0}(\theta) \right\| \right) dt dt' d\omega \\ &\leq 2 \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \left(\tilde{C} \|\omega\| + 2\tilde{C}^2 \|\omega\|^2 \right) d\omega < \infty, \end{aligned}$$

with $\tilde{C} > 0$ from Lemma 5.10. Thus, $\text{Hess}_\theta \tilde{N}(\theta)$ has finite operator norm. From (6.78), we have for $\theta = \theta_0$ that

$$\begin{aligned}
H_N &= \text{Hess}_\theta \tilde{N}(\theta_0) \\
&= 2 \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \int_0^1 \int_0^1 \mathbf{b}_{T,\omega}^{t,0}(\theta_0) (\mathbf{b}_{T,\omega}^{t,0}(\theta_0) - \mathbf{b}_{T,\omega}^{t',0}(\theta_0))^\top dt dt' d\omega \\
&= 2 \int_{\mathbb{R}^2} |\mathcal{F}_f(\omega)|^2 \left[\int_0^1 \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,0}(\theta_0)^\top dt \right. \\
&\quad \left. - \left(\int_0^1 \mathbf{b}_{T,\omega}^{t,0}(\theta_0) dt \right) \left(\int_0^1 \mathbf{b}_{T,\omega}^{t',0}(\theta_0) dt' \right)^\top \right] d\omega. \tag{6.79}
\end{aligned}$$

Since matrices of the form xx^\top with $x \in \mathbb{R}^{d_1}$ are symmetric, it follows that H_N is symmetric.

Now, let $\xi \in \mathbb{R}^{d_1} \setminus \{0\}$. By Assumption 6.28, there is an open Borel set $B \subseteq \mathbb{R}^2$ with positive Lebesgue-measure such that $\mathcal{F}_f(\omega) \neq 0$ for all $\omega \in B$. If the Assumptions 6.36 and 6.18 (implying that the rotation and scaling functions themselves do not vanish on a Borel set with positive Lebesgue-measure) hold, there is another Borel set $B' \subseteq [0, 1]$ with positive Lebesgue-measure such that

$$\begin{aligned}
B' \rightarrow \mathbb{R}, \quad t \mapsto \langle \xi, \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \rangle &= \frac{2\pi}{\sigma_t^{\hat{\alpha}_T}} \sum_{m=1}^{d_1} \xi_m \left(\left(\omega_1 \cos(\rho_t^{\hat{\phi}_T}) + \omega_2 \sin(\rho_t^{\hat{\phi}_T}) \right) \frac{\partial(\delta_t^\theta)_1}{\partial \theta_m} \right. \\
&\quad \left. + \left(\omega_2 \cos(\rho_t^{\hat{\phi}_T}) - \omega_1 \sin(\rho_t^{\hat{\phi}_T}) \right) \frac{\partial(\delta_t^\theta)_2}{\partial \theta_m} \right) \Big|_{\theta=\theta_0}
\end{aligned}$$

is not constant for all $\omega \in \mathbb{R}^2 \setminus \Omega^0$ with some Lebesgue null-set $\Omega^0 \subseteq \mathbb{R}^2$. Since the Cauchy-Schwarz inequality implies that $(\int_0^1 g(t) dt)^2 \leq \int_0^1 g(t)^2 dt$ for all integrable functions $g: [0, 1] \rightarrow \mathbb{R}$, with equality if and only if g is constant a.e., we get

$$\int_0^1 \langle \xi, \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \rangle^2 dt - \left(\int_0^1 \langle \xi, \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \rangle dt \right)^2 > 0.$$

Hence,

$$\xi^\top H_N \xi \geq 2 \int_B |\mathcal{F}_f(\omega)|^2 \left[\int_0^1 \langle \xi, \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \rangle^2 dt - \left(\int_0^1 \langle \xi, \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \rangle dt \right)^2 \right] d\omega > 0,$$

since the integrand (as a function in ω) is strictly positive on B . We conclude that H_N is symmetric and positive definite and as such invertible. \square

Assumption 6.40. We assume that $f \in H^2(\mathbb{R}^2)$.

Theorem 6.41. Under the Assumption 5.9, let $(\hat{\theta}_T^*)_{T \in \mathbb{N}}$ a sequence of random vectors with values in U' , such that $\hat{\theta}_T^* \xrightarrow{T \rightarrow \infty} \theta_0$ in probability. Suppose, that the Assumptions of Theorem 6.35 and the Assumptions 6.25 and 6.40 hold. Then, with H_N from (6.76),

$$\left\| \text{Hess}_\theta \tilde{N}_T(\hat{\theta}_T^*) - H_N \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.}$$

Proof. We will use the decomposition $\tilde{N}_T = A_T + B_T + C_T$ from (6.54) and show that

$$\begin{aligned} & \left\| \text{Hess}_\theta A_T(\hat{\theta}_T^*) - H_N \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability,} \\ & \sup_{\theta \in U'} \left\| \text{Hess}_\theta B_T(\theta) + \text{Hess}_\theta C_T(\theta) \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \end{aligned}$$

With Lemma 5.10, Lemma A.14 in the Appendix, ω_T^t from (6.63), and $h_{\omega, T, t}^\theta$ from 6.77, we get

$$\begin{aligned} & \text{Hess}_\theta A_T(\theta) \\ = & \text{Hess}_\theta \left[- \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 h_{\omega, T, t}^\theta \mathcal{F}_f(\omega_T^t) \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \overline{h_{\omega, T, t'}^\theta \mathcal{F}_f(\omega_T^{t'})} d\omega \right] \\ = & -2 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right. \\ & \cdot \left. \left[\text{Hess}_\theta h_{\omega, T, t}^\theta \overline{h_{\omega, T, t'}^\theta} + \text{grad}_\theta h_{\omega, T, t}^\theta \overline{\text{grad}_\theta^\top h_{\omega, T, t'}^\theta} \right] \right\} d\omega \\ = & -2 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right. \\ & \cdot h_{\omega, T, t}^\theta \overline{h_{\omega, T, t'}^\theta} \left[\mathbf{iH}_{T, \omega}^{t, 0}(\theta) - \mathbf{b}_{T, \omega}^{t, 0}(\theta) \mathbf{b}_{T, \omega}^{t, 0}(\theta)^\top + \mathbf{b}_{T, \omega}^{t, 0}(\theta) \mathbf{b}_{T, \omega}^{t', 0}(\theta)^\top \right] \left. \right\} d\omega \\ = & -2 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right. \\ & \cdot h_{\omega, T, t}^\theta \overline{h_{\omega, T, t'}^\theta} \left[\mathbf{iH}_{T, \omega}^{t, 0}(\theta) - \mathbf{b}_{T, \omega}^{t, 0}(\theta) \mathbf{b}_{T, \omega}^{t', 0}(\theta)^\top \right] \left. \right\} d\omega. \end{aligned} \quad (6.80)$$

where we used that $\mathbf{b}_{T, \omega}^{t, 0}(\theta_0) - \mathbf{b}_{T, \omega}^{t', 0}(\theta_0) = \mathbf{b}_{T, \omega}^{t, t'}(\theta_0)$. From (6.64), we get

$$|\mathcal{F}_f(\omega_T^t)| \leq |\mathcal{F}_f(\omega_T^t) - \mathcal{F}_f(\omega)| + |\mathcal{F}_f(\omega)| = \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} + |\mathcal{F}_f(\omega)| \right). \quad (6.81)$$

Moreover, the continuity of $\theta \mapsto \delta_t^\theta$ (Assumption 6.29), the Lipschitz-continuity of h_ω (Lemma A.8 in the Appendix), and the continuous mapping theorem (Theorem B.8 in the Appendix) yield that

$$\begin{aligned} \left| h_{\omega, T, t}^{\hat{\theta}_T^*} - 1 \right| &= \left| h_\omega \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}}(\delta_t^{\hat{\theta}_T^*} - \delta_t^{\theta_0}) \right) - 1 \right| \\ &= \left| \exp \left(2\pi i \left\langle \omega, 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}}(\delta_t^{\hat{\theta}_T^*} - \delta_t^{\theta_0}) \right\rangle \right) - 1 \right| \\ &= \left| \exp \left(2\pi i \left\langle 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{\rho_t^{\hat{\phi}_T}} \omega, \delta_t^{\hat{\theta}_T^*} - \delta_t^{\theta_0} \right\rangle \right) - 1 \right| \\ &\leq 2\pi\sqrt{2} \left\| 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{\rho_t^{\hat{\phi}_T}} \omega \right\| \left\| \delta_t^{\hat{\theta}_T^*} - \delta_t^{\theta_0} \right\| \\ &\leq \frac{2\pi\sqrt{2}}{\sigma_{\min}} \|\omega\| \left\| \delta_t^{\hat{\theta}_T^*} - \delta_t^{\theta_0} \right\| = o_{\mathbb{P}}(\|\omega\|) \end{aligned}$$

and thus

$$h_{\omega, T, t}^{\hat{\theta}_T^*} \overline{h_{\omega, T, t'}^{\hat{\theta}_T^*}} = 1 + o_{\mathbb{P}}(\|\omega\|^2). \quad (6.82)$$

In particular, the imaginary part converges to 0. With (6.65), (6.81), and (6.82), we have

$$\begin{aligned}
& \left\| \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \cdot h_{\omega,T,t}^{\hat{\theta}_T^*} \overline{h_{\omega,T,t'}^{\hat{\theta}_T^*}} \cdot \mathbf{iH}_{T,\omega}^{t,0}(\hat{\theta}_T^*) \right\} \right\|_1 \\
&= \left| \Im \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \cdot h_{\omega,T,t}^{\hat{\theta}_T^*} \overline{h_{\omega,T,t'}^{\hat{\theta}_T^*}} \right\} \right| \left\| \mathbf{H}_{T,\omega}^{t,0}(\hat{\theta}_T^*) \right\|_1 \\
&= \left| \Im \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right\} \Re \left\{ h_{\omega,T,t}^{\hat{\theta}_T^*} \overline{h_{\omega,T,t'}^{\hat{\theta}_T^*}} \right\} \right. \\
&\quad \left. + \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right\} \Im \left\{ h_{\omega,T,t}^{\hat{\theta}_T^*} \overline{h_{\omega,T,t'}^{\hat{\theta}_T^*}} \right\} \right| \left\| \mathbf{H}_{T,\omega}^{t,0}(\hat{\theta}_T^*) \right\|_1 \\
&\leq \left| \Im \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right\} \right| \left\| \mathbf{H}_{T,\omega}^{t,0}(\hat{\theta}_T^*) \right\|_1 \\
&\quad + \left| \mathcal{F}_f(\omega_T^t) \right| \left| \mathcal{F}_f(\omega_T^{t'}) \right| \left| \Im \left\{ h_{\omega,T,t}^{\hat{\theta}_T^*} \overline{h_{\omega,T,t'}^{\hat{\theta}_T^*}} \right\} \right| \left\| \mathbf{H}_{T,\omega}^{t,0}(\hat{\theta}_T^*) \right\|_1 \\
&= \mathcal{O}_{\mathbb{P}} \left(\frac{\|\omega\|}{T} + \frac{\|\omega\| |\mathcal{F}_f(\omega)|}{\sqrt{T}} \right) + o_{\mathbb{P}} \left(\|\omega\|^3 |\mathcal{F}_f(\omega)|^2 \right).
\end{aligned}$$

It follows that

$$\begin{aligned}
& \left\| \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \right. \\
&\quad \left. \cdot \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \cdot h_{\omega,T,t}^{\hat{\theta}_T^*} \overline{h_{\omega,T,t'}^{\hat{\theta}_T^*}} \cdot \mathbf{iH}_{T,\omega}^{t,0}(\hat{\theta}_T^*) \right\} d\omega \right\|_1 \\
&\leq \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^4 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \mathcal{O}_{\mathbb{P}} \left(\frac{\|\omega\|}{T} + \frac{\|\omega\| |\mathcal{F}_f(\omega)|}{\sqrt{T}} \right) + o_{\mathbb{P}} \left(\|\omega\|^3 |\mathcal{F}_f(\omega)|^2 \right) d\omega \\
&= \mathcal{O}_{\mathbb{P}} \left(\frac{r_T^3}{T} + \sqrt{\frac{r_T}{T}} \int_{\Omega_T} \|\omega\| |\mathcal{F}_f(\omega)|^2 d\omega \right) + o_{\mathbb{P}} \left(\int_{\Omega_T} \|\omega\|^3 |\mathcal{F}_f(\omega)|^2 d\omega \right) \\
&= o_{\mathbb{P}}(1), \tag{6.83}
\end{aligned}$$

because of $r_T^3/T \xrightarrow{T \rightarrow \infty} 0$ and Assumption 6.40. Similarly,

$$\begin{aligned}
\left\| \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \cdot \mathbf{iH}_{T,\omega}^{t,0}(\theta_0) \right\} \right\|_1 &= \left| \Im \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right\} \right| \left\| \mathbf{H}_{T,\omega}^{t,0}(\theta_0) \right\|_1 \\
&= \mathcal{O}_{\mathbb{P}} \left(\frac{\|\omega\|}{T} + \frac{\|\omega\| |\mathcal{F}_f(\omega)|}{\sqrt{T}} \right),
\end{aligned}$$

which implies that

$$\left\| \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \cdot \mathbf{iH}_{T,\omega}^{t,0}(\theta_0) \right\} d\omega \right\|_1 = o_{\mathbb{P}}(1). \tag{6.84}$$

Moreover, by Assumption 5.9 and the continuous mapping theorem (Theorem B.8 in the Appendix), we get that

$$\begin{aligned}
& \left\| \mathbf{b}_{T,\omega}^{t,0}(\hat{\theta}_T^*) \mathbf{b}_{T,\omega}^{t,t'}(\hat{\theta}_T^*)^\top - \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top \right\|_1 \\
&\leq \left\| \mathbf{b}_{T,\omega}^{t,0}(\hat{\theta}_T^*) \left[\mathbf{b}_{T,\omega}^{t,t'}(\hat{\theta}_T^*) - \mathbf{b}_{T,\omega}^{t,t'}(\theta_0) \right]^\top \right\|_1 + \left\| \left[\mathbf{b}_{T,\omega}^{t,0}(\hat{\theta}_T^*) - \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \right] \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top \right\|_1 \\
&\leq d_1 \tilde{C} \|\omega\| \left(\left\| \mathbf{b}_{T,\omega}^{t,t'}(\hat{\theta}_T^*) - \mathbf{b}_{T,\omega}^{t,t'}(\theta_0) \right\| + \left\| \mathbf{b}_{T,\omega}^{t,0}(\hat{\theta}_T^*) - \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \right\| \right)
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{2\pi d_1 \tilde{C}}{\sigma_{\min}} \|\omega\| \left(2 |\omega_1| \left| \cos(\rho_t^{\hat{\phi}^T}) \right| \left\| \text{grad}_\theta(\delta_t^\theta)_1|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_t^\theta)_1|_{\theta=\theta_0} \right\| \right. \\
&\quad + 2 |\omega_1| \left| \sin(\rho_t^{\hat{\phi}^T}) \right| \left\| \text{grad}_\theta(\delta_t^\theta)_2|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_t^\theta)_2|_{\theta=\theta_0} \right\| \\
&\quad + 2 |\omega_2| \left| \sin(-\rho_t^{\hat{\phi}^T}) \right| \left\| \text{grad}_\theta(\delta_t^\theta)_1|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_t^\theta)_1|_{\theta=\theta_0} \right\| \\
&\quad + 2 |\omega_2| \left| \cos(\rho_t^{\hat{\phi}^T}) \right| \left\| \text{grad}_\theta(\delta_t^\theta)_2|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_t^\theta)_2|_{\theta=\theta_0} \right\| \\
&\quad + |\omega_1| \left| \cos(\rho_{t'}^{\hat{\phi}^T}) \right| \left\| \text{grad}_\theta(\delta_{t'}^\theta)_1|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_{t'}^\theta)_1|_{\theta=\theta_0} \right\| \\
&\quad + |\omega_1| \left| \sin(\rho_{t'}^{\hat{\phi}^T}) \right| \left\| \text{grad}_\theta(\delta_{t'}^\theta)_2|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_{t'}^\theta)_2|_{\theta=\theta_0} \right\| \\
&\quad + |\omega_2| \left| \sin(-\rho_{t'}^{\hat{\phi}^T}) \right| \left\| \text{grad}_\theta(\delta_{t'}^\theta)_1|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_{t'}^\theta)_1|_{\theta=\theta_0} \right\| \\
&\quad \left. + |\omega_2| \left| \cos(\rho_{t'}^{\hat{\phi}^T}) \right| \left\| \text{grad}_\theta(\delta_{t'}^\theta)_2|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_{t'}^\theta)_2|_{\theta=\theta_0} \right\| \right) \\
&\leq \frac{2\pi d_1^{3/2} \tilde{C}}{\sigma_{\min}} \|\omega\|^2 \left(2 \left\| \text{grad}_\theta(\delta_t^\theta)_1|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_t^\theta)_1|_{\theta=\theta_0} \right\| \right. \\
&\quad + 2 \left\| \text{grad}_\theta(\delta_t^\theta)_2|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_t^\theta)_2|_{\theta=\theta_0} \right\| \\
&\quad + \left\| \text{grad}_\theta(\delta_{t'}^\theta)_1|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_{t'}^\theta)_1|_{\theta=\theta_0} \right\| \\
&\quad \left. + \left\| \text{grad}_\theta(\delta_{t'}^\theta)_2|_{\theta=\hat{\theta}_T^*} - \text{grad}_\theta(\delta_{t'}^\theta)_2|_{\theta=\theta_0} \right\| \right) \\
&= o_{\mathbb{P}}(\|\omega\|^2).
\end{aligned}$$

With (6.80), (6.81), (6.82), (6.83), and (6.84), it follows that

$$\begin{aligned}
&\left\| \text{Hess}_\theta A_T(\hat{\theta}_T^*) - \text{Hess}_\theta A_T(\theta_0) \right\|_1 \\
&= \left\| 2 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right. \right. \\
&\quad \left. \left. \cdot h_{\omega,T,t}^{\hat{\theta}_T^*} \overline{h_{\omega,T,t'}^{\hat{\theta}_T^*}} \left(\mathbf{b}_{T,\omega}^{t,0}(\hat{\theta}_T^*) \mathbf{b}_{T,\omega}^{t,t'}(\hat{\theta}_T^*)^\top - \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top \right) \right\} d\omega \right\|_1 + o_{\mathbb{P}}(1) \\
&\leq 2 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^4 |\mathcal{F}_f(\omega_T^t)| |\mathcal{F}_f(\omega_T^{t'})| o_{\mathbb{P}}(\|\omega\|^4) d\omega + o_{\mathbb{P}}(1) \\
&= o_{\mathbb{P}} \left(\int_{\Omega_T} \|\omega\|^4 \left(\frac{1}{T} + \frac{|\mathcal{F}_f(\omega)|}{\sqrt{T}} + |\mathcal{F}_f(\omega)|^2 \right) d\omega \right) \\
&= o_{\mathbb{P}} \left(\frac{r_T^6}{T} + \sqrt{\frac{r_T^4}{T}} \int_{\Omega_T} \|\omega\|^4 |\mathcal{F}_f(\omega)|^2 d\omega + \int_{\Omega_T} \|\omega\|^4 |\mathcal{F}_f(\omega)|^2 d\omega \right) \\
&= o_{\mathbb{P}}(1), \tag{6.85}
\end{aligned}$$

because of $r_T^6/T \xrightarrow{T \rightarrow \infty} 0$ and Assumption 6.40. By Lemma 6.39, H_N has finite operator norm. In particular, the components of H_N are finite. Hence, with (6.79),

$$\left\| 2 \int_{\mathbb{R}^2 \setminus \Omega_T} |\mathcal{F}_f(\omega)|^2 \int_0^1 \int_0^1 \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top dt dt' d\omega \right\|_1 = o(1),$$

as $T \rightarrow \infty$. With (6.80) and (6.85), it follows that

$$\begin{aligned}
& \left\| \text{Hess}_\theta A_T(\hat{\theta}_T^*) - H_N \right\|_1 \\
&= \left\| \text{Hess}_\theta A_T(\theta_0) - H_N \right\|_1 + o_{\mathbb{P}}(1) \\
&\leq 2 \int_{\Omega_T} \left\| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right\} \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top \right. \\
&\quad \left. - |\mathcal{F}_f(\omega)|^2 \int_0^1 \int_0^1 \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top dt dt' \right\|_1 d\omega + o_{\mathbb{P}}(1) \\
&\leq 2 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left\| \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top \right\|_1 \left| \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right\} - |\mathcal{F}_f(\omega)|^2 \right| \\
&\quad + |\mathcal{F}_f(\omega)|^2 \left\| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top - \int_0^1 \int_0^1 \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top dt dt' \right\|_1 d\omega \\
&\quad + o_{\mathbb{P}}(1). \tag{6.86}
\end{aligned}$$

By the continuous mapping theorem (Theorem B.8 in the Appendix), we have that

$$\left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \xrightarrow{T \rightarrow \infty} 1 \quad \text{in probability.}$$

Together with Lemma 5.10 and (6.81), this implies that

$$\begin{aligned}
& \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left\| \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t,t'}(\theta_0)^\top \right\|_1 \left| \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re \left\{ \mathcal{F}_f(\omega_T^t) \overline{\mathcal{F}_f(\omega_T^{t'})} \right\} - |\mathcal{F}_f(\omega)|^2 \right| d\omega \\
&\leq \int_{\Omega_T} \tilde{C}^2 \|\omega\|^2 \left| (1 + o_{\mathbb{P}}(1)) \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{T}} + |\mathcal{F}_f(\omega)| \right)^2 - |\mathcal{F}_f(\omega)|^2 \right| d\omega \\
&= \mathcal{O}_{\mathbb{P}} \left(\frac{r_T^2}{T} + \sqrt{\frac{r_T^2}{T} \int_{\Omega_T} \|\omega\|^2 |\mathcal{F}_f(\omega)|^2 d\omega} \right) + o_{\mathbb{P}} \left(\int_{\Omega_T} \|\omega\|^2 |\mathcal{F}_f(\omega)|^2 d\omega \right) \\
&= o_{\mathbb{P}}(1), \tag{6.87}
\end{aligned}$$

since $r_T^2/T \xrightarrow{T \rightarrow \infty} 0$ and Assumption 6.40 holds. Furthermore, by Lemma A.7 in the Appendix, Lemma A.9 in the Appendix, and Assumption 6.25, we get that

$$\begin{aligned}
& \left\| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t',0}(\theta_0)^\top - \int_0^1 \int_0^1 \mathbf{b}_{T,\omega}^{t,0}(\theta_0) \mathbf{b}_{T,\omega}^{t',0}(\theta_0)^\top dt dt' \right\|_1 \\
&= \sum_{m,m'=1}^{d_1} \left| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} (\mathbf{b}_{T,\omega}^{t,0}(\theta_0))_m (\mathbf{b}_{T,\omega}^{t',0}(\theta_0))_{m'} - \int_0^1 \int_0^1 (\mathbf{b}_{T,\omega}^{t,0}(\theta_0))_m (\mathbf{b}_{T,\omega}^{t',0}(\theta_0))_{m'} dt dt' \right| \\
&\leq \tilde{C} \|\omega\| \frac{\beta_T}{T} \sum_{m,m'=1}^{d_1} \left[\text{TV} \left(t \mapsto (\mathbf{b}_{T,\omega}^{t,0}(\theta_0))_m \right) + \text{TV} \left(t' \mapsto (\mathbf{b}_{T,\omega}^{t',0}(\theta_0))_{m'} \right) \right]
\end{aligned}$$

$$\begin{aligned}
&\leq 2d_1\tilde{C}\|\omega\|\frac{\beta_T}{T}\sum_{m=1}^{d_1}\left[|\omega_1|\mathrm{TV}\left(t\mapsto 1/\sigma_t^{\hat{\alpha}T}\left(\cos(\rho_t^{\hat{\phi}T})\frac{\partial(\delta_t^\theta)_1}{\partial\theta_m}+\sin(\rho_t^{\hat{\phi}T})\frac{\partial(\delta_t^\theta)_2}{\partial\theta_m}\right)\Big|_{\theta=\theta_0}\right) \right. \\
&\quad \left. +|\omega_2|\mathrm{TV}\left(t\mapsto 1/\sigma_t^{\hat{\alpha}T}\left(\sin(-\rho_t^{\hat{\phi}T})\frac{\partial(\delta_t^\theta)_1}{\partial\theta_m}+\cos(\rho_t^{\hat{\phi}T})\frac{\partial(\delta_t^\theta)_2}{\partial\theta_m}\right)\Big|_{\theta=\theta_0}\right)\right] \\
&= \mathcal{O}_{\mathbb{P}}\left(\|\omega\|^2\frac{\beta_T}{T}\right), \tag{6.88}
\end{aligned}$$

with $\tilde{C} > 0$ from Lemma 5.10, where the final line in (6.88) follows from Assumption 6.11 and Lemma A.5 in the Appendix. Similarly, with Lemma A.5 in the Appendix and Lemma A.6 in the Appendix,

$$\begin{aligned}
&\left\|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\mathbf{b}_{T,\omega}^{t,0}(\theta_0)\mathbf{b}_{T,\omega}^{t,0}(\theta_0)^\top - \int_0^1\mathbf{b}_{T,\omega}^{t,0}(\theta_0)\mathbf{b}_{T,\omega}^{t,0}(\theta_0)^\top dt\right\|_1 \\
&\leq \frac{\beta_T}{T}\sum_{m,m'=1}^{d_1}\mathrm{TV}\left(t\mapsto (\mathbf{b}_{T,\omega}^{t,0}(\theta_0))_m(\mathbf{b}_{T,\omega}^{t,0}(\theta_0))_{m'}\right) \\
&\leq 2d_1\tilde{C}\|\omega\|\frac{\beta_T}{T}\sum_{m=1}^{d_1}\left[|\omega_1|\mathrm{TV}\left(t\mapsto 1/\sigma_t^{\hat{\alpha}T}\left(\cos(\rho_t^{\hat{\phi}T})\frac{\partial(\delta_t^\theta)_1}{\partial\theta_m}+\sin(\rho_t^{\hat{\phi}T})\frac{\partial(\delta_t^\theta)_2}{\partial\theta_m}\right)\Big|_{\theta=\theta_0}\right) \right. \\
&\quad \left. +|\omega_2|\mathrm{TV}\left(t\mapsto 1/\sigma_t^{\hat{\alpha}T}\left(\sin(-\rho_t^{\hat{\phi}T})\frac{\partial(\delta_t^\theta)_1}{\partial\theta_m}+\cos(\rho_t^{\hat{\phi}T})\frac{\partial(\delta_t^\theta)_2}{\partial\theta_m}\right)\Big|_{\theta=\theta_0}\right)\right] \\
&= \mathcal{O}_{\mathbb{P}}\left(\|\omega\|^2\frac{\beta_T}{T}\right). \tag{6.89}
\end{aligned}$$

Collecting (6.86), (6.87), (6.88), and (6.89), it follows that

$$\left\|\mathrm{Hess}_\theta A_T(\hat{\theta}_T^*) - H_N\right\|_1 = \mathcal{O}_{\mathbb{P}}\left(\frac{\beta_T}{T}\int_{\Omega_T}\|\omega\|^2|\mathcal{F}_f(\omega)|^2 d\omega\right) = o_{\mathbb{P}}(1). \tag{6.90}$$

Next, we show that $\sup_{\theta\in U'}\|\mathrm{Hess}_\theta B_T(\theta)\|_1 \xrightarrow{T\rightarrow\infty} 0$ in probability. By Lemma 5.10,

$$\sup_{\theta\in U'}\left\|\mathrm{Hess}_\theta h_\omega\left(\tau_T^t(\delta_t^\theta - \delta_t^{\theta_0}) - \tau_T^{t'}\delta_{t'}^\theta\right)\right\|_1 = \mathcal{O}\left(\|\omega\|^2\right). \tag{6.91}$$

Because of (6.81), (6.91), and Lemma 5.4, we have

$$\begin{aligned}
\sup_{\theta\in U'}\|\mathrm{Hess}_\theta B_T(\theta)\|_1 &= \sup_{\theta\in U'}\left\|\int_{\Omega_T}2\Re\left[\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\mathrm{Hess}_\theta h_\omega\left(\tau_T^t(\delta_t^\theta - \delta_t^{\theta_0}) - \tau_T^{t'}\delta_{t'}^\theta\right) \right. \right. \\
&\quad \left. \left. \cdot\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}T}}\right)^2\mathcal{F}_f(\omega_t^t)\overline{V^{t'}(\omega)}\right]d\omega\right\|_1 \\
&\leq 2\sup_{\theta\in U'}\int_{\Omega_T}\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\left\|\mathrm{Hess}_\theta h_\omega\left(\tau_T^t(\delta_t^\theta - \delta_t^{\theta_0}) - \tau_T^{t'}\delta_{t'}^\theta\right)\right\|_1 \\
&\quad \cdot\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^2|\mathcal{F}_f(\omega_t^t)||V^{t'}(\omega)|d\omega \\
&= \mathcal{O}_{\mathbb{P}}\left(\int_{\Omega_T}\frac{\|\omega\|^2}{\sqrt{\beta_T}}\left\{\frac{1}{\sqrt{T}}+|\mathcal{F}_f(\omega)|\right\}d\omega\right) \\
&= \mathcal{O}_{\mathbb{P}}\left(\frac{r_T^4}{\sqrt{T}\beta_T}+\frac{r_T}{\sqrt{\beta_T}}\int_{\Omega_T}\|\omega\||\mathcal{F}_f(\omega)|d\omega\right).
\end{aligned}$$

Because of $r_T^4/\sqrt{T}\beta_T \xrightarrow{T \rightarrow \infty} 0$, $r_T/\sqrt{\beta_T} \xrightarrow{T \rightarrow \infty} 0$, and Assumption 6.28, we conclude that

$$\sup_{\theta \in U'} \|\text{Hess}_\theta B_T(\theta)\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \quad (6.92)$$

Finally, we consider $\text{Hess}_\theta C_T(\theta)$. Because of (6.91) and Lemma 5.4, it is

$$\begin{aligned} & \sup_{\theta \in U'} \|\text{Hess}_\theta C_T(\theta)\|_1 \\ &= \sup_{\theta \in U'} \left\| - \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \text{Hess}_\theta h_\omega(\tau_T^t \delta_t^\theta - \tau_T^{t'} \delta_{t'}^\theta) V_T^t(\omega) \overline{V_T^{t'}(\omega)} \, d\omega \right\|_1 \\ &\leq \sup_{\theta \in U'} \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t, t' \in \mathbb{T}} \left\| \text{Hess}_\theta h_\omega(\tau_T^t \delta_t^\theta - \tau_T^{t'} \delta_{t'}^\theta) \right\|_1 \left| V_T^t(\omega) \overline{V_T^{t'}(\omega)} \right| \, d\omega \\ &= \int_{\Omega_T} \mathcal{O}_{\mathbb{P}} \left(\frac{\|\omega\|^2}{\beta_T} \right) \, d\omega = \mathcal{O}_{\mathbb{P}} \left(\frac{r_T^4}{\beta_T} \right). \end{aligned}$$

Since $r_T^4/\beta_T \xrightarrow{T \rightarrow \infty} 0$, we get

$$\sup_{\theta \in U'} \|\text{Hess}_\theta C_T(\theta)\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.} \quad (6.93)$$

From (6.90), (6.92), and (6.93), we conclude that

$$\left\| \text{Hess}_\theta \tilde{N}_T(\theta_0) - H_N \right\|_1 \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability.}$$

□

Theorem 6.42 (Uniform tightness of drift estimators). *Under the Assumptions 6.36 and 6.40 and the Assumptions of Theorem 6.35, if $\hat{\theta}_T \xrightarrow{T \rightarrow \infty} \theta_0$ in probability, the sequence $(\sqrt{T}(\hat{\theta}_T - \theta_0))_{T \in \mathbb{N}}$ is uniformly tight.*

Proof. By Assumption 6.25, there is a convex open neighborhood $U' \subseteq \Theta$ of θ_0 such that \tilde{N}_T is twice continuously differentiable on U' . In particular, if $\theta \in U'$ is a minimizer of \tilde{N}_T , then $\text{grad}_\theta \tilde{N}_T(\theta) = 0$. We define

$$\tilde{G}_T(\theta) := \begin{cases} \text{grad}_\theta \tilde{N}_T(\theta), & \text{if } \theta \in U', \\ \text{grad}_\theta \tilde{N}_T(\theta_0), & \text{if } \theta \in \Theta \setminus U'. \end{cases}$$

Because, by definition, $\hat{\theta}_T$ is a minimizer of \tilde{N}_T , and because $\hat{\theta}_T \xrightarrow{T \rightarrow \infty} \theta_0 \in U'$ in probability, we have that

$$\begin{aligned} \mathbb{P} \left(\sqrt{T} \tilde{G}_T(\hat{\theta}_T) > \epsilon \right) &= \mathbb{P} \left(\sqrt{T} \text{grad}_\theta \tilde{N}_T(\hat{\theta}_T) > \epsilon, \hat{\theta}_T \in U' \right) \\ &\quad + \mathbb{P} \left(\sqrt{T} \text{grad}_\theta \tilde{N}_T(\theta_0) > \epsilon, \hat{\theta}_T \notin U' \right) \\ &\leq \mathbb{P} \left(\hat{\theta}_T \notin U' \right), \end{aligned}$$

which converges to 0 for all $\epsilon > 0$, that is,

$$\sqrt{T} \tilde{G}_T(\hat{\theta}_T) = o_{\mathbb{P}}(1). \quad (6.94)$$

If $\theta \in U'$, the mean value theorem for real functions of multiple variables applied to each component of $\text{grad}_\theta \tilde{N}_T(\theta)$ yields that

$$\text{grad}_\theta \tilde{N}_T(\theta) = \text{grad}_\theta \tilde{N}_T(\theta_0) + \text{Hess}_\theta \tilde{N}_T(\theta^\dagger)(\theta - \theta_0), \quad (6.95)$$

with a component-wise convex combination $\theta^\dagger \in U$ of θ and θ_0 . By (6.95), on the event $\{\hat{\theta}_T \in U'\}$, there is $\hat{\theta}_T^\dagger \in U'$ between $\hat{\theta}_T$ and θ_0 with

$$\text{grad}_\theta \tilde{N}_T(\hat{\theta}_T) = \text{grad}_\theta \tilde{N}_T(\theta_0) + \text{Hess}_\theta \tilde{N}_T(\hat{\theta}_T^\dagger)(\hat{\theta}_T - \theta_0). \quad (6.96)$$

With the definitions

$$\hat{\theta}_T^* := \begin{cases} \hat{\theta}_T^\dagger, & \text{if } \hat{\theta}_T \in U', \\ \theta_0, & \text{if } \hat{\theta}_T \in \Theta \setminus U', \end{cases}$$

and

$$\tilde{H}_T := \begin{cases} \text{Hess}_\theta \tilde{N}_T(\hat{\theta}_T^*) = \text{Hess}_\theta \tilde{N}_T(\hat{\theta}_T^\dagger), & \text{if } \hat{\theta}_T \in U', \\ 0, & \text{if } \hat{\theta}_T \in \Theta \setminus U', \end{cases}$$

and equation (6.94), it follows that

$$\sqrt{T} \text{grad}_\theta \tilde{N}_T(\theta_0) + \tilde{H}_T \sqrt{T}(\hat{\theta}_T - \theta_0) = \sqrt{T} \tilde{G}_T(\hat{\theta}_T) = o_{\mathbb{P}}(1), \quad (6.97)$$

which is true on the event $\{\hat{\theta}_T \in U'\}$ by (6.96) and on the complement $\{\hat{\theta}_T \notin U'\}$ by design of \tilde{G}_T and \tilde{H}_T . From (6.97) and Theorem 6.35, we get that

$$\tilde{H}_T \sqrt{T}(\hat{\theta}_T - \theta_0) = -\sqrt{T} \text{grad}_\theta \tilde{N}_T(\theta_0) + o_{\mathbb{P}}(1) = \mathcal{O}_{\mathbb{P}}(1). \quad (6.98)$$

Because $\hat{\theta}_T$ is a consistent estimator for θ_0 and $\hat{\theta}_T^*$ is between $\hat{\theta}_T$ and θ_0 , it follows at once that $\hat{\theta}_T^* \xrightarrow{T \rightarrow \infty} \theta_0$ in probability. Since H_N is invertible by Lemma 6.39, Theorem 6.41 implies that $\tilde{H}_T \xrightarrow{T \rightarrow \infty} H_N$ in probability. From this, (6.98), and Lemma A.17 in the Appendix, we conclude that $\sqrt{T}(\hat{\theta}_T - \theta_0) = \mathcal{O}_{\mathbb{P}}(1)$. \square

Chapter 7

Model selection and motion blur

To select fitting parametric models for the drift, rotation, and scaling functions, we can estimate with different models and choose the one which leads to the estimator with the lowest contrast value. This can be interpreted in the following way. We explain this for drift, the situation for the rotation and scaling is similar. Instead of minimizing the contrast functional $N_T(\theta)$ over $\theta \in \Theta$ for, for example, linear drift functions δ^θ , we could also minimize the contrast as a functional of the drift function itself, $N_T(\delta)$, over all linear functions δ on $[0, 1]$. This can now easily be extended to encompass a variety of drift models, for example, by considering the set of all polynomial drift functions on $[0, 1]$ with degree less or equal to 3 (which we used in the Chapters 9 and 10). We can then estimate the drift by the parametric model (and the associated parameter estimator) with the overall smallest contrast value. Of course, since polynomial models of high degree contain those of lesser degree, this model selection should be superfluous if we only consider polynomial models. However, since we have to deal with statistical errors and we minimize the contrast numerically, a too complex model (e.g., assuming a cubic polynomial for the drift function when it is actually linear) might yield a worse reconstruction than a simpler one, for example, due to over-fitting.

Furthermore, we use an objective means of measuring how well the estimation method reduced the blur of the superimposed image due to the affine transformations. We will also use this measure to compare our results to those of alternative estimation methods which are described in Chapter 8.

We will explain the procedure for translation, first, since the application to rotation and scaling is based on it.

7.1 Drift blur

The following was proposed in (Xu et al., 2013) which is based on the work of (Chen et al., 2010).

Definition 7.1 (Motion blur measure). *Let $I: \{x_j \mid j \in \{1, \dots, n\}\} \rightarrow \mathbb{R}$ a (grey value) pixel image, $\psi \in [0, 2\pi)$, and*

$$J(\psi) := \sum_{j=1}^n \left(\Delta I((x_j)_1, (x_j)_2)_\psi \right)^2$$

the average squared directional derivative of I in direction $(\cos(\psi), \sin(\psi))^\top$. We define

the motion blur measure as

$$m_2(I) := \log \left(\frac{J(\psi_{\max})}{J(\psi_{\min})} \right),$$

where we have an estimated motion direction $\psi_{\min} \in \operatorname{argmin}_{\psi \in [0, 2\pi)} J(\psi)$ and ψ_{\max} is the direction perpendicular to ψ_{\min} .

Remark 7.2. Note, that $J(\psi) = 0$ if and only if I is constant in direction ψ . An advantage of m_2 is that $m_2(aI) = m_2(I)$ for all $a > 0$ which means that it does not depend on the grey value scale of the image. The fact that ψ_{\min} is selected as a minimizer of J is based on the idea that the image is blurred in the direction of the motion and thus the image intensity changes little in this direction (on average), while it varies much more in the perpendicular direction.

The minimizer ψ_{\min} can be obtained as follows. Rewrite

$$J(\psi) = (\cos(\psi), \sin(\psi)) D (\cos(\psi), \sin(\psi))^{\top},$$

where

$$D = \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \quad \text{with} \quad d_{rs} := \sum_{j=1}^n \frac{\partial I}{\partial (x)_r}((x_j)_1, (x_j)_2) \cdot \frac{\partial I}{\partial (x)_s}((x_s)_j, (x_j)_2),$$

for $r, s \in \{1, 2\}$. Then,

$$J(\psi) = d_{11} + d_{12} \sin(2\psi) + (d_{22} - d_{11})(\sin(\psi))^2.$$

We get the minimum value of J by setting

$$\frac{dJ(\psi)}{d\psi} = d_{12} 2 \cos(2\psi) + (d_{22} - d_{11}) \sin(2\psi) = 0,$$

which yields $\psi = \psi_m + (r\pi)/2$, $r \in \mathbb{Z}$, with $\psi_m = \arctan(2d_{12}/(d_{11} - d_{22}))/2$. The motion direction is then determined by

$$\psi_{\min} := \begin{cases} \psi_m & \text{if } J(\psi_m) \leq J(\psi_m + \pi/2), \\ \psi_m + \pi/2 & \text{if } J(\psi_m) > J(\psi_m + \pi/2). \end{cases}$$

The $J(\psi_{\max})$ also keeps the blur measure value low in the case of an image that is (almost) constant over wide areas (where the directional derivative is small in any direction).

Remark 7.3 (Flaw of m_2). While m_2 is a good measure of blurring in a known direction, it has a major flaw when it comes to determining the motion direction. Consider an image I which contains a vertical white bar ten pixels long and 1 pixel wide and is black everywhere else. Then, the motion blur measure m_2 will register the long bright object as a vertically blurred shorter and even brighter object. If we now blur this image horizontally such that it now contains a less bright (i.e., grey) bar of length 10 pixels and width 5 pixels, the motion blur value of that image (in vertical direction) will be less than the one of the original image because the total squared directional derivative in horizontal direction (i.e., $J(\psi_{\max})$) becomes much smaller due to the lower grey value of the bar.

This problem occurs when there is a stretched object of relatively high brightness in the image, as is the case for the curved lines in the test image in Figure 9.1 and for the

dense structure in the middle lower third of the data shown in Figure 10.3. We can thus apply m_2 only if we already know a meaningful motion direction. In simulations, this is easy since we know the true drift, rotation, and scaling and can thus use the average true motion direction. In our application to SMS data, we will use the average motion direction of fiducial markers which mostly follow the true and unknown motion of the specimen they are attached to (see 8), to assess the quality of our estimation results.

7.2 Rotation and scaling blur

We already know that a rotation and scaling in the image domain results in the same rotation and an inverse scaling in the Fourier domain. After taking the squared Fourier magnitudes (to get rid of the drift) and after performing a log-polar coordinate transform, the rotation and scaling became a translation. This means that we can apply the motion blur measure m_2 to $|\mathcal{F}_{\tilde{f}_T}|^2 \circ \mathcal{LP}$ to get an estimate of the quality of the rotation and scaling correction, where \tilde{f}_T is the point-wise average of the rotation and scaling corrected frames.

7.3 Motion blur measure in application

In our simulation study (Chapter 9), since we already know the true drift function $\delta_t(\theta_0)$, we can choose the average drift direction $\int_0^1 \partial \delta_t(\theta_0) / \partial t dt = \delta_1(\theta_0)$ as the motion direction (after normalization). Similarly, we can choose the average rotation and logarithmic scaling direction to compute the motion blur of the superposition of log-polar transformed squared Fourier magnitudes of the images after rotation and scaling correction, but before drift correction. Hence, in our context (where, for drift, I is either \tilde{f}_T or the superimposed image, and similar for rotation and scaling, see Tables 9.7 and 9.8) we get the motion blur measure

$$\tilde{m}_2 := \log \left(\frac{\sum_{j=1}^n \langle \text{grad}_x I(x_j), R_{\pi/2} \delta_1(\theta_0) / \|\delta_1(\theta_0)\| \rangle^2}{\sum_{j=1}^n \langle \text{grad}_x I(x_j), \delta_1(\theta_0) / \|\delta_1(\theta_0)\| \rangle^2} \right). \quad (7.1)$$

Note that the average drift direction used to determine the motion blur (7.1) in the case of a drift function with jump is (before normalization)

$$t_0 \delta_{t_0}(\theta) + (1 - t_0) \left(\delta_1(\theta) - \lim_{t \searrow t_0} \delta_t(\theta) \right)$$

instead of just $\delta_1(\theta)$, where t_0 is the time at which the jump occurs. We calculated an approximation of $\text{grad}_x I$ as follows (see e.g., (Gonzalez, R.C. and Woods, R.E., 2002)).

Let I be a pixel image of size $N \times N$. For every pixel location (i, j) , $i, j \in \{1, \dots, N\}$, the gradient of I is defined as $\nabla I(i, j) := (G_x(i, j), G_y(i, j))^T$ with

$$G_x(i, j) := \sum_{i', j'=-1}^1 S_x(i' + 2, j' + 2) I(i + i', j + j'),$$

$$G_y(i, j) := \sum_{i', j'=-1}^1 S_y(i' + 2, j' + 2) I(i + i', j + j'),$$

where we extend the image periodically, that is, $I(0, j) := I(N, j)$, $I(N + 1, j) := I(1, j)$, $I(i, 0) := I(i, N)$, and $I(i, N + 1) := I(i, 1)$ and so on. Here, S_x and S_y are the Sobel

masks

$$S_x := \frac{1}{8} \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \quad S_y := \frac{1}{8} \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}.$$

Often, especially if I is noisy, it is beneficial to smooth the image first, for example, with a Gauss kernel

$$K := \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

This means that we replace every $I(i, j)$ with the weighted average

$$\bar{I}(i, j) := \sum_{i', j'=-1}^1 K(i' + 2, j' + 2) I(i + i', j + j')$$

of the 3×3 pixel area centred on it. Because our images (see Chapters 9 and 10) are very noisy, we smooth twice before applying the motion blur measure.

Chapter 8

Other estimation methods

In this chapter, we take a look at two other methods which can be used to estimate drift, rotation, and scaling in an image sequence. The first uses cross correlation to compare all possible pairs of images, while the second relies on tracking the positions of markers that are fixed on the specimen and appear in every single frame.

8.1 Cross Correlation

Definition 8.1 (Cross correlation). *Let $g_1, g_2: \mathbb{R}^2 \rightarrow \mathbb{C}$ integrable. The cross correlation of g_1 and g_2 is defined as*

$$g_1 \star g_2: \mathbb{R}^2 \rightarrow \mathbb{C}, \quad \tau \mapsto \int_{\mathbb{R}^2} \overline{g_1(x)} g_2(x + \tau) dx,$$

where the argument $\tau \in \mathbb{R}^2$ is called the lag.

In applications, we can either implement a loop which computes the cross correlation for all possible lags $\tau \in \{x_j - x_{j'} \mid j, j' \in \{1, \dots, n\}\}$, with n equidistant pixel locations x_j , or we use a trick. The following theorem gives an efficient way to compute the cross correlation numerically, since it allows to make use of the well-known fast Fourier transform algorithm (FFT, see e.g., Cooley and Tukey (1965)).

The following theorem is similar to the convolution theorem (see e.g., Rudin (1990, Theorem 1.2.4 (b))).

Theorem 8.2. *For $g_1, g_2: \mathbb{R}^2 \rightarrow \mathbb{C}$, we have $g_1 \star g_2 = \mathcal{F}^{-1}(\overline{\mathcal{F}_{g_1}} \cdot \mathcal{F}_{g_2})$.*

Proof. Since $\overline{ab} = \bar{a} \cdot \bar{b}$ for all $a, b \in \mathbb{C}$, we have for all $\omega \in \mathbb{R}^2$ that

$$\begin{aligned} \overline{\mathcal{F}_{g_1}(\omega)} \cdot \mathcal{F}_{g_2}(\omega) &= \overline{\int_{\mathbb{R}^2} g_1(x) e^{-2\pi i \langle \omega, x \rangle} dx} \cdot \int_{\mathbb{R}^2} g_2(y) e^{-2\pi i \langle \omega, y \rangle} dy \\ &= \left(\int_{\mathbb{R}^2} \overline{g_1(x)} e^{2\pi i \langle \omega, x \rangle} dx \right) \cdot \int_{\mathbb{R}^2} g_2(y) e^{-2\pi i \langle \omega, y \rangle} dy \\ &= \int_{\mathbb{R}^2} \overline{g_1(x)} e^{2\pi i \langle \omega, x \rangle} \left(\int_{\mathbb{R}^2} g_2(y) e^{-2\pi i \langle \omega, y \rangle} dy \right) dx \\ &= \int_{\mathbb{R}^2} \overline{g_1(x)} e^{2\pi i \langle \omega, x \rangle} \left(\int_{\mathbb{R}^2} g_2(x + \tau) e^{-2\pi i \langle \omega, x + \tau \rangle} d\tau \right) dx \\ &= \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} \overline{g_1(x)} g_2(x + \tau) dx \right) e^{-2\pi i \langle \omega, \tau \rangle} d\tau = \mathcal{F}_{g_1 \star g_2}(\omega), \end{aligned}$$

with $\tau = y - x$. Applying the inverse Fourier transform to both sides yields the assertion. \square

Cross correlation can obviously be useful in our SMS-setup for motion correction. Consider a (binned) sequence of frames $(O^t = \{O_j^t \mid j \in J_T^t\})_{t \in \mathbb{T}}$, translated by vectors $(\delta^t)_{t \in \mathbb{T}}$, rotated by angles $(\phi^t)_{t \in \mathbb{T}}$, and scaled by factors $(\sigma^t)_{t \in \mathbb{T}}$, each frame representing a pixel image of size $n_1 \times n_2$, with $n_1 \geq n_2$ and a total pixel number of $n = n_1 n_2$. Similarly to the M-estimation method summarized in Algorithm 4.10, we first compute the analytical Fourier-Mellin transformation of the squared Fourier magnitude of each frame, where we also transform into log-polar coordinates to convert the rotation and scaling into a translation (where we assume that the rotation shows in the n_1 -direction and the scaling in the n_2 -direction). We then maximize the cross correlation between any two Fourier-Mellin frames, for example, at times $t, t' \in \mathbb{T}$, to get an optimal lag $\tau^{t,t'}$ between the t -th and the t' -th Fourier-Mellin frames. Then, $\frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} \tau^{t,t'}$ estimates the average translation of the t -th Fourier-Mellin frame with respect to the whole sequence. By subtracting the average initial lag, $\frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} \tau^{0,t'}$, we get

$$\bar{\tau}^{t,\bullet} := \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} (\tau^{t,t'} - \tau^{0,t'}).$$

Now, we have to reconstruct the rotation angles ρ^t and scaling factors σ^t from $(\bar{\tau}^{t,\bullet})_{t \in \mathbb{T}}$. For the rotation, this essentially means rescaling the range from n_1 to 2π . Since we want to estimate angles in $(-\pi, \pi]$, we further need to shift by $\pi/2$, count modulo π , and subtract $\pi/2$ again. We thus estimate ρ^t by

$$\hat{\rho}^t := \left(\left(\frac{2\pi}{n_1} \bar{\tau}_1^{t,\bullet} + \frac{\pi}{2} \right) \bmod \pi \right) - \frac{\pi}{2}, \quad t \in \mathbb{T}.$$

For the scaling factor, due to the log-polar coordinate transformation, the corresponding components of the $\bar{\tau}^{t,\bullet}$'s are in a logarithmic scale. Since the maximal possible radius is $n_2/2$, we have

$$\hat{\sigma}^t := \left(\frac{n_2}{2} \right)^{-\bar{\tau}_2^{t,\bullet}/(n_2/2)}, \quad t \in \mathbb{T}.$$

Next, we calibrate the original frames O^t , rotating them by $-\hat{\rho}^t$ and scaling them with $1/\hat{\sigma}^t$. Straightforward cross correlation on the set of calibrated frames results in optimal lags $\tilde{\tau}^{t,t'}$ and, subsequently, translation estimates

$$\hat{\delta}^t := \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} (\tilde{\tau}^{t,t'} - \tilde{\tau}^{0,t'}).$$

Shifting the already for rotation and scaling corrected frames by $-\hat{\delta}^t$ and, thereafter, overlaying all of them yields an estimator for the image f .

We summarize this procedure in the following algorithm.

Algorithm 8.3 (Image correction via cross correlation). *We assume data in the form of pixel images of size $n_1 \times n_2$, with $n_1 \geq n_2$ and a total pixel number of $n = n_1 n_2$.*

1. Given a (binned) sequence of frames $(O^t = \{O_j^t \mid j \in J_T^t\})_{t \in \mathbb{T}}$, compute the squared Fourier magnitudes $|\mathcal{F}_{O^t}^{\mathbb{X}^t}|^2$, where $\mathbb{X}^t = \{x_j \mid j \in J_T^t\}$, using fast Fourier transform (FFT).

2. Choose an appropriate $\gamma > 0$, transform the $|\mathcal{F}_{O^t}^{\mathbb{X}^t}|^2$'s into log-polar coordinates, multiply with $e^{l\gamma}$, where l is the logarithm of the radius, and apply FFT again to get $\mathcal{F}_{G^t} = \mathcal{AFM}_{|\mathcal{F}_{O^t}^{\mathbb{X}^t}|^2}$, where

$$G^t: \mathbb{R} \times (-\pi, \pi] \rightarrow [0, \infty), \quad (l, \psi) \mapsto e^{l\gamma} \left(\left| \mathcal{F}_{O^t}^{\mathbb{X}^t} \right|^2 \circ \mathcal{LP} \right) (l, \psi).$$

3. For each pair $(t, t') \in \mathbb{T}^2$, compute the optimal lag

$$\tau^{t,t'} = (\tau_1^{t,t'}, \tau_2^{t,t'}) \in \arg \max_{\tau \in \mathbb{R}^2} (G^t \star G^{t'}) (\tau) = \arg \max_{\tau \in \mathbb{R}^2} \mathcal{F}^{-1}(\overline{\mathcal{F}_{G^t}} \cdot \mathcal{F}_{G^{t'}}) (\tau).$$

If $\tau^{t,t'}$ is not unique, take the one with the smallest length $\|\tau^{t,t'}\|$.

4. For all $t \in \mathbb{T}$, let $\bar{\tau}^{t,\bullet} := \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} (\tau^{t,t'} - \tau^{0,t'})$ and estimate the rotation angles and scaling factors with

$$\hat{\rho}^t := \left(\left(\frac{2\pi}{n_1} \bar{\tau}_1^{t,\bullet} + \frac{\pi}{2} \right) \bmod \pi \right) - \frac{\pi}{2}, \quad \hat{\sigma}^t := \left(\frac{n_2}{2} \right)^{-\bar{\tau}_2^{t,\bullet}/(n_2/2)}, \quad t \in \mathbb{T}.$$

5. For each $t \in \mathbb{T}$, scale O^t with $1/\hat{\sigma}^t$ and rotate it by $-\hat{\rho}^t$, using for example, bilinear interpolation, calling the resulting calibrated frames P^t .

6. For each pair $(t, t') \in \mathbb{T}^2$, compute the optimal lag

$$\tilde{\tau}^{t,t'} \in \arg \max_{\tau \in \mathbb{R}^2} (P^t \star P^{t'}) (\tau) = \arg \max_{\tau \in \mathbb{R}^2} \mathcal{F}^{-1}(\overline{\mathcal{F}_{P^t}} \cdot \mathcal{F}_{P^{t'}}) (\tau).$$

If $\tilde{\tau}^{t,t'}$ is not unique, take the one with the smallest length $\|\tilde{\tau}^{t,t'}\|$.

7. Estimate the translation vector by $\hat{\delta}^t := \frac{\beta_T}{T} \sum_{t' \in \mathbb{T}} (\tilde{\tau}^{t,t'} - \tilde{\tau}^{0,t'})$.

8. For each $t \in \mathbb{T}$, translate P^t by $-\hat{\delta}^t$, calling the resulting calibrated frames Q^t .

9. Compute the point-wise average $\hat{f}(x) := \frac{\beta_T}{T} \sum_{t \in \mathbb{T}} Q^t(x)$.

8.2 Fiducial marker tracking

As was pointed out in the introduction, it is common practice to incorporate fiducial markers (bright fluorescent spheres) into the specimen. Since we observe these fiducial markers in every frame, we can track their positions and use this information to calibrate the images. For each fiducial marker μ , we can visually identify a small rectangular subset $[a, b] \times [c, d] \subseteq [0, 1]^2$ of the image domain such that (presumably all of) the data coming from μ during the entire data acquisition time is inside this rectangle. For each $t \in \mathbb{T}$, let $\mu^t \in [0, 1]^2$ the true and unknown position of the fiducial marker and let $\{z_i^t \mid i \in \{1, \dots, k^t\}\} \subseteq [a, b] \times [c, d]$ the recorded positions in $[a, b] \times [c, d]$ at time t , where $k^t \in \mathbb{N}$. We estimate μ^t by the spatial mean of all z_i^t , that is,

$$\hat{\mu}^t := \frac{1}{k^t} \sum_{i=1}^{k^t} z_i^t \in [a, b] \times [c, d].$$

This gives us a sequence $(\hat{\mu}^t)_{t \in \mathbb{T}}$ of estimated fiducial marker positions.

Now, assume that we observe two fiducial markers, μ_1 and μ_2 , each with their respective rectangular occurrence areas, $[a_1, b_1] \times [c_1, d_1] \subseteq [0, 1]^2$ and $[a_2, b_2] \times [c_2, d_2] \subseteq [0, 1]^2$, such that

$$([a_1, b_1] \times [c_1, d_1]) \cap ([a_2, b_2] \times [c_2, d_2]) = \emptyset, \quad (8.1)$$

and they contain k_1^t and k_2^t measurements, respectively. Then, for each of these fiducial markers, we get a sequence of estimated positions, $(\hat{\mu}_1^t)_{t \in \mathbb{T}}$ and $(\hat{\mu}_2^t)_{t \in \mathbb{T}}$. From these, we can estimate rotation and scaling of the specimen in the following way. Define the difference vectors

$$\hat{d}^t = ((\hat{d}^t)_1, (\hat{d}^t)_2)^\top := \hat{\mu}_1^t - \hat{\mu}_2^t \in \mathbb{R}^2, \quad t \in \mathbb{T}.$$

Because of (8.1), $\hat{d}^t \neq 0$ for all $t \in \mathbb{T}$. Homogeneous scaling of the whole specimen by a factor $\sigma > 0$ implies that $\|\hat{d}^t\|$ is also scaled by σ . Similarly, rotating the specimen by an angle $\rho \in (-\pi, \pi]$ rotates \hat{d}^t by the same angle. Note, that neither of this does depend on the position of the origin. Hence, we get a sequence of scaling and rotation estimators, $(\hat{\sigma}^t)_{t \in \mathbb{T}}$ and $(\hat{\rho}^t)_{t \in \mathbb{T}}$, with

$$\hat{\sigma}^t := \|\hat{d}^t\| / \|\hat{d}^0\|, \quad \hat{\rho}^t := \text{atan2}((\hat{d}^t)_2, (\hat{d}^t)_1) - \text{atan2}((\hat{d}^0)_2, (\hat{d}^0)_1),$$

where

$$\text{atan2}: \mathbb{R}^2 \setminus \{0\} \rightarrow (-\pi, \pi], \quad (y, x) \mapsto \begin{cases} \arctan(y/x), & \text{if } x > 0, \\ \arctan(y/x) + \pi, & \text{if } x < 0 \text{ and } y \geq 0, \\ \arctan(y/x) - \pi, & \text{if } x < 0 \text{ and } y < 0, \\ \pi/2, & \text{if } x = 0 \text{ and } y > 0, \\ -\pi/2, & \text{if } x = 0 \text{ and } y < 0, \end{cases}$$

is the angle between a vector $(x, y) \in \mathbb{R}^2 \setminus \{0\}$ and the x -axis. We can then calibrate the images with those estimators by rotating the t -th frame by $-\hat{\rho}^t$ and scaling it by $1/\hat{\sigma}^t$. This changes the estimated marker positions to

$$\tilde{\mu}_j^t := \frac{1}{\hat{\sigma}^t} \cdot R_{-\hat{\rho}^t} \hat{\mu}_j^t, \quad j \in \{1, 2\}.$$

Finally, we correct the images for drift by translating the t -th frame by $-\tilde{\mu}_1^t$ (or $-\tilde{\mu}_2^t$). Since the estimated marker positions are subject to statistical errors and measurement errors, the average $-(\tilde{\mu}_1^t + \tilde{\mu}_2^t)/2$ is even better suited for drift correction since it reduces the variance by a factor of (about) $1/2$. The estimated image is then given by the point-wise average of all frames.

We summarize this in the following algorithm.

Algorithm 8.4 (Image correction via fiducial marker tracking).

1. Specify the rectangles containing the two fiducial markers, $[a_j, b_j] \times [c_j, d_j] \subseteq [0, 1]^2$, $j \in \{1, 2\}$, and collect the data points $\{z_{i,j}^t \mid t \in \mathbb{T}, i \in \{1, \dots, k_j^t\}\}$ within them.
2. For all $t \in \mathbb{T}$, compute the spatial means $\hat{\mu}_j^t := \frac{1}{k_t} \sum_{i=1}^{k_t} z_{i,j}^t$, $j \in \{1, 2\}$, and their differences, $\hat{d}^t := \hat{\mu}_1^t - \hat{\mu}_2^t$.
3. For each $t \in \mathbb{T}$, scale the t -th frame with $1/\hat{\sigma}^t$ and rotate it by $-\hat{\rho}^t$, given by

$$1/\hat{\sigma}^t := \|\hat{d}^0\| / \|\hat{d}^t\|, \quad -\hat{\rho}^t := \text{atan2}((\hat{d}^0)_2, (\hat{d}^0)_1) - \text{atan2}((\hat{d}^t)_2, (\hat{d}^t)_1).$$

4. For each $t \in \mathbb{T}$, translate the t -th frame by $-(\tilde{\mu}_1^t + \tilde{\mu}_2^t)/2$, where $\tilde{\mu}_j^t := \frac{1}{\sigma^t} \cdot R_{-\hat{\rho}^t} \hat{\mu}_j^t$, $j \in \{1, 2\}$.
5. Compute the point-wise average of all frames.

We will apply the motion blur measure from Chapter 7 to compare the results of fiducial marker tracking with those of our M-estimation method and those from cross correlation.

Chapter 9

Simulation Study

To investigate the finite sample properties of the M-estimation method proposed in Chapter 4, we conduct a simulation study with images of size $n = N^2$ pixels with $N = 512$. We opt for $T \in \{20, 50, 100\}$ in order to reduce computational time, as our implementation requires several minutes to compute $\hat{\theta}_T$ on a 512×512 image for $T = 500$, say. In order to make these simulations comparable to the data in Chapter 10, we choose drift, rotation, and scaling parameters θ_0 , ϕ_0 , and α_0 such that the total motion (i.e., the pixel shift between the first and the last image) has comparable scale to the ones observed in our SMS data.

We consider the model (3.2) with four different motion types: linear, quadratic, and cubic motion functions, as well as a piecewise linear motion with a jump at an unknown time. Note that the motion with jump violates the Lipschitz properties in the Assumptions 6.10 and 6.31.

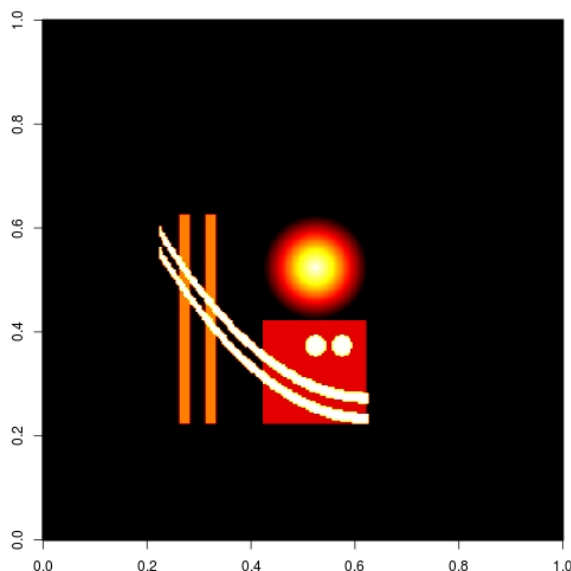


Figure 9.1: Test image \tilde{f} with grey scale values (rescaled to the unit interval), represented by colours ranging from black (0) over red and yellow to white (1).

We use the test image displayed in Figure 9.1, with image intensity \tilde{f} ranging from zero to one (the average image intensity is about 0.045), and apply three error models. We aim to apply our method to SMS microscopy and therefore, following the model (3.2), introduce a randomness of the selected pixel locations at each time point by multiplying

the noisy grey values with Bernoulli random variables. More precisely, for every pixel location x_j , $j \in \{1, \dots, n\}$, and every time point $t \in \mathbb{T}$, we observe

$$\tilde{O}_j^t := \left(\tilde{f}^t(x_j) + \tilde{\nu}_j^t \tilde{\epsilon}_j^t \right) B_j^t, \quad (9.1)$$

where the B_j^t 's are independent Bernoulli random variables with parameter $p = 5/T$ meaning that every pixel is seen five times on average over the whole observation period. Note that \mathbb{T} is defined in such a way that the whole time interval is always $[0, 1]$, that is, if $T = 100$, the time between two subsequent frames is exactly half as long as if T were only 50.

First, we employ a Gaussian error model, which means that the $\tilde{\epsilon}_j^t$ in (9.1) are independent standard normal random variables.

Secondly, in order to test the robustness of our estimation method against outliers, we assume that the $\tilde{\epsilon}_j^t$ in (9.1) are i.i.d. t -distributed with 2 degrees of freedom.

Finally, we simulate a Poisson error model,

$$\tilde{O}_j^t \sim \text{Pois} \left(p \tilde{f}^t(x_j) \right), \quad (9.2)$$

where the \tilde{O}_j^t are mutually independent and the Poisson intensity contains the Bernoulli-probability p such that the average image intensity of the two models (9.1) and (9.2) is the same.

For the Poisson model (9.2), we use a variance stabilizing transformation, $\sqrt{\tilde{O}_j^t + 1/4}$ (see e.g., (Frick et al., 2013, page 378)). We minimize discretized versions of the contrast functionals from Definitions 4.1 and 4.6 and use fast Fourier transform (FFT) which can be performed in $O(N^2 \cdot 2 \log(N))$ steps, exploiting Remark 2.15 so we can apply FFT for the rotation and scaling contrast. The minimizers are evaluated by a standard Nelder-Mead-type algorithm as implemented in the statistical software **R**. We specify the parameters as follows: $\nu = 0.1$ (Gaussian and t_2 -distributed errors), $r_T = u_T = v_T = 16$. As start value for the minimization algorithm we choose $0 \in \mathbb{R}^d$, where $d \in \{d_1, d_2 + d_3\}$ is the dimension of the drift parameter θ_0 or the rotation and scaling parameter (ϕ_0, α_0) , respectively.

Polynomial motion models have been described in the Examples 6.19 and 6.37. In the linear model we have $\delta_t(\theta) = \theta t$. We choose $\theta_0 = (50/512, 35/512)$, $\phi_0 = \pi/8$, and $\alpha_0 = 0.278$, that is, the image is shifted by 50 pixels in x_1 -direction and by 35 pixels in x_2 -direction, rotated by an angle of $\pi/8$ and scaled up to a factor of 1.278 over the time interval $[0, 1]$. In the quadratic drift model we set

$$\begin{aligned} \delta_t(\theta) &= (\theta_{11}, \theta_{21})^\top t + (\theta_{12}, \theta_{22})^\top t^2, \\ \rho_t(\phi) &= \phi_1 t + \phi_2 t^2, \\ \sigma_t(\alpha) &= 1 + \alpha_1 t + \alpha_2 t^2. \end{aligned}$$

For the drift in x_1 -direction we choose $(\theta_{11}, \theta_{12}) = (30/512, 30/512)$, in x_2 -direction $(\theta_{21}, \theta_{22}) = (10/512, 40/512)$. For the rotation we choose $(\phi_1, \phi_2) = (\pi/16, \pi/8)$ and for the scaling $(\alpha_1, \alpha_2) = (0.15, 0.1)$. Similarly, we employ a cubic model for drift, $\delta_t(\theta) = (\theta_{11}, \theta_{21})t + (\theta_{12}, \theta_{22})t^2 + (\theta_{13}, \theta_{23})t^3$, and for rotation and scaling accordingly. In this case, we choose the motion parameters $(\theta_{11}, \theta_{12}, \theta_{13}) = (20/512, 0, 25/512)$ as well as $(\theta_{21}, \theta_{22}, \theta_{23}) = (12/512, 20/512, 5/512)$, $(\phi_1, \phi_2, \phi_3) = (\pi/16, \pi/32, 5\pi/32)$, and $(\alpha_1, \alpha_2, \alpha_3) = (0, 0.1, 0.2)$.

The results of one estimate are reported in the Tables 9.1 and 9.2, the means of 100 simulations each in Tables 9.3 and 9.4. As recorded in the Tables 9.5 and 9.6, with

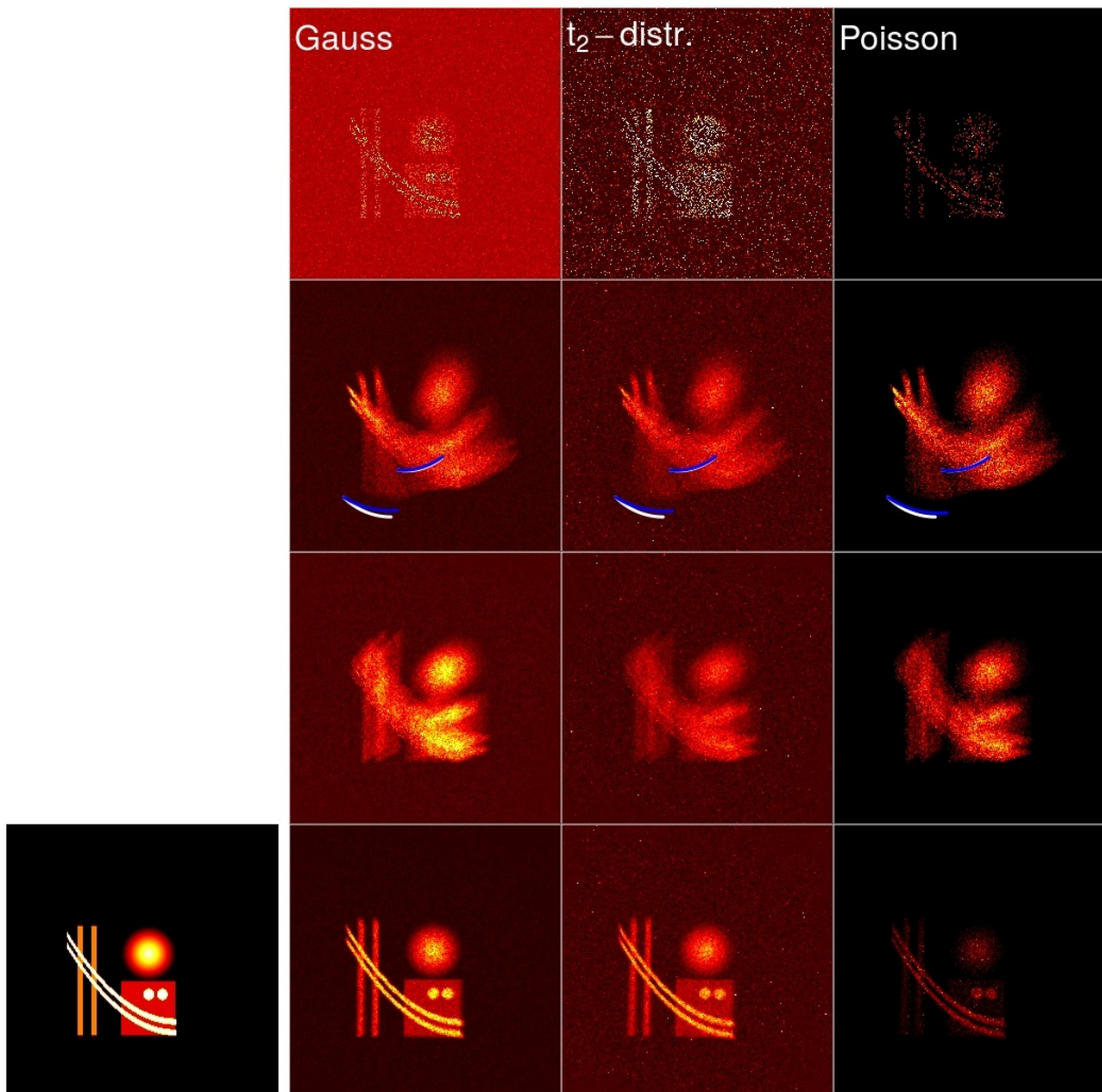


Figure 9.2: *Left (fourth row): Test image \tilde{f} with grey scale values (rescaled to the unit interval), represented by colours ranging from black (0) over red and yellow to white (1). Right: The first row shows the first frames of sequences of $T = 20$ noisy images subject to linear motion functions (from left to right: Gaussian noise, Student- t_2 noise, Poisson model). The second row depicts the corresponding superimposed images (i.e., the averages over all frames in the sequence). The true motion curves of two pixels are shown as white curve segments on top of which we plot the estimated motion in blue. The true and the estimated parameters are reported in the Tables 9.1 and 9.2. The third row shows the superimposed images after the rotation and scaling correction while the fourth row displays the images after the additional drift correction.*

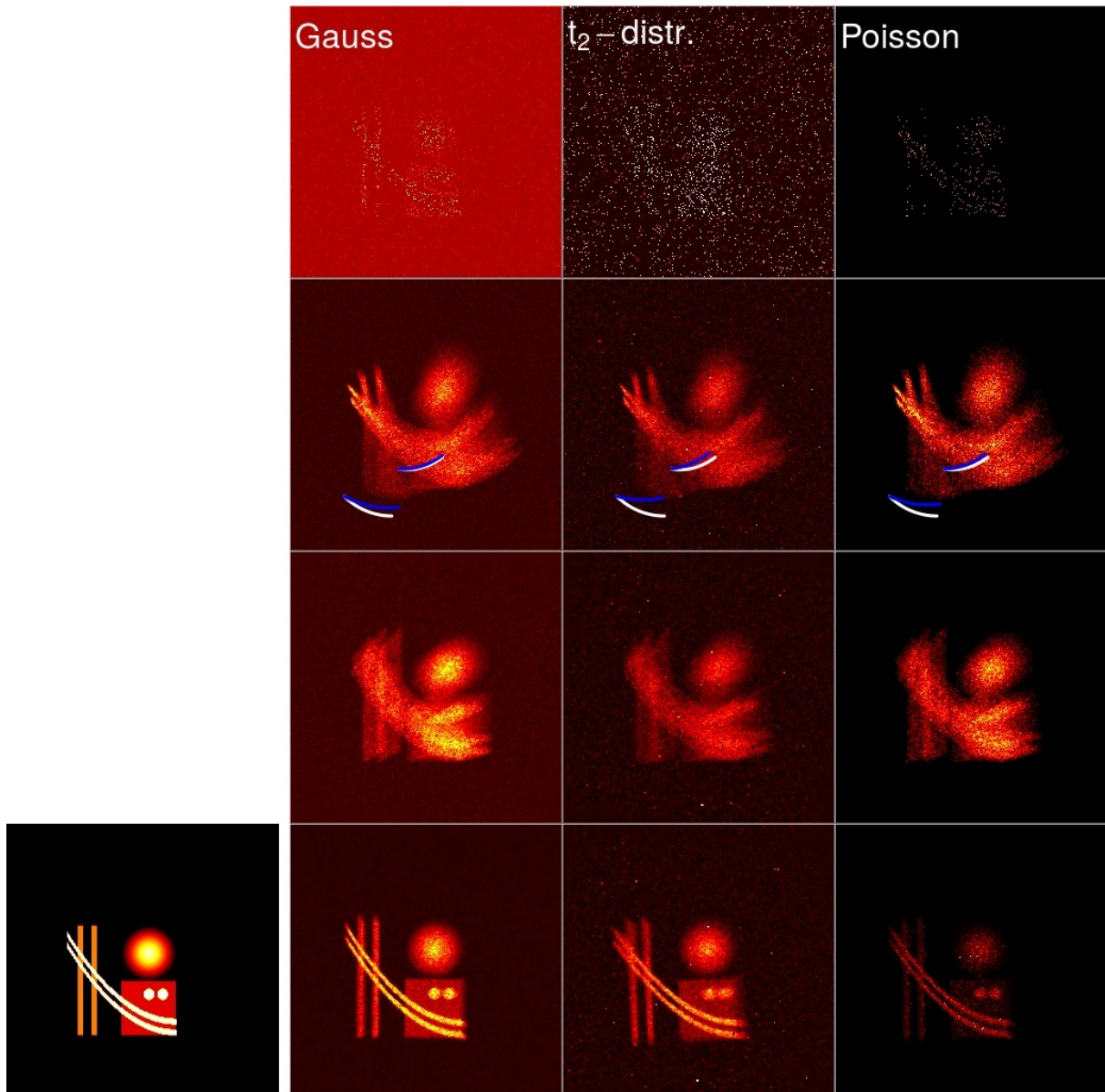


Figure 9.3: *Left (fourth row): Test image \tilde{f} with grey scale values (rescaled to the unit interval), represented by colours ranging from black (0) over red and yellow to white (1). Right: The first row shows the first frames of sequences of $T = 50$ noisy images subject to linear motion functions (from left to right: Gaussian noise, Student- t_2 noise, Poisson model). The second row depicts the corresponding superimposed images (i.e., the averages over all frames in the sequence). The true motion curves of two pixels are shown as white curve segments on top of which we plot the estimated motion in blue. The true and the estimated parameters are reported in the Tables 9.1 and 9.2. The third row shows the superimposed images after the rotation and scaling correction while the fourth row displays the images after the additional drift correction.*

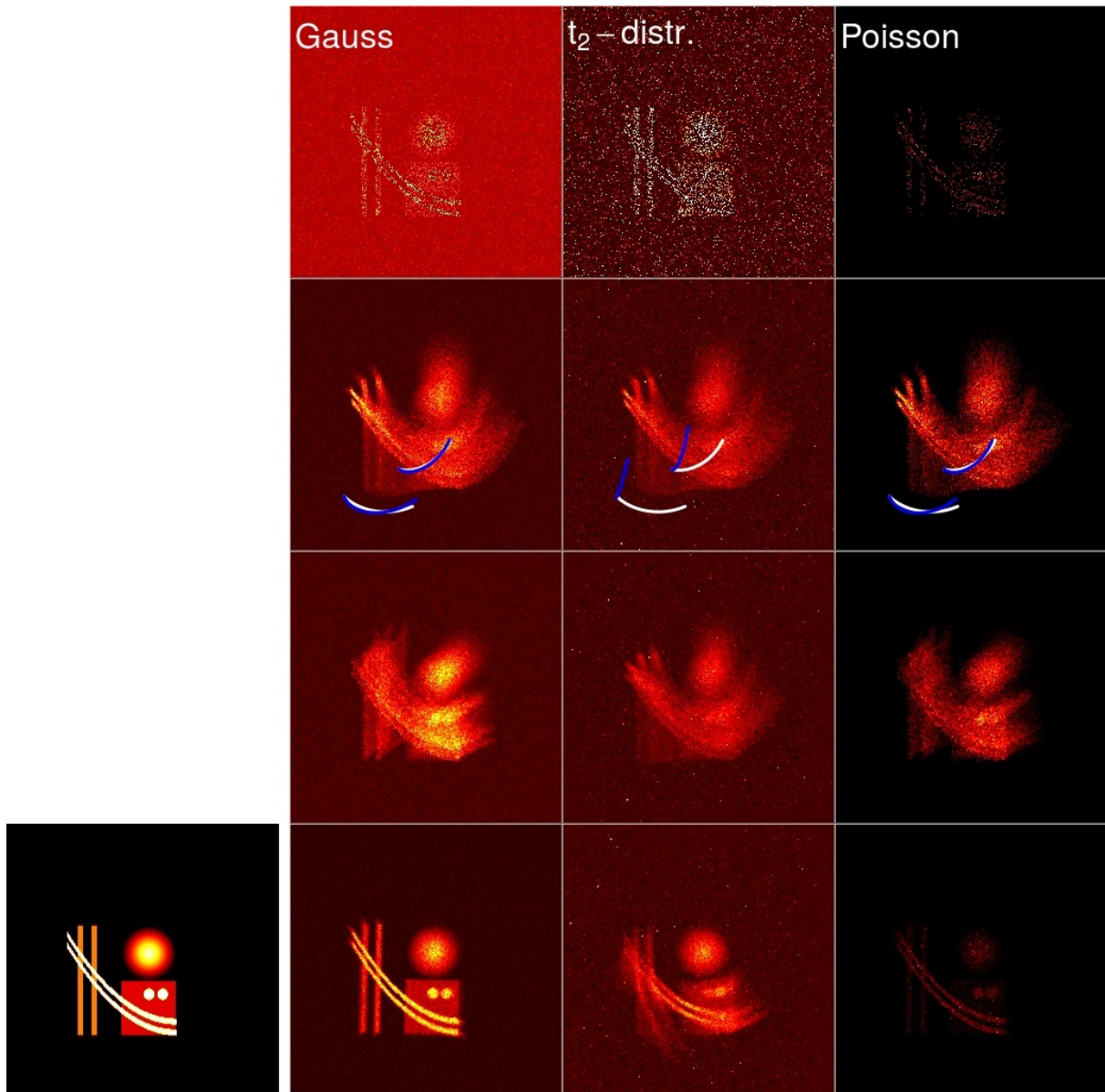


Figure 9.4: *Left (fourth row): Test image \tilde{f} with grey scale values (rescaled to the unit interval), represented by colours ranging from black (0) over red and yellow to white (1). Right: The first row shows the first frames of sequences of $T = 20$ noisy images subject to quadratic motion functions (from left to right: Gaussian noise, Student- t_2 noise, Poisson model). The second row depicts the corresponding superimposed images (i.e., the averages over all frames in the sequence). The true motion curves of two pixels are shown as white curve segments on top of which we plot the estimated motion in blue. The true and the estimated parameters are reported in the Tables 9.1 and 9.2. The third row shows the superimposed images after the rotation and scaling correction while the fourth row displays the images after the additional drift correction.*

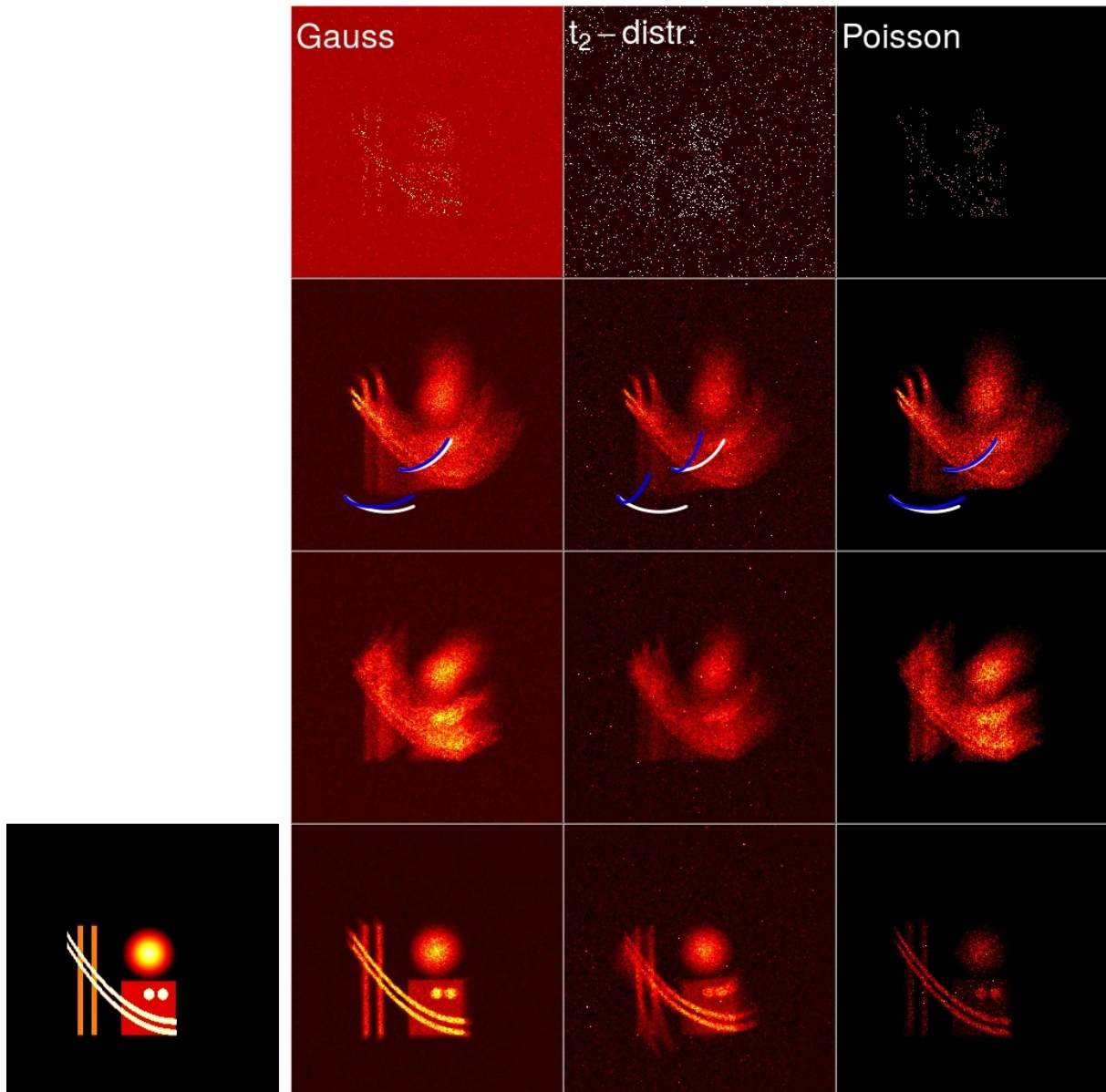


Figure 9.5: *Left (fourth row): Test image \tilde{f} with grey scale values (rescaled to the unit interval), represented by colours ranging from black (0) over red and yellow to white (1). Right: The first row shows the first frames of sequences of $T = 50$ noisy images subject to quadratic motion functions (from left to right: Gaussian noise, Student- t_2 noise, Poisson model). The second row depicts the corresponding superimposed images (i.e., the averages over all frames in the sequence). The true motion curves of two pixels are shown as white curve segments on top of which we plot the estimated motion in blue. The true and the estimated parameters are reported in the Tables 9.1 and 9.2. The third row shows the superimposed images after the rotation and scaling correction while the fourth row displays the images after the additional drift correction.*

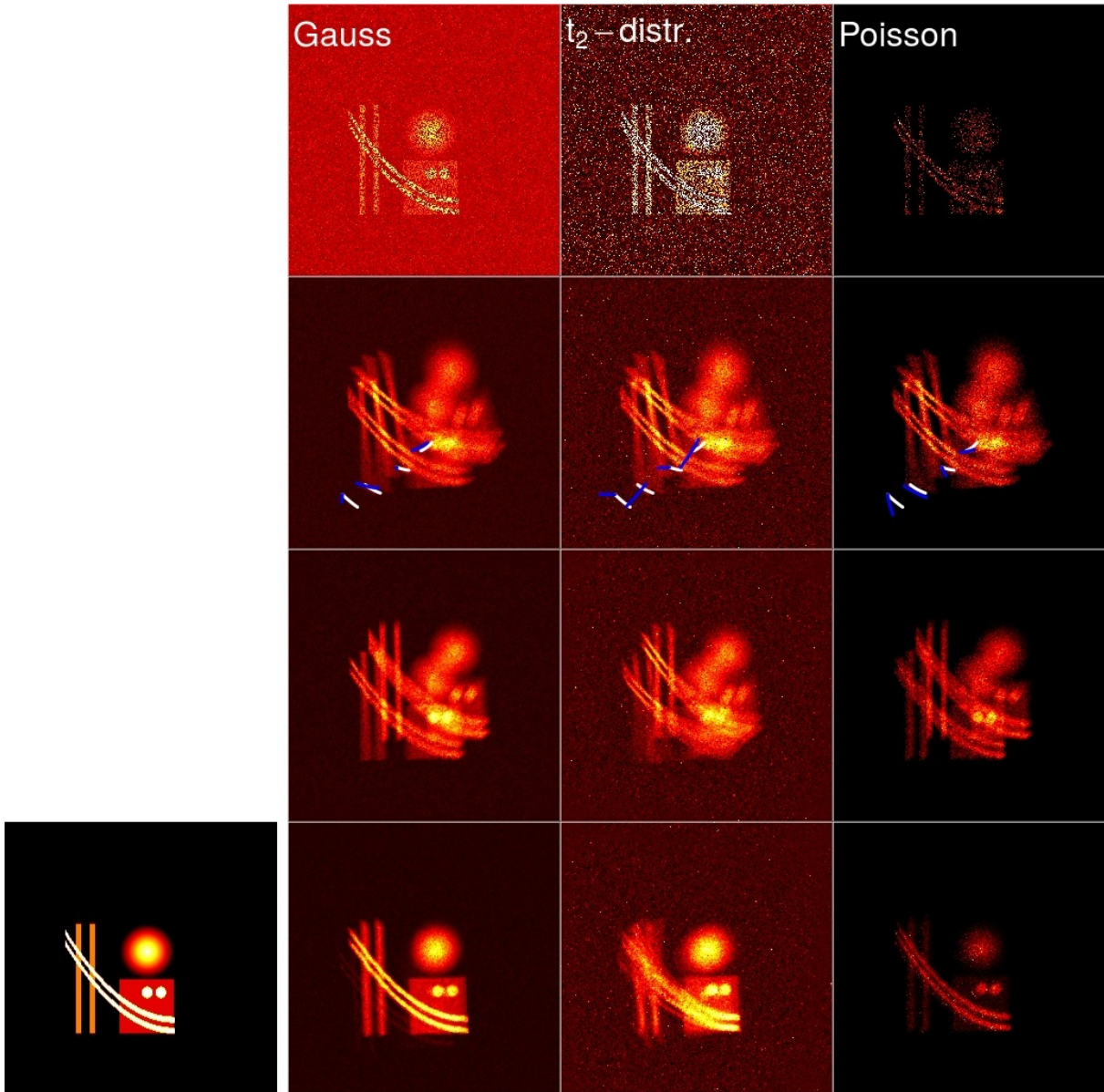


Figure 9.6: *Left (fourth row): Test image \tilde{f} with grey scale values (rescaled to the unit interval), represented by colours ranging from black (0) over red and yellow to white (1). Right: The first row shows the first frames of sequences of $T = 20$ noisy images subject to piecewise linear motion functions with a common change point (from left to right: Gaussian noise, Student- t_2 noise, Poisson model). The second row depicts the corresponding superimposed images (i.e., the averages over all frames in the sequence). The true motion curves of two pixels are shown as white curve segments. The left pixel moves to the right and down, then jumps a little to the right and up and continues its motion, as depicted with the two leftmost curve segments. The two curve segments to the right show the similar motion of the second pixel. On top of these white curves we plot the estimated motion in blue. The true and the estimated parameters are reported in the Tables 9.1 and 9.2. The third row shows the superimposed images after the rotation and scaling correction while the fourth row displays the images after the additional drift correction.*

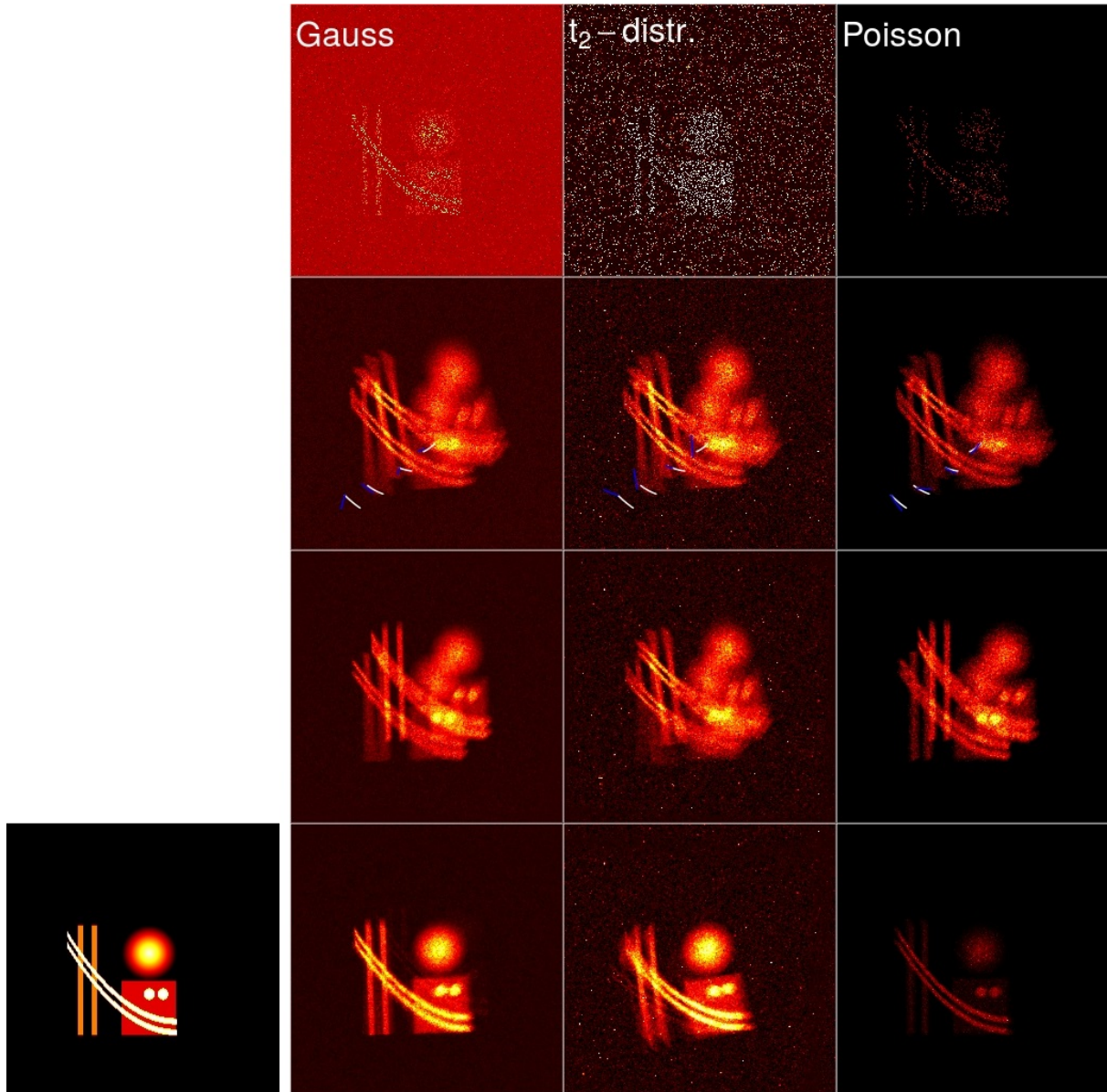


Figure 9.7: *Left (fourth row): Test image \tilde{f} with grey scale values (rescaled to the unit interval), represented by colours ranging from black (0) over red and yellow to white (1). Right: The first row shows the first frames of sequences of $T = 50$ noisy images subject to piecewise linear motion functions with a common change point (from left to right: Gaussian noise, Student- t_2 noise, Poisson model). The second row depicts the corresponding superimposed images (i.e., the averages over all frames in the sequence). The true motion curves of two pixels are shown as white curve segments. The left pixel moves to the right and down, then jumps a little to the right and up and continues its motion, as depicted with the two leftmost curve segments. The two curve segments to the right show the similar motion of the second pixel. On top of these white curves we plot the estimated motion in blue. The true and the estimated parameters are reported in the Tables 9.1 and 9.2. The third row shows the superimposed images after the rotation and scaling correction while the fourth row displays the images after the additional drift correction.*

increasing degree of the polynomials, the mean squared error increases. Nevertheless, visual inspection of the estimated images in Figure 9.5 exhibits good reconstruction quality for the Gaussian and Poisson models.

To evaluate our motion correction we use versions of the motion blur measure m_2 proposed by Xu et al. (2013) which is based on the work of Chen et al. (2010) (for details see Chapter 7). The motion blur values of the superimposed images, the corresponding rotation and scaling corrected images, and the final estimated images are reported in the Tables 9.7 and 9.8. Evidently, the blurring is reduced in every correction step, although visual inspection of the resulting images shows that in the case of t_2 -distributed noise, the reconstruction is not very good for motion models which are more complex than linear.

Motion model with jump. Finally, in order to analyse the robustness of our method, for example, when a smooth motion abruptly jumps due to an external shock, we consider a piecewise linear motion model with a jump at an unknown time,

$$\begin{aligned}\delta_t(\theta) &= \begin{cases} (\theta_{11}, \theta_{21})^\top t & \text{if } t \leq t_0, \\ (\theta_{12}, \theta_{22})^\top (t - t_0) + (\theta_{13}, \theta_{23})^\top & \text{if } t > t_0, \end{cases} \\ \rho_t(\phi) &= \begin{cases} \phi_1 t & \text{if } t \leq t_0, \\ \phi_2(t - t_0) + \phi_3 & \text{if } t > t_0, \end{cases} \\ \sigma_t(\alpha) &= \begin{cases} 1 + \alpha_1 t & \text{if } t \leq t_0, \\ 1 + \alpha_2(t - t_0) + \alpha_3 & \text{if } t > t_0, \end{cases}\end{aligned}$$

with $(\phi_1, \phi_2, \phi_3, \alpha_1, \alpha_2, \alpha_3, t_0) \in \Phi \times \mathbb{A} \times [0, 1] \subset \mathbb{R}^7$ and $(\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23}) \in \Theta \subset \mathbb{R}^6$, that is, the rotation function jumps to the angle ϕ_3 at the unknown time point t_0 and so on. As mentioned before, this type of motion does not meet our assumptions, for example, the Lipschitz properties in the Assumptions 6.10 and 6.31 are not fulfilled as one can easily see by perturbing the parameter t_0 . For these simulations, we double the Bernoulli probability to $p = 10/T$.

For the simulation, we choose the drift parameters $(\theta_{11}, \theta_{21}) = (20/512, 10/512)$, $(\theta_{12}, \theta_{22}) = (15/512, 30/512)$, $(\theta_{13}, \theta_{23}) = (40/512, 36/512)$, the rotation angle parameters $(\phi_1, \phi_2, \phi_3) = (\pi/8, \pi/8, \pi/40)$, and the scaling parameters $(\alpha_1, \alpha_2, \alpha_3) = (0.1, 0.2, 0.07)$, as well as the jump time $t_0 = 0.4$. We estimate the jump time t_0 together with the rotation and scaling parameters and take the resulting estimate as given for the drift estimation. Once again, we use the Gaussian noise, the t -distributed noise with 2 degrees of freedom, and the Poisson model. The superimposed images as well as their reconstructions are visualized in Figures 9.6 and 9.7. In the Tables 9.1 and 9.2, the estimation results are summarized. The blur values are reported in the Tables 9.7 and 9.8.

Computational time. For polynomial motion, simulating a sample and computing the estimates required about 2 minutes for $T = 20$ frames on a Core AMD Opteron with 2.6 GHz, where the most expensive part is the rotation and scaling of the images, which means that the computational time scales almost linearly with the number of frames T . The numerical optimization itself takes about 10 seconds. For the motion with jump, we considered jump times \hat{t}_0 on the grid $\{2/T, \dots, 98/T\}$ and, given \hat{t}_0 , minimized the contrast functional w.r.t. (ϕ, α) to find the estimator for (ϕ_0, α_0, t_0) with overall minimal contrast. This leads to higher computational times between about 3 minutes ($T = 20$) and 30 minutes ($T = 100$). Our simulations show that the proposed estimation method works well and significantly reduces blurring. This has been demonstrated for a polynomial drift even if we observe just a small part of the shifted image at every time point. We

have also obtained good results for reconstruction in a Poisson model. The simulations with t_2 -distributed noise show that the estimation method is not very robust against outliers, especially if the observed images are very sparse. Finally, we studied the case of a piecewise linear drift with a jump at an unknown time point, that is, a discontinuous drift. Although the Assumptions 6.10 and 6.31 are not satisfied in this case, we found that even in this setting our estimator performs quite well.

par. ($\phi_0; \alpha_0; t_0$)	linear rot./scal. (0.393; 0.278; -)	quadratic rot./scal. (0.196, 0.393; 0.15, 0.1; -)	cubic rot./scal. (0.196, 0.098, 0.491; 0, 0.1, 0.2; -)	rot./scal. with jump (0.393, 0.393, 0.079; 0.1, 0.2, 0.07; 0.4)
error type	$(\hat{\phi}_T; \hat{\alpha}_T)$	$(\hat{\phi}_T; \hat{\alpha}_T)$	$(\hat{\phi}_T; \hat{\alpha}_T)$	$(\hat{\phi}_T; \hat{\alpha}_T)$
Gaussian	20 (0.406; 0.155) 50 (0.405; 0.119) 100 (0.328; 0.091)	(0.317, 0.280; 0.322, -0.201) (0.168, 0.388; 0.221, -0.129) (0.032, 0.001; 0.154, -0.067)	(0.145, 0.829, -0.257; 0.429, -0.579, 0.324) (0.120, 0.873, -0.753; 0.658, -1.235, 0.769) (0.332, 0.109, 0.131; 0.554, -1.002, 0.631)	(0.231, 0.347, 0.064; 0.145, 0.076, 0.052; 0.35) (0.192, 0.345, 0.058; 0.207, 0.067, 0.070; 0.44) (-0.114, 0.210, 0.017; 0.135, 0.047, 0.047; 0.42)
t -distr.	20 (0.433; 0.143) 50 (0.312; 0.059) 100 (0.046; 0.108)	(-0.049, 0.079; 0.302, -0.177) (0.160, 0.052; 0.260, -0.171) (-0.041, 0.024; 0.126, -0.055)	(0.120, -0.076, 0.037; 0.253, -0.106, -0.026) (0.027, -0.056, -0.015; 0.189, -0.099, 0.023) (-0.058, 0.034, 0.033; 0.038, -0.037, 0.062)	(-0.212, 0.096, 0.158; 0.290, 0.049, 0.094; 0.2) (-0.317, 0.044, 0.046; 0.283, 0.063, 0.054; 0.32) (-0.073, -0.045, -0.063; 0.106, 0.119, 0.101; 0.41)
Poisson	20 (0.454; 0.154) 50 (0.368; 0.110) 100 (0.345; 0.092)	(0.357, 0.196; 0.335, -0.214) (0.181, 0.413; 0.207, -0.106) (0.057, 0.028; 0.137, -0.052)	(0.393, 0.431, -0.057; 0.373, -0.462, 0.249) (0.220, 1.342, -1.400; 0.633, -1.182, 0.749) (-0.749, 0.044, 0.695; 0.625, -1.157, 0.730)	(0.416, 0.424, 0.072; 0.199, 0.067, 0.073; 0.40) (0.349, 0.310, 0.103; 0.197, 0.037, 0.065; 0.40) (0.489, 0.176, 0.139; 0.111, 0.066, 0.036; 0.40)

Table 9.1: Displaying the estimated $(\hat{\phi}_T, \hat{\alpha}_T)$ for one simulation in different rotation/scaling models. We have considered image sequences with $T \in \{20, 50, 100\}$ time points as well as Gaussian and Student- t_2 error models with variance 0.1^2 and a Poisson model as explained in detail in the text.

true parameter	θ_0	linear drift	quadratic drift	cubic drift	drift with jump
error type	T	θ_T	θ_T	θ_T	θ_T
Gaussian	20 (0.099; 0.058) 50 (0.097; 0.056) 100 (0.096; 0.063)	(0.053, 0.065; 0.024, 0.65) (0.034, 0.081; 0.027, 0.064) (0.066, -0.001; 0.000, 0.163)	(0.059, 0.059, 0.020, 0.078)	(0.039, 0.000, 0.049, 0.019, 0.039, 0.009)	(-0.035, 0.055, 0.047; 0.035, 0.048, 0.070) (-0.031, 0.001, 0.063; 0.011, 0.029, 0.072) (-0.034, -0.014, 0.062; -0.001, 0.054, 0.072)
t -distr.	20 (0.099; 0.055) 50 (0.094; 0.063) 100 (0.087; 0.103)	(0.834, -0.018; 0.014, 0.145) (0.068, 0.023; -0.006, 0.145) (0.105, -0.049; -0.012, 0.185)	(0.056, 0.018, -0.042; 0.029, 0.032, 0.067) (0.045, 0.038, -0.075; 0.022, 0.048, 0.063) (0.047, 0.025, -0.051; 0.010, 0.040, 0.087)	(-0.121, 0.088, 0.016; 0.014, 0.110, 0.026) (0.012, -0.005, 0.069; 0.039, 0.112, 0.047) (0.046, -0.022, 0.071; 0.048, 0.116, 0.081)	
Poisson	20 (0.099; 0.053) 50 (0.098; 0.061) 100 (0.095; 0.062)	(0.061, 0.051; 0.017, 0.073) (0.039, 0.076; 0.026, 0.059) (0.113, -0.032; -0.006, 0.163)	(0.020, 0.047, 0.011; 0.014, 0.031, 0.010) (0.040, 0.020, -0.024; -0.003, 0.024, 0.080) (-0.003, 0.031, -0.027; 0.070, 0.027, 0.028)	(-0.034, 0.030, 0.059; 0.003, 0.026, 0.067) (0.037, 0.021, 0.080; 0.019, 0.055, 0.060) (0.034, 0.014, 0.078; 0.012, 0.084, 0.057)	

Table 9.2: Displaying the estimated $\hat{\theta}_T$ for one simulation in different drift models. We have considered image sequences with $T \in \{20, 50, 100\}$ time points as well as Gaussian and Student- t_2 error models with variance 0.1^2 and a Poisson model as explained in detail in the text.

par. ($\phi_0; \alpha_0; t_0$)	linear rot./scal.	quadratic rot./scal.	cubic rot./scal.	rot./scal. with jump
T	$(\hat{\phi}_T; \hat{\alpha}_T)$	$(\hat{\phi}_T; \hat{\alpha}_T)$	$(\hat{\phi}_T; \hat{\alpha}_T)$	$(\hat{\phi}_T; \hat{\alpha}_T; (\hat{t}_0)_T)$
Gaussian	20 (0.413; 0.154) 50 (0.381; 0.119) 100 (0.328; 0.091)	(0.278, 0.310; 0.338, -0.202) (0.268, 0.294; 0.224, -0.127) (0.040, 0.054; 0.164, -0.079)	(0.211, 0.583, -0.045; 0.366, -0.416, 0.208) (0.063, 0.206, -0.245; 0.423, -0.645, 0.382) (0.382, 0.327, -0.479; 0.399, -0.650, 0.408)	(0.357, 0.363, 0.107; 0.141, 0.061, 0.065; 0.445) (0.129, 0.281, 0.059; 0.219, 0.060, 0.066; 0.378) (0.123, 0.169, 0.076; 0.136, 0.049, 0.051; 0.418)
t -distr.	20 (0.335; 0.162) 50 (0.182; 0.110) 100 (0.017; 0.106)	(0.055, 0.074; 0.328, -0.216) (0.037, 0.028; 0.261, -0.156) (0.016, -0.021; 0.117, -0.045)	(0.123, 0.004, -0.031; 0.249, -0.083, -0.034) (0.037, -0.005, 0.008; 0.260, -0.194, 0.060) (-0.011, -0.029, -0.015; 0.148, -0.028, 0.002)	(-0.277, 0.096, 0.158; 0.548, 0.056, 0.092; 0.225) (-0.195, 0.040, -0.019; 0.290, 0.063, 0.074; 0.350) (-0.130, 0.036, 0.020; 0.121, 0.065, 0.033; 0.382)
Poisson	20 (0.413; 0.153) 50 (0.362; 0.113) 100 (0.308; 0.095)	(0.325, 0.250; 0.382, -0.257) (0.203, 0.309; 0.232, -0.129) (0.085, 0.028; 0.140, -0.053)	(0.271, 0.264, -0.037; 0.337, -0.306, 0.113) (0.087, 0.373, -0.286; 0.447, -0.695, 0.413) (0.168, 0.252, -0.641; 0.560, -0.013, 0.640)	(-0.062, 0.317, 0.051; 0.391, 0.060, 0.079; 0.300) (0.349, 0.311, 0.103; 0.198, 0.037, 0.066; 0.400) (0.489, 0.176, 0.139; 0.111, 0.066, 0.037; 0.400)

Table 9.3: Setting as in Table 9.1. Displaying the means of the estimators $(\hat{\phi}_T; \hat{\alpha}_T)$ from 100 simulations each.

true parameter	θ_0	linear drift	quadratic drift	cubic drift	drift with jump
error type	T	$\hat{\theta}_T$	$\hat{\theta}_T$	$\hat{\theta}_T$	$\hat{\theta}_T$
Gaussian	20 (0.098; 0.059) 50 (0.098; 0.059) 100 (0.094; 0.066)	(0.058, 0.059; 0.024, 0.064) (0.042, 0.074; 0.020, 0.071) (0.080, -0.005; -0.011, 0.165)	(0.030, 0.043, 0.005; 0.018, 0.033, 0.008) (0.051, 0.036, -0.068; 0.025, 0.026, 0.076) (0.047, 0.034, -0.035; 0.004, 0.012, 0.093)	(0.017, 0.020, 0.076; 0.017, 0.053, 0.066) (-0.006, 0.015, 0.065; 0.028, 0.061, 0.065) (0.006, 0.008, 0.069; 0.017, 0.064, 0.071)	
t -distr.	20 (0.094; 0.068) 50 (0.086; 0.084) 100 (0.079; 0.107)	(0.085, -0.009; 0.009, 0.139) (0.083, -0.010; -0.004, 0.161) (0.090, -0.027; -0.007, 0.176)	(0.052, 0.032, -0.052; 0.026, 0.042, 0.059) (0.045, 0.039, -0.063; 0.027, 0.039, 0.063) (0.050, 0.034, -0.076; 0.023, 0.051, 0.064)	(-0.093, 0.037, 0.024; 0.045, 0.108, 0.039) (-0.009, -0.014, 0.060; 0.049, 0.096, 0.067) (0.008, -0.001, 0.064; 0.019, 0.098, 0.067)	
Poisson	20 (0.098; 0.059) 50 (0.097; 0.062) 100 (0.095; 0.067)	(0.063, 0.051; 0.020, 0.069) (0.055, 0.057; 0.014, 0.084) (0.093, -0.013; -0.014, 0.167)	(0.038, 0.040, -0.016; 0.018, 0.040, 0.026) (0.042, 0.045, -0.051; 0.024, 0.018, 0.073) (0.053, 0.071, -0.141; 0.025, 0.020, 0.089)	(0.013, 0.046, 0.053; 0.008, 0.095, 0.041) (-0.001, 0.025, 0.068; 0.036, 0.054, 0.068) (0.000, 0.008, 0.066; 0.012, 0.084, 0.057)	

Table 9.4: Setting as in Table 9.2. Displaying the means of the estimators $\hat{\theta}_T$ from 100 simulations each.

	Gaussian noise			t_2 noise			Poisson model		
	$T = 20$	$T = 50$	$T = 100$	$T = 20$	$T = 50$	$T = 100$	$T = 20$	$T = 50$	$T = 100$
Linear rot./scal.	1e-2	2e-2	3e-2	3e-2	1e-1	1e-1	1e-2	2e-2	4e-2
Quadratic rot./scal.	1e-1	9e-2	1e-1	2e-1	2e-1	2e-1	2e-1	9e-2	1e-1
Cubic rot./scal.	1e-0	1e-0	2e-0	4e-1	6e-1	4e-1	8e-1	2e-0	3e-0
Rot./scal. with jump	4e-2	1e-1	1e-1	1e-0	6e-1	4e-1	6e-1	7e-2	2e-1

Table 9.5: Root of the mean squared error $\mathbb{E} \left\| (\hat{\phi}_T - \phi_0, \hat{\alpha}_T - \alpha_0) \right\|^2$ of the estimators $(\hat{\phi}_T, \hat{\alpha}_T)$ from 100 simulations each.

	Gaussian noise			t_2 noise			Poisson model		
	$T = 20$	$T = 50$	$T = 100$	$T = 20$	$T = 50$	$T = 100$	$T = 20$	$T = 50$	$T = 100$
Linear drift	9e-5	8e-5	3e-5	3e-4	1e-3	2e-3	1e-4	6e-5	3e-5
Quadratic drift	5e-4	7e-4	1e-2	1e-2	1e-2	2e-2	3e-4	1e-3	1e-2
Cubic drift	4e-3	2e-2	2e-2	1e-2	1e-2	2e-2	7e-3	2e-2	6e-2
Drift with jump	4e-3	5e-3	5e-3	3e-2	1e-2	7e-3	1e-2	4e-3	5e-3

Table 9.6: Root of the mean squared error $\mathbb{E} \left\| \hat{\theta}_T - \theta_0 \right\|^2$ of the estimators $\hat{\theta}_T$ from 100 simulations each.

	Gaussian noise			t_2 noise			Poisson model			
	$T = 20$	$T = 50$	$T = 100$	$T = 20$	$T = 50$	$T = 100$	$T = 20$	$T = 50$	$T = 100$	
SI	Linear rot./scal.	-0.950	-0.960	-0.950	-0.955	-0.933	-0.944	-0.952	-0.954	-0.952
	Quadratic rot./scal.	-0.410	-0.413	-0.411	-0.410	-0.419	-0.397	-0.413	-0.409	-0.419
	Cubic rot./scal.	0.183	0.181	0.183	0.186	0.181	0.180	0.182	0.182	0.183
\hat{f}_T^*	Rot./scal. with jump	-0.705	-0.704	-0.704	-0.699	-0.704	-0.714	-0.708	-0.704	-0.706
	Linear rot./scal.	-1.055	-1.049	-1.017	-1.066	-1.011	-0.970	-1.057	-1.030	-1.013
	Quadratic rot./scal.	-0.494	-0.465	-0.425	-0.434	-0.451	-0.413	-0.488	-0.468	-0.427
Rot./scal. with jump	Cubic rot./scal.	0.100	0.129	0.115	0.169	0.170	0.174	0.089	0.113	0.152
		-0.758	-0.768	-0.738	-0.770	-0.733	-0.734	-0.779	-0.772	-0.774

Table 9.7: Blur measure values (of the log-polar transformed Fourier magnitudes) of the superimposed images (SI) and the rotation and scaling corrected images \hat{f}_T^* . The corresponding estimators ($\hat{\phi}_T, \hat{\alpha}_T$) are reported in Table 9.1. The images for linear rotation/scaling, quadratic rotation/scaling, and rotation/scaling with jump and $T \in \{20, 50\}$ are shown in Figures 9.2 through 9.7.

	Gaussian noise			t_2 noise			Poisson model			
	$T = 20$	$T = 50$	$T = 100$	$T = 20$	$T = 50$	$T = 100$	$T = 20$	$T = 50$	$T = 100$	
\hat{f}_T^*	Linear drift	0.037	0.033	0.065	0.015	0.040	-0.039	-0.022	0.034	0.048
	Quadratic drift	-0.004	0.040	-0.425	-0.031	-0.011	-0.095	0.026	-0.008	-0.076
	Cubic drift	-0.042	-0.207	-0.059	-0.114	-0.108	-0.065	-0.008	-0.192	-0.174
\hat{f}_T	Drift with jump	-0.366	-0.302	-0.651	-0.383	-0.312	-0.269	-0.174	-0.198	-0.457
	Linear drift	-1.178	-1.120	-0.878	-0.807	-0.244	-0.179	-0.758	-0.631	-0.833
	Quadratic drift	-1.254	-1.112	-0.477	-0.203	-0.301	-0.185	-0.705	-0.868	-0.375
Drift with jump	Cubic drift	-0.996	-0.579	-0.728	-0.276	-0.140	-0.072	-0.514	-0.608	-0.037
		-1.162	-0.669	-0.349	-0.176	-1.188	-0.963	-0.560	-1.084	-1.101

Table 9.8: Blur measure values of the rotation and scaling corrected images \hat{f}_T^* and the estimated images \hat{f}_T . The corresponding estimators $\hat{\phi}_T$ are reported in Table 9.2. The images for linear drift, quadratic drift, and drift with jump and $T \in \{20, 50\}$ are shown in Figures 9.2 through 9.7.

Chapter 10

Application to SMS data

In this chapter, we demonstrate how the estimation method proposed in Chapter 4 can be used to process SMS nanoscopy data. In particular, we address suitable choices for the drift, rotation, and scaling models δ_t^θ , ρ_t^ϕ , and σ_t^α as well as computational issues.

We used a standard SMS-setup for this study (see e.g., Geisler et al. (2012); Hartmann et al. (2016)) which was equipped with a home-built stable sample holder ensuring that the sample drift is well below the expected average localization accuracy. For image acquisition, a series of $T \in \{29,000; 35,000\}$ frames was taken from a microtubule network (β -tubulin) of HeLa-cells with a frame exposure time of 15 ms, resulting in a total image acquisition time of about 7.25 and 8.75 minutes, respectively. During this time an experimental drift and rotation were applied by moving the sample with respect to the objective lens.

The lateral positions of the fluorescent markers were then calculated from the single frames by a mask-fitting of the respective Airy spot (Thompson et al., 2002). These locations were tabulated together with the respective time of detection $t \in \mathbb{T}$.

We analyse two data sets (networks I and II) from β -tubulin networks in HeLa-cells. The position histogram of the first data set is shown in Figure 1.2. It contains 1,243,170 positions recorded in 29,000 frames which are distributed over an area of approximately $32 \mu\text{m} \times 30 \mu\text{m}$. The second dataset (see Figure 10.3) contains 1,074,516 positions recorded in 35,000 frames. The data of this set are distributed over an area of about $49 \mu\text{m} \times 42 \mu\text{m}$. The positions of the fiducials were used to compare the quality of our method to the current state of the art of rigid motion correction.

To analyse the data with our M-estimation method we create $T = 2000$ position histograms of $n = N^2 = 512^2$ bins of the first data set and $T = 500$ position histograms of $n = N^2 = 512^2$ bins of the second data set, that is, we look at $T = 2000$ position histograms which are composed of the data points of $\beta_T = 14$ frames each (see Figure 1.1) and $T = 500$ position histograms composed of $\beta_T = 70$ frames each. Our empirical analysis shows that the estimates are not strongly influenced by the choices of T and N , however too small values circumvent the registration of small movements and for large values computational problems arise in terms of speed. This is in accordance with our previous simulation results.

As exemplarily demonstrated in Figure 10.1 for the second data set shown in Figure 10.3, the number of recorded markers n_t varies slightly as the experiment continues. Since switched on markers bleach after emitting light, one has to increase the switching laser intensity occasionally to get a roughly constant number of observations per frame, which was quite successful in this experiment.

Note that in the experiments, fiducial markers were included into the sample, that is,

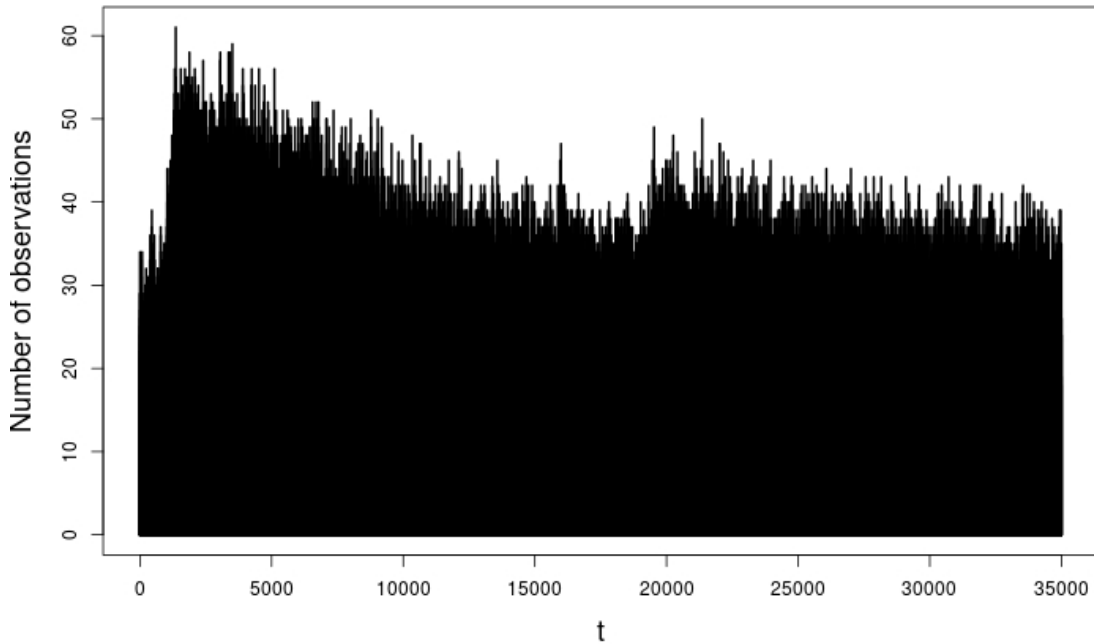


Figure 10.1: *Number of data points (registered markers) n_t per frame (network II)*

a persistent fluorescence source, which enables us to track the motion easily, for testing purposes (see Chapter 8). We stress that this is currently state of the art technology to align SMS images over time (see the Introduction). In order to investigate the validity of our method, we delete the data originating from the fiducial markers from the observed sample and use it for verification only.

First, we try the cross correlation method described in Chapter 8 on the data from network II. It turns out that this method relies heavily on rather complete images to estimate the rotation and scaling of the images which is not the case in pure drift correction (Geisler et al., 2012). Figure 10.2 shows the correction based on a binned sequence of 5 binned frames (i.e., $\beta_T = 7,000$). Even halving the bin size to $\beta_T = 3,500$ lead to “reconstructions” that were far blurrier than the uncorrected image. The reconstruction in Figure 10.2 seems to improve the image resolution a little but due to the large value of β_T , the motion blur in each binned frame is too large to enable a satisfactory reconstruction.

As reported in Tables 10.1 and 10.2, we apply three different motion models to the second data set and choose the one with the smallest contrast values. Moreover, we display the corresponding motion blur values m_2 (see Chapter 7) to compare the image estimators with fiducial marker tracking. The time required for the computation of $\hat{\theta}_T$, $\hat{\phi}_T$, and $\hat{\alpha}_T$ may last up to about 45 minutes (on a Core AMD Opteron with 2.6 GHz), depending on the bin width. Since we do not know the true motion functions (as was the case in the simulation study, Chapter 9), we determine the motion direction as the average motion direction obtained from fiducial marker tracking (once for rotation and log-scaling and once more for drift, see Chapters 7 and 8 for details). In particular, in the Tables 10.1 and 10.2, we report the m_2 -value for the correction via fiducial tracking, too. The result indicates that our estimation method is competitive with tracking of the fiducial movement. The reconstructions of the image for fiducial tracking and a linear motion model are displayed in Figure 10.4.

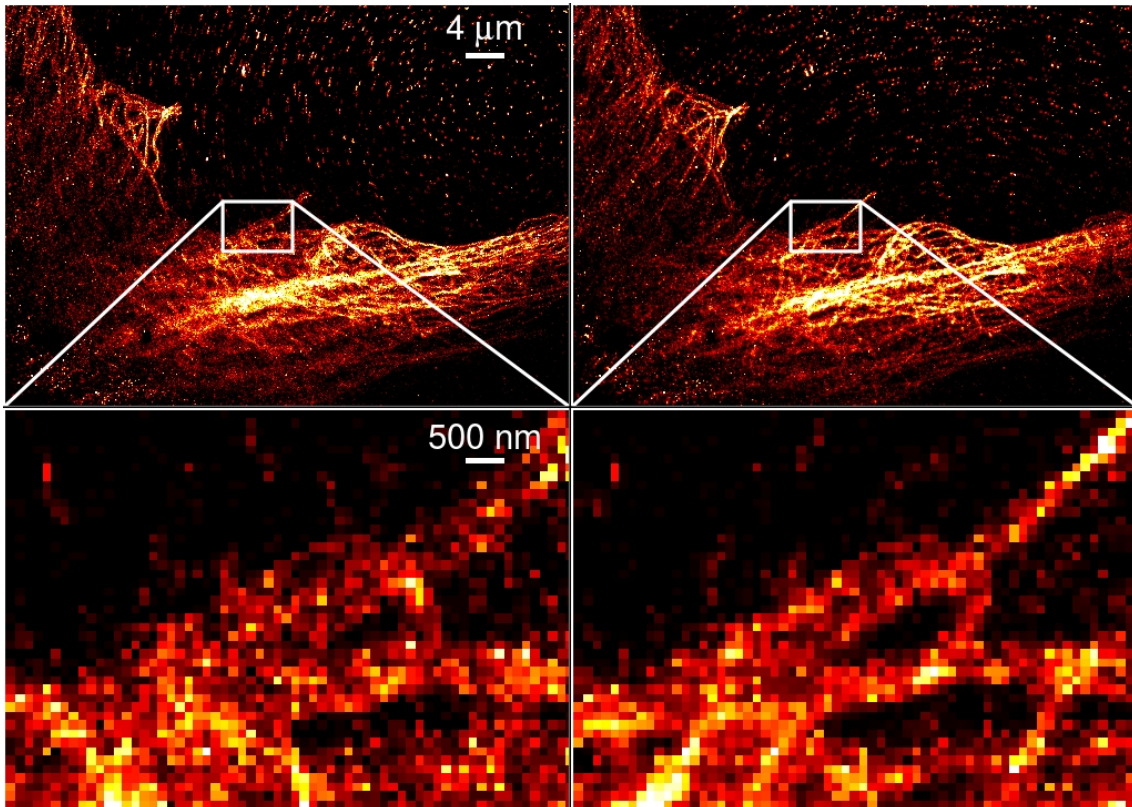


Figure 10.2: *First row: motion blurred position histogram of network II with $n = 512^2$ bins (left) and reconstruction using cross correlation with 5 binned frames, that is, with bin size $\beta_T = 7,000$ (right). Second row: detailed views inside the white boxes.*

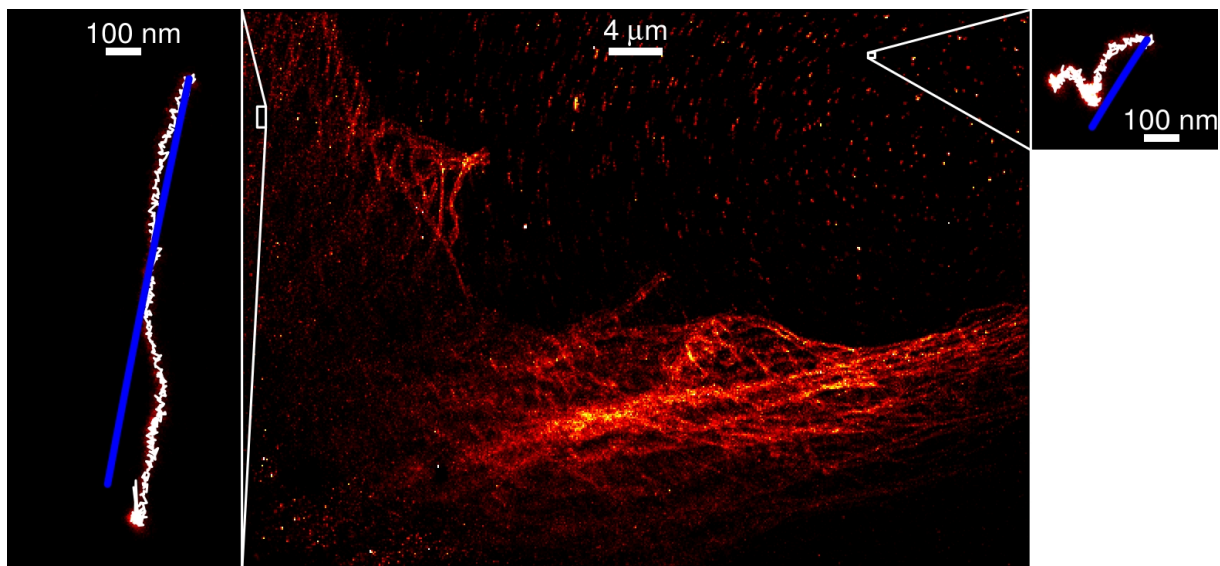


Figure 10.3: *Motion blurred position histogram of network II with $n = 512^2$ bins (middle) and close up on two areas with fiducials (left and right). These fiducials were used for tracking. Their estimated positions are drawn over the fiducial data as white curves. The blue curves demonstrate the estimated linear motion (our M -estimator) of the two fiducial starting positions according to a linear motion model.*

rot./scal. models		rot.			scal.			contrast	motion blur
		$\hat{\phi}_{T;3}$	$\hat{\phi}_{T;2}$	$\hat{\phi}_{T;1}$	$\hat{\alpha}_{T;3}$	$\hat{\alpha}_{T;2}$	$\hat{\alpha}_{T;1}$	M_T	\tilde{m}_2
linear	linear	-	-	0.018	-	-	0.003	273890e13	2.124
quadratic	quadratic	-	0.003	0.014	-	-0.005	0.014	273891e13	2.125
cubic	cubic	-0.002	-0.009	0.023	-0.002	0.004	0.001	273893e13	2.125
fiducial tracking								-	2.108
superimposed image								273895e13	2.131

Table 10.1: *Rotation and scaling estimation results for the β -tubulin network I in Figure 1.2 for several drift models.*

drift models		x_1 -dir.			x_2 -dir.			contrast	motion blur
x_1 -dir.	x_2 -dir.	$\hat{\theta}_{T;13}$	$\hat{\theta}_{T;12}$	$\hat{\theta}_{T;11}$	$\hat{\theta}_{T;23}$	$\hat{\theta}_{T;22}$	$\hat{\theta}_{T;21}$	N_T	m_2
linear	linear	-	-	0.001	-	-	-0.015	125e6	-0.366
quadratic	quadratic	-	-0.002	0.007	-	0.013	-0.026	125e6	-0.352
cubic	cubic	-0.014	0.024	-0.006	0.001	0.004	-0.020	126e6	-0.349
fiducial tracking								-	-0.408
rot./scal.-correct. image								128e6	-0.057

Table 10.2: *Drift estimation results for the β -tubulin network II with fiducial markers, see Figure 10.3, for several drift models. The displayed motion blur values are for the respective images with fiducial markers removed.*

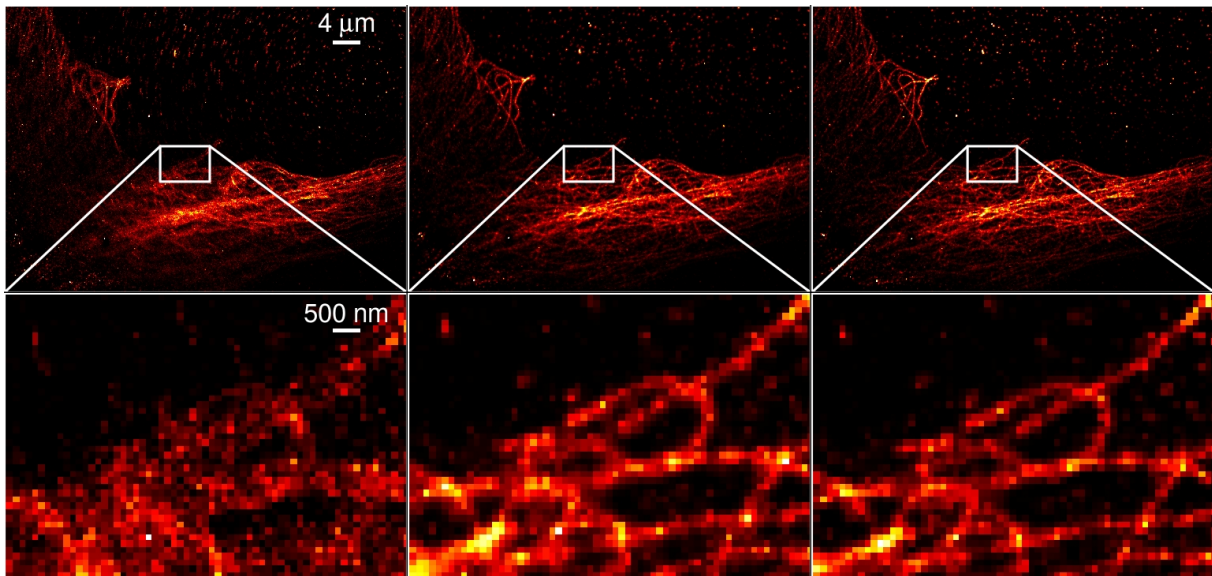


Figure 10.4: *Drift blurred network II (top left), with the assumption of a linear motion estimated image (top center), and by fiducial marker tracking corrected image (top right), as well as detailed views inside the white boxes (bottom row)*

Chapter 11

Bootstrap confidence bands

Given the parameter and image estimators $\hat{\theta}_T$, $\hat{\phi}_T$, $\hat{\alpha}_T$, and \hat{f}_T from the Definitions 4.3, 4.8, and 4.9 and thus estimators $(\delta_t^{\hat{\theta}_T})_1$, $(\delta_t^{\hat{\theta}_T})_2$, $\rho_t^{\hat{\phi}_T}$, and $\sigma_t^{\hat{\alpha}_T}$ for the drift, rotation, and scaling functions $(\delta_t^{\theta_0})_1$, $(\delta_t^{\theta_0})_2$, $\rho_t^{\phi_0}$, and $\sigma_t^{\alpha_0}$, we can construct bootstrap confidence bands for these functions using the method described in Hall and Pittelkow (1990). Here, we give a short summary of that method with our application to motion functions in mind, using the drift function in x_1 -direction as an example (the confidence bands for the other functions are computed in the exactly same way). For notational simplicity following (3.4), we index the spatial location by a single index $j \in \{1, \dots, n\}$. Note that this method assumes homoscedasticity (i.e., $(\nu_j^t)^2 \equiv \nu^2 > 0$) and we use it for simplicity's sake. To account for heteroscedasticity, one could make use of a wild bootstrap procedure (see e.g., Wu (1986); Liu (1988); Mammen (1993)). First, we consider the standardized difference $\Delta_t := ((\delta_t^{\hat{\theta}_T})_1 - (\delta_t^{\theta_0})_1) / \hat{\nu}$, where $\hat{\nu}$ is the empirical standard deviation of the *residuals*

$$r_j^t := O_j^t - \hat{f}_T\left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}}(x_j - \delta_t^{\hat{\theta}_T})\right), \quad t \in \mathbb{T}, j \in \{1, \dots, n\}, \quad (11.1)$$

and thus an estimator for the standard deviation of the errors

$$\epsilon_j^t = O_j^t - f\left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(x_j - \delta_t^{\theta_0})\right), \quad t \in \mathbb{T}, j \in \{1, \dots, n\}.$$

Obviously, constructing a confidence band for $(\delta_t^{\theta_0})_1$ is equivalent to constructing one for Δ_t . Next we choose the shape of the confidence band in terms of two functions $g_+, g_- : [0, 1] \rightarrow [0, \infty)$ such that $(\delta_t^{\hat{\theta}_T})_1 + \hat{\nu}u_+g_+(t)$ and $(\delta_t^{\hat{\theta}_T})_1 - \hat{\nu}u_-g_-(t)$ represent the upper and lower border, respectively, of the confidence band for $(\delta_t^{\theta_0})_1$, with appropriate positive numbers u_+, u_- . For a confidence level $\eta \in (0, 1)$ we minimize $u_+ + u_-$ under the constraint

$$\mathbb{P}\left((\delta_t^{\theta_0})_1 \in [(\delta_t^{\hat{\theta}_T})_1 + \hat{\nu}u_+g_+(t), (\delta_t^{\hat{\theta}_T})_1 - \hat{\nu}u_-g_-(t)] \text{ for all } t \in [0, 1]\right) \geq 1 - \eta,$$

or, equivalently under

$$\mathbb{P}(\Delta_t \in [-u_+g_+(t), u_-g_-(t)] \text{ for all } t \in [0, 1]) \geq 1 - \eta.$$

Since the distribution of Δ_t is unknown, we approximate it by bootstrapping $B \in \mathbb{N}$ times from the residuals (11.1), that is, for every $b \in \{1, \dots, B\}$ and every $j \in \{1, \dots, n\}$, $t \in \mathbb{T}$, we draw $(\epsilon_j^t)^{(b)}$ independently with replacement from the set of all residuals $\{r_{j'}^{t'} \mid j' \in \{1, \dots, n\}, t' \in \mathbb{T}\}$. We thus obtain

$$(O_j^t)^{(b)} := \hat{f}_T\left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}}(x_j - \delta_t^{\hat{\theta}_T})\right) + (\epsilon_j^t)^{(b)}, \quad t \in \mathbb{T}, j \in \{1, \dots, n\}.$$

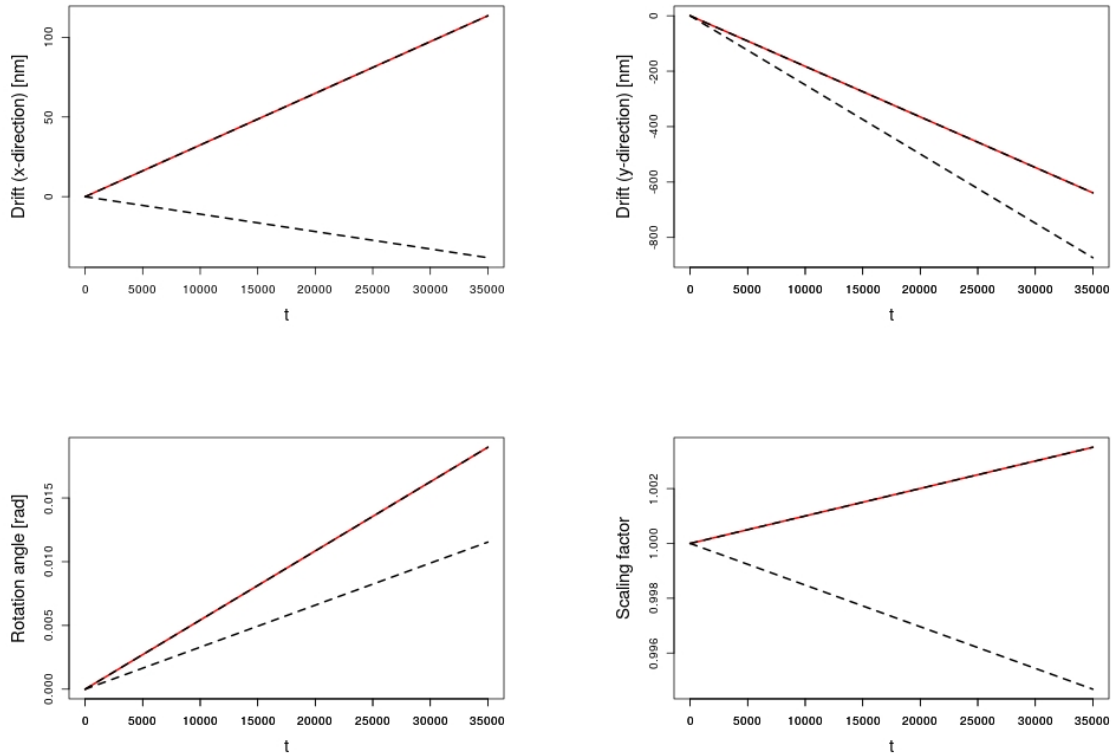


Figure 11.1: *Estimated motion functions (red) and confidence bands (dashed).*

Applying our estimation method to the $(O_j^t)^{(b)}$ we obtain bootstrap replicates $\hat{\theta}_T^{(b)}$, $\hat{\phi}_T^{(b)}$, and $\hat{\alpha}_T^{(b)}$ and thus replicates $\hat{f}_T^{(b)}$, $b \in \{1, \dots, B\}$. This in turn leads to bootstrap replicates

$$\begin{aligned}
 (r_j^t)^{(b)} &:= (O_j^t)^{(b)} - \hat{f}_T^{(b)} \left(1/\sigma_t^{\hat{\alpha}_T^{(b)}} \cdot R_{-\rho_t, \hat{\phi}_T^{(b)}}(x_j - \delta_t^{\hat{\theta}_T^{(b)}}) \right), \\
 \hat{\nu}^{(b)} &:= \sqrt{\frac{1}{nT} \sum_{j=1}^n \sum_{t \in \mathbb{T}} \left((r_j^t)^{(b)} - \frac{1}{nT} \sum_{j'=1}^n \sum_{t' \in \mathbb{T}} (r_{j'}^{t'})^{(b)} \right)^2}, \\
 \Delta_t^{(b)} &:= \left((\delta_t^{\hat{\theta}_T^{(b)}})_1 - (\delta_T^{\hat{\theta}_T})_1 \right) / \hat{\nu}^{(b)},
 \end{aligned}$$

which allow for minimization of $u_+ + u_-$ such that

$$\# \left\{ b \in \{1, \dots, B\} \mid \Delta_t^{(b)} \in [-u_+ g_+(t), u_- g_-(t)] \text{ for all } t \in [0, 1] \right\} \geq (1 - \eta)B.$$

Note that approximating the distribution of the drift parameter estimator $\hat{\theta}_T$ is not theoretically justified because we have not proved a central limit theorem for it. In the case of pure drift estimation, however, we were able to show asymptotic normality of $\hat{\theta}_T$ (Hartmann et al., 2016), and we are confident that it can also be derived in the scenario presented here with some additional work. Thus, we compute bootstrap confidence bands for the drift functions, too.

Because we assume the motion functions to be fixed at time $t = 0$ (at 0 for drift and rotation and at 1 for scaling), we can employ a confidence band which has width zero at $t = 0$. Since, in our application, we look at polynomial motion functions only, and since

on $[0, 1]$ the linear part dominates the others in the sense that $t \geq t^p$ for all $p > 1$ and $t \in [0, 1]$, we will choose $g_+(t) = g_-(t) = t$.

The thus obtained confidence bands for the above data set of Figure 10.4, linear drift, rotation, and scaling models, $B = 200$ and $\eta = 0.05$ are shown in Figure 11.1.

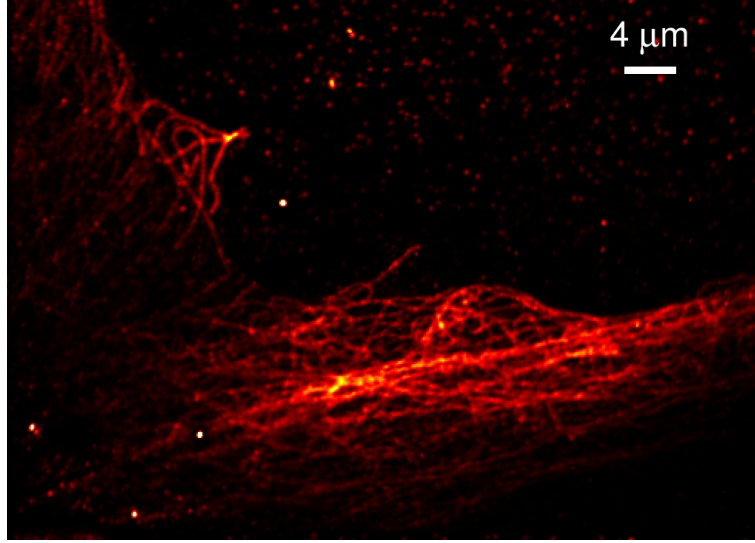


Figure 11.2: Average of the bootstrap replicates $\hat{f}_T^{(b_1)}, \dots, \hat{f}_T^{(b_m)}$ of the estimated image \hat{f}_T corresponding to the $m = \lceil (1 - \eta)B \rceil$ motion curves $t \mapsto ((\delta_t^{\hat{\theta}_T^{(b)}})_1, (\delta_t^{\hat{\theta}_T^{(b)}})_2, \rho_t^{\hat{\phi}_T^{(b)}}, \sigma_t^{\hat{\alpha}_T^{(b)}})$ nearest to the estimator $t \mapsto ((\delta_t^{\hat{\theta}_T})_1, (\delta_t^{\hat{\theta}_T})_2, \rho_t^{\hat{\phi}_T}, \sigma_t^{\hat{\alpha}_T})$ with respect to the supremum norm distance.

To further visualize the confidence statement we take a look at the average of (most of) the bootstrap replicates $\hat{f}_T^{(b)}$ of the estimator \hat{f}_T . For that we choose the 0.95-proportion of the bootstrap replicates

$$t \mapsto \left((\delta_t^{\hat{\theta}_T^{(b)}})_1, (\delta_t^{\hat{\theta}_T^{(b)}})_2, \rho_t^{\hat{\phi}_T^{(b)}}, \sigma_t^{\hat{\alpha}_T^{(b)}} \right), \quad b \in \{1, \dots, B\},$$

with the smallest supremum norm distances

$$\sup_{t \in [0, 1]} \left\| \left((\delta_t^{\hat{\theta}_T^{(b)}})_1, (\delta_t^{\hat{\theta}_T^{(b)}})_2, \rho_t^{\hat{\phi}_T^{(b)}}, \sigma_t^{\hat{\alpha}_T^{(b)}} \right) - \left((\delta_t^{\hat{\theta}_T})_1, (\delta_t^{\hat{\theta}_T})_2, \rho_t^{\hat{\phi}_T}, \sigma_t^{\hat{\alpha}_T} \right) \right\|$$

to the original motion function estimator. We denote the corresponding indices with b_1, \dots, b_m , where $m = \lceil (1 - \eta)B \rceil$. The convex hull of the corresponding motion curves resembles a four-dimensional bootstrap confidence band. Figure 11.2 shows the average of the images $\hat{f}_T^{(b_1)}, \dots, \hat{f}_T^{(b_m)}$ which thus “contains” the true image with a probability of about 0.95.

A visual comparison with 10.4 shows that this “bootstrap image” is blurrier than the estimator \hat{f}_T , but still better defined than the uncorrected image.

Chapter 12

Conclusion

We proposed a method for rigid motion (drift, rotation, and scaling) estimation and correction in sparse dynamic imaging and derived its asymptotic distributional properties. On the one hand, sparse acquisition is beneficial for improved spatial resolution and an important feature of any SMS microscopy. On the other hand, we have seen that this provides a significant burden as it is well known that the specimens move over time due to, for example, thermal inhomogeneity inside the sample and external systematic movements of the optical device. This raises a particular challenge for image registration as sparse acquisition and motion over time provide a conflicting situation. Currently, this is solved by technically incorporating bright fiducial markers into the specimen and registering their tracks. We claim that this can be completely discarded in many applications and it is sufficient to apply the proposed statistical method to estimate the drift, rotation, and scaling and finally to obtain the image from simply correcting the data by this estimated motion. The proposed method has been investigated in simulations and in real world examples from SMS microscopy and produces results similar to fiducial tracking. In general, reconstructions are quite satisfying. In particular, the results show a certain degree of stability with respect to parameter choices, for example, the thresholds r_T , u_T , and v_T . Consistency of the proposed estimators (and asymptotic normality for the rotation and scaling estimators) have been established which allows to quantify the statistical error of the motion estimates and the final image. To this end, simple bootstrap methods can be used.

It remains to further work to investigate higher order properties of the proposed estimator as well as properties of non parametric estimators, for example if the motion functions are estimated by splines. Note that the proposed method (at least the drift and scaling estimation) can in principle be applied to higher dimensions, in particular three dimensional measurements. However, computationally this appears to be much more demanding.

Appendix A

Mathematical Tools

The following are simple (but maybe not obvious) statements about complex numbers or functions.

Lemma A.1. For $a, b \in \mathbb{C}$ we have $|a \pm b|^2 = |a|^2 \pm 2\Re(a\bar{b}) + |b|^2$.

Proof. Write $a = a_1 + ia_2$ and $b = b_1 + ib_2$ with $a_1, a_2, b_1, b_2 \in \mathbb{R}$. Then,

$$|a \pm b|^2 = (a_1 \pm b_1)^2 + (a_2 \pm b_2)^2 = a_1^2 + a_2^2 + b_1^2 + b_2^2 \pm 2(a_1b_1 + a_2b_2) = |a|^2 + |b|^2 \pm 2\Re(a\bar{b}).$$

□

Corollary A.2. For $a, b \in \mathbb{C}$ we have $|a + b|^2 \leq 2|a|^2 + 2|b|^2$.

Proof. From Lemma A.1, we get

$$|a|^2 + |b|^2 - 2\Re(a\bar{b}) = |a - b|^2 \geq 0,$$

implying that $|a|^2 + |b|^2 \geq 2\Re(a\bar{b})$, and thus (again with Lemma A.1)

$$|a + b|^2 = |a|^2 + 2\Re(a\bar{b}) + |b|^2 \leq 2|a|^2 + 2|b|^2.$$

□

Corollary A.3. Let $g: [0, 1] \rightarrow \mathbb{C}$ integrable. Then,

$$\int_0^1 \left| g(t) - \int_0^1 g(t') dt' \right|^2 dt = \int_0^1 |g(t)|^2 dt - \left| \int_0^1 g(t) dt \right|^2.$$

Proof. Using Lemma A.1 and $a\bar{a} = |a|^2$ for all $a \in \mathbb{C}$, we have

$$\begin{aligned} & \int_0^1 \left| g(t) - \int_0^1 g(t') dt' \right|^2 dt \\ &= \int_0^1 \left(|g(t)|^2 dt + \left| \int_0^1 g(t') dt' \right|^2 - 2\Re \left(g(t) \overline{\int_0^1 g(t') dt'} \right) \right) dt \\ &= \int_0^1 |g(t)|^2 dt + \left| \int_0^1 g(t') dt' \right|^2 - 2\Re \left(\int_0^1 g(t) dt \overline{\int_0^1 g(t') dt'} \right) \\ &= \int_0^1 |g(t)|^2 dt - \left| \int_0^1 g(t') dt' \right|^2. \end{aligned}$$

□

Lemma A.4. *Let $a, b \in \mathbb{C}$ and $C > 0$ such that $|a|, |b| \leq C$. Then, the following inequality holds, $|a|^2 - |b|^2 \leq 2C|a - b|$.*

Proof. Because $|a| + |b| \leq 2C$, we have

$$|a|^2 - |b|^2 = (|a| + |b|)(|a| - |b|) \leq 2C(|a| - |b|) \leq 2C|a - b|,$$

where the last inequality comes from the second triangle inequality. \square

Lemma A.5. *Let $X \subseteq \mathbb{R}$ compact and $g_1, g_2: X \rightarrow \mathbb{C}$ bounded functions.*

1. *If g_1, g_2 are Lipschitz-continuous with Lipschitz-constants L_1 and L_2 , respectively, then $g_1 \cdot g_2$ is Lipschitz-continuous with Lipschitz-constant $L_1 \|g_2\|_\infty + L_2 \|g_1\|_\infty$.*
2. $\text{TV}(g_1 \cdot g_2) \leq \|g_2\|_\infty \text{TV}(g_1) + \|g_1\|_\infty \text{TV}(g_2)$.

Proof.

1. For all $x, y \in X$ we have

$$\begin{aligned} |g_1(x)g_2(x) - g_1(y)g_2(y)| &\leq |g_1(x)g_2(x) - g_1(y)g_2(x)| + |g_1(y)g_2(x) - g_1(y)g_2(y)| \\ &\leq (L_1 \|g_2\|_\infty + L_2 \|g_1\|_\infty) |x - y|. \end{aligned}$$

2. Let $\mathbb{X} = \{x_0, \dots, x_n\}$ a partition of X . Then,

$$\begin{aligned} &\sum_{i=1}^n |g_1(x_i)g_2(x_i) - g_1(x_{i-1})g_2(x_{i-1})| \\ &\leq \sum_{i=1}^n |g_1(x_i)g_2(x_i) - g_1(x_{i-1})g_2(x_i)| + |g_1(x_{i-1})g_2(x_i) - g_1(x_{i-1})g_2(x_{i-1})| \\ &\leq \|g_2\|_\infty \sum_{i=1}^n |g_1(x_i) - g_1(x_{i-1})| + \|g_1\|_\infty \sum_{i=1}^n |g_2(x_i) - g_2(x_{i-1})|. \end{aligned}$$

Taking the supremum over all possible partitions \mathbb{X} of X on both sides yields

$$\text{TV}(g_1 \cdot g_2) \leq \|g_2\|_\infty \text{TV}(g_1) + \|g_1\|_\infty \text{TV}(g_2).$$

\square

Lemma A.6. *Let $g: [0, 1] \rightarrow \mathbb{C}$, $T \in \mathbb{N}$, and $t_i := i/T$ for $i \in \{0, 1, \dots, T\}$. Then,*

$$\left| \frac{1}{T} \sum_{i=0}^{T-1} g(t_i) - \int_0^1 g(t) dt \right| \leq \frac{\text{TV}(g)}{T}.$$

Proof. We have

$$\begin{aligned} \left| \frac{1}{T} \sum_{i=0}^{T-1} g(t_i) - \int_0^1 g(t) dt \right| &= \left| \frac{1}{T} \sum_{i=0}^{T-1} g(t_i) - \sum_{i=0}^{T-1} \int_0^{1/T} g(t_i + t) dt \right| \\ &= \left| \int_0^{1/T} \sum_{i=0}^{T-1} (g(t_i) - g(t_i + t)) dt \right|. \end{aligned}$$

It follows that

$$\begin{aligned}
\left| \frac{1}{T} \sum_{i=0}^{T-1} g(t_i) - \int_0^1 g(t) dt \right| &\leq \int_0^{1/T} \sum_{i=0}^{T-1} |g(t_i) - g(t_i + t)| dt \\
&= \int_0^{1/T} \sum_{i=0}^{T-1} (|g(t_{i+1}) - g(t_i + t)| + |g(t_i + t) - g(t_i)|) dt \\
&\leq \int_0^{1/T} \sum_{j=0}^{2T-1} |g(\tilde{t}_{j+1}) - g(\tilde{t}_j)| dt \\
&\leq \int_0^{1/T} \text{TV}(g) dt = \frac{\text{TV}(g)}{T},
\end{aligned}$$

where

$$(\tilde{t}_0, \dots, \tilde{t}_{2T}) := \left(0, t, \frac{1}{T}, \frac{1}{T} + t, \frac{2}{T}, \frac{2}{T} + t, \dots, \frac{T-1}{T} + t, 1 \right), \quad t \in (0, 1/T),$$

defines a partition of $[0, 1]$. \square

Lemma A.7. Let $g_1, g_2: [0, 1] \rightarrow \mathbb{C}$ and $C > 0$ such that $|g_1(t)| \leq C$ and $|g_2(t)| \leq C$ for all $t \in [0, 1]$. Let $T \in \mathbb{N}$ and $t_i := i/T$ for $i \in \{0, 1, \dots, T\}$. Then,

$$\left| \left(\frac{1}{T} \sum_{i=0}^{T-1} g_1(t_i) \right) \left(\frac{1}{T} \sum_{i'=0}^{T-1} g_2(t_{i'}) \right) - \int_0^1 g_1(t) dt \int_0^1 g_2(t') dt' \right| \leq \frac{C}{T} (\text{TV}(g_1) + \text{TV}(g_2)).$$

Proof. With Lemma A.6, we have

$$\begin{aligned}
&\left| \left(\frac{1}{T} \sum_{i=0}^{T-1} g_1(t_i) \right) \left(\frac{1}{T} \sum_{i'=0}^{T-1} g_2(t_{i'}) \right) - \int_0^1 g_1(t) dt \int_0^1 g_2(t') dt' \right| \\
&\leq \left| \left(\frac{1}{T} \sum_{i=0}^{T-1} g_1(t_i) \right) \left(\frac{1}{T} \sum_{i'=0}^{T-1} g_2(t_{i'}) \right) - \left(\frac{1}{T} \sum_{i=0}^{T-1} g_1(t_i) \right) \int_0^1 g_2(t') dt' \right| \\
&\quad + \left| \left(\frac{1}{T} \sum_{i=0}^{T-1} g_1(t_i) \right) \int_0^1 g_2(t') dt' - \int_0^1 g_1(t) dt \int_0^1 g_2(t') dt' \right| \\
&= \left| \frac{1}{T} \sum_{i=0}^{T-1} g_1(t_i) \right| \left| \frac{1}{T} \sum_{i'=0}^{T-1} g_2(t_{i'}) - \int_0^1 g_2(t') dt' \right| \\
&\quad + \left| \frac{1}{T} \sum_{i=0}^{T-1} g_1(t_i) - \int_0^1 g_1(t) dt \right| \left| \int_0^1 g_2(t') dt' \right| \\
&\leq \frac{C}{T} (\text{TV}(g_1) + \text{TV}(g_2)).
\end{aligned}$$

\square

Lemma A.8. Let $a \in \mathbb{R}^2$. The function $g: \mathbb{R}^2 \rightarrow \mathbb{C}$, $x \mapsto e^{i\langle a, x \rangle}$, is Lipschitz-continuous with Lipschitz-constant $\sqrt{2} \|a\|$.

Proof. For $x, y \in \mathbb{R}^2$, we have

$$\begin{aligned}
|g(x) - g(y)|^2 &= |\cos(\langle a, x \rangle) - \cos(\langle a, y \rangle)|^2 + |\sin(\langle a, x \rangle) - \sin(\langle a, y \rangle)|^2 \\
&= 2 \langle a, x - y \rangle^2 \leq 2 \|a\|^2 \|x - y\|^2,
\end{aligned}$$

where we used the Lipschitz-continuity of \cos and \sin with Lipschitz-constant 1 and the Cauchy-Schwarz inequality. \square

Lemma A.9. *Let $g_1, g_2: [0, 1] \rightarrow \mathbb{R}$, $a, b \in \mathbb{C}$, and $g := ag_1 + bg_2$. Then,*

$$\text{TV}(g) \leq |a| \text{TV}(g_1) + |b| \text{TV}(g_2).$$

Proof. Let $\{t_0, \dots, t_n\}$ a partition of $[0, 1]$. We have

$$\begin{aligned} \sum_{i=0}^{n-1} |g(t_{i+1}) - g(t_i)| &= \sum_{i=0}^{n-1} |ag_1(t_{i+1}) + bg_2(t_{i+1}) - ag_1(t_i) - bg_2(t_i)| \\ &= \sum_{i=0}^{n-1} |a(g_1(t_{i+1}) - g_1(t_i)) + b(g_2(t_{i+1}) - g_2(t_i))| \\ &\leq |a| \sum_{i=0}^{n-1} |g_1(t_{i+1}) - g_1(t_i)| + |b| \sum_{i=0}^{n-1} |g_2(t_{i+1}) - g_2(t_i)|. \end{aligned}$$

Taking the supremum over all possible partitions yields the assertion. \square

Lemma A.10. *Let $g: \mathbb{R} \rightarrow \mathbb{R}$ differentiable, $a, b \in \mathbb{R}$ with $a < b$, and $C > 0$ such that $|g'(x)| \leq C$ for all $x \in [a, b]$. Then, the restriction $g|_{[a, b]}$ of g to $[a, b]$ is Lipschitz-continuous with Lipschitz-constant C .*

Proof. Let $x, y \in [a, b]$. By the mean value theorem, there is a $z \in [x, y]$ such that

$$|g(x) - g(y)| = |g(y) + (x - y)g'(z) - g(y)| = |g'(z)| |x - y| \leq C |x - y|.$$

\square

Lemma A.11. *Let $g_1: [0, 1] \rightarrow \mathbb{C}$ and $g_2: [0, 1] \rightarrow (0, \infty)$ integrable such that $|g_1| \leq 1$ and*

$$\left| \int_0^1 g_1(t) g_2(t) dt \right| = \int_0^1 g_2(t) dt. \quad (\text{A.1})$$

Then, there is $c \in \mathbb{C}$ with $|c| = 1$ such that $g_1(t) = c$ a.e. on $[0, 1]$.

Proof. From (A.1), we get

$$\int_0^1 g_2(t) dt = \left| \int_0^1 g_1(t) g_2(t) dt \right| \leq \int_0^1 |g_1(t)| g_2(t) dt$$

and therefore,

$$\int_0^1 (|g_1(t)| - 1) g_2(t) dt \geq 0.$$

Since $|g_1(t)| - 1 \leq 0$ and $g_2(t) > 0$ for all $t \in [0, 1]$, this implies that $|g_1(t)| - 1 = 0$ a.e., that is, $|g_1(t)| = 1$ a.e. In particular,

$$\overline{g_1(t)} = g_1(t)^{-1} \quad \text{a.e. on } [0, 1]. \quad (\text{A.2})$$

Because of

$$\begin{aligned} \int_0^1 \int_0^1 g_2(t) g_2(t') dt dt' &= \left(\int_0^1 g_2(t) dt \right)^2 = \left| \int_0^1 g_1(t) g_2(t) dt \right|^2 \\ &= \int_0^1 \int_0^1 g_1(t) \overline{g_1(t')} g_2(t) g_2(t') dt dt', \end{aligned}$$

we get

$$\begin{aligned}
0 &= \int_0^1 \int_0^1 \left(g_1(t) \overline{g_1(t')} - 1 \right) g_2(t) g_2(t') \, dt \, dt' \\
&= \int_0^1 \int_0^1 \left(\Re(g_1(t) \overline{g_1(t')}) - 1 \right) g_2(t) g_2(t') \, dt \, dt' \\
&\quad + i \int_0^1 \int_0^1 \Im(g_1(t) \overline{g_1(t')}) g_2(t) g_2(t') \, dt \, dt'.
\end{aligned} \tag{A.3}$$

This means that both the real and imaginary parts on the right hand side have to be zero. Because of

$$\Re(g_1(t) \overline{g_1(t')}) - 1 \leq \left| \Re(g_1(t) \overline{g_1(t')}) \right| - 1 \leq |g_1(t)| |g_1(t')| - 1 = 0,$$

and $g_2(t)g_2(t') > 0$, it follows from (A.3) that $\Re(g_1(t) \overline{g_1(t')}) = 1$ a.e. Since we also have that $\left| g_1(t) \overline{g_1(t')} \right| = |g_1(t)| |g_1(t')| = 1$ a.e., we get $g_1(t) \overline{g_1(t')} = 1$ a.e. Because of (A.2), this means that

$$g_1(t) = \left(\overline{g_1(t')} \right)^{-1} = \left(g_1(t')^{-1} \right)^{-1} = g_1(t') \quad \text{a.e. on } [0, 1],$$

which implies that $g_1(t) = c$ a.e. on $[0, 1]$ for some $c \in \mathbb{C}$ with $|c| = 1$. \square

Lemma A.12. *Let $g: \mathbb{R}^2 \rightarrow \mathbb{C}$ such that*

$$L_g := 2\pi\sqrt{2} \int_{\mathbb{R}^2} \|x\| |g(x)| \, dx < \infty.$$

Then, the Fourier transform $\mathcal{F}_g: \mathbb{R}^2 \rightarrow \mathbb{C}$ is Lipschitz-continuous with Lipschitz-constant L_g .

Proof. Let $\omega, \omega' \in \mathbb{R}^2$. Then,

$$\begin{aligned}
|\mathcal{F}_g(\omega) - \mathcal{F}_g(\omega')| &= \left| \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega, x \rangle} g(x) \, dx - \int_{\mathbb{R}^2} e^{-2\pi i \langle \omega', x \rangle} g(x) \, dx \right| \\
&= \left| \int_{\mathbb{R}^2} \left(e^{-2\pi i \langle \omega, x \rangle} - e^{-2\pi i \langle \omega', x \rangle} \right) g(x) \, dx \right| \\
&\leq \int_{\mathbb{R}^2} \left| e^{-2\pi i \langle \omega, x \rangle} - e^{-2\pi i \langle \omega', x \rangle} \right| |g(x)| \, dx \\
&\leq \int_{\mathbb{R}^2} 2\pi\sqrt{2} \|x\| \|\omega - \omega'\| |g(x)| \, dx = L_g \|\omega - \omega'\|,
\end{aligned}$$

where we used Lemma A.8 for the second inequality. \square

Lemma A.13. *Let $g: \mathbb{R}^2 \rightarrow \mathbb{C}$ such that*

$$L'_g := \sqrt{2} \int_0^\infty \int_0^{2\pi} \|(2\pi\psi, \log(r))\| r^\gamma |(g \circ \mathcal{P})(r, \psi)| \, d\psi \frac{dr}{r} < \infty.$$

Then, the analytical Fourier-Mellin transform $\mathcal{AFM}_g: \mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{C}$ is Lipschitz-continuous with Lipschitz-constant L'_g .

Proof. Let $(u, v), (u', v') \in \mathbb{Z} \times \mathbb{R}$. Then,

$$\begin{aligned}
& |\mathcal{AFM}_g(u, v) - \mathcal{AFM}_g(u', v')| \\
&= \left| \int_0^\infty \int_0^{2\pi} \left(e^{-2\pi i u \psi} r^{-iv} - e^{-2\pi i u' \psi} r^{-iv'} \right) r^\gamma (g \circ \mathcal{P})(r, \psi) \, d\psi \frac{dr}{r} \right| \\
&\leq \int_0^\infty \int_0^{2\pi} \left| e^{-2\pi i u \psi - iv \log(r)} - e^{-2\pi i u' \psi - iv' \log(r)} \right| r^\gamma |(g \circ \mathcal{P})(r, \psi)| \, d\psi \frac{dr}{r} \\
&\leq \int_0^\infty \int_0^{2\pi} \sqrt{2} \|(2\pi \psi, \log(r))\| \|(u, v) - (u', v')\| r^\gamma |(g \circ \mathcal{P})(r, \psi)| \, d\psi \frac{dr}{r} \\
&\leq L'_g \|(u, v) - (u', v')\|,
\end{aligned}$$

where we used Lemma A.8 for the second inequality. \square

Lemma A.14. *Let I a finite index set, $d \in \mathbb{N}$, and $g_i: \mathbb{R}^d \rightarrow \mathbb{C}$ for $i \in I$.*

1. *If g_i is differentiable for all $i \in I$, then, for all $m \in \{1, \dots, d\}$,*

$$\frac{\partial}{\partial x_m} \sum_{i, i' \in I} g_i(x) \overline{g_{i'}(x)} = 2 \sum_{i, i' \in I} \Re \left(\frac{\partial g_i}{\partial x_m}(x) \overline{g_{i'}(x)} \right).$$

2. *If g_i is twice differentiable for all $i \in I$, then, for all $k, m \in \{1, \dots, d\}$,*

$$\frac{\partial^2}{\partial x_k \partial x_m} \sum_{i, i' \in I} g_i(x) \overline{g_{i'}(x)} = 2 \sum_{i, i' \in I} \Re \left(\frac{\partial^2 g_i}{\partial x_k \partial x_m}(x) \overline{g_{i'}(x)} + \frac{\partial g_i}{\partial x_m}(x) \overline{\frac{\partial g_{i'}}{\partial x_k}(x)} \right).$$

Proof.

1. We have

$$\begin{aligned}
\frac{\partial}{\partial x_m} \sum_{i, i' \in I} g_i(x) \overline{g_{i'}(x)} &= \frac{\partial}{\partial x_m} \sum_{i, i' \in I} \left[\Re(g_i(x)) + i \Im(g_i(x)) \right] \left[\Re(g_{i'}(x)) - i \Im(g_{i'}(x)) \right] \\
&= \frac{\partial}{\partial x_m} \sum_{i, i' \in I} \left[\Re(g_i(x)) \Re(g_{i'}(x)) + \Im(g_i(x)) \Im(g_{i'}(x)) \right] \\
&= 2 \sum_{i, i' \in I} \left[\Re \left(\frac{\partial g_i}{\partial x_m}(x) \right) \Re(g_{i'}(x)) + \Im \left(\frac{\partial g_i}{\partial x_m}(x) \right) \Im(g_{i'}(x)) \right] \\
&= 2 \sum_{i, i' \in I} \Re \left(\frac{\partial g_i}{\partial x_m}(x) \overline{g_{i'}(x)} \right),
\end{aligned}$$

where we used the symmetry of the sum in i and i' for the second and third equality.

2. This follows directly from the first part of the Lemma and the product rule. \square

Lemma A.15. *Let $d \in \mathbb{N}$ and let $g: \mathbb{R}^d \times [0, 1] \rightarrow \mathbb{C}$ such that*

1. *$t \mapsto g(x, t)$ is integrable for all $x \in \mathbb{R}^d$,*

2. *$x \mapsto g(x, t)$ is continuously differentiable a.e. on $[0, 1]$,*

3. there is an integrable $h: [0, 1] \rightarrow [0, \infty)$ such that, for all $(x, t) \in \mathbb{R}^d \times [0, 1]$ and for all $m \in \{1, \dots, d\}$,

$$\left| \frac{\partial g}{\partial x_m}(x, t) \right| \leq h(t).$$

Then,

$$\frac{\partial}{\partial x_m} \int_0^1 \int_0^1 g(x, t) \overline{g(x, t')} dt dt' = 2 \int_0^1 \int_0^1 \Re \left(\frac{\partial g}{\partial x_m}(x, t) \overline{g(x, t')} \right) dt dt'.$$

Proof. First, note that if the partial derivatives of g are bounded by h , then so are the real and imaginary parts of those partial derivatives. We have

$$\begin{aligned} & \frac{\partial}{\partial x_m} \int_0^1 \int_0^1 g(x, t) \overline{g(x, t')} dt dt' \\ &= \frac{\partial}{\partial x_m} \int_0^1 \int_0^1 \left(\Re(g(x, t)) + i\Im(g(x, t)) \right) \left(\Re(g(x, t')) - i\Im(g(x, t')) \right) dt dt' \\ &= \frac{\partial}{\partial x_m} \int_0^1 \int_0^1 \Re(g(x, t))\Re(g(x, t')) + \Im(g(x, t))\Im(g(x, t')) dt dt' \\ &= \frac{\partial}{\partial x_m} \left[\int_0^1 \Re(g(x, t)) dt \right]^2 + \frac{\partial}{\partial x_m} \left[\int_0^1 \Im(g(x, t)) dt \right]^2 \\ &= 2 \int_0^1 \Re \left(\frac{\partial g}{\partial x_m}(x, t) \right) dt \int_0^1 \Re(g(x, t')) dt' \\ &\quad + 2 \int_0^1 \Im \left(\frac{\partial g}{\partial x_m}(x, t) \right) dt \int_0^1 \Im(g(x, t')) dt' \\ &= 2 \int_0^1 \int_0^1 \Re \left(\frac{\partial g}{\partial x_m}(x, t) \overline{g(x, t')} \right) dt dt', \end{aligned}$$

where we used the symmetry of the integrals in t and t' for the second equality and the chain rule as well as Theorem B.2 for the fourth equality, while the last one stems from the fact that $\Re(a)\Re(b) + \Im(a)\Im(b) = \Re(a\bar{b})$ for all $a, b \in \mathbb{C}$. \square

The following two lemmas are useful for showing asymptotic normality in Chapter 6.

Lemma A.16. *Let $(\nu_T)_{T \in \mathbb{N}}$ a sequence in $(0, \infty)$ and $\nu \in (0, \infty)$ such that $\nu_T \xrightarrow{T \rightarrow \infty} \nu$. Furthermore, let $X_T \sim \mathcal{N}(0, \nu_T)$ and $X \sim \mathcal{N}(0, \nu)$. Then, $X_T \xrightarrow{\mathcal{D}} X$.*

Proof. The characteristic functions of X_T and X are

$$\phi_{X_T}(y) = \exp \left(-\frac{\nu_T^2 y^2}{2} \right), \quad \phi_X(y) = \exp \left(-\frac{\nu^2 y^2}{2} \right).$$

Since $\nu_T \xrightarrow{T \rightarrow \infty} \nu$ and the functions $\tilde{\nu} \mapsto \exp(-\tilde{\nu}^2 y^2/2)$ are continuous, the statement follows from Theorem B.4. \square

Lemma A.17. *For $d \in \mathbb{N}$ and a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, let $(X_T)_{T \in \mathbb{N}}$ a sequence of random matrices $X_T: \Omega \rightarrow \mathbb{R}^{d \times d}$ and $X \in \mathbb{R}^{d \times d}$ such that X is invertible and $X_T \xrightarrow{T \rightarrow \infty} X$ in probability. Furthermore, let $(Y_T)_{T \in \mathbb{N}}$ a sequence of random vectors in \mathbb{R}^d , such that $(X_T Y_T)_{T \in \mathbb{N}}$ is uniformly tight. Then, $(Y_T)_{T \in \mathbb{N}}$ is uniformly tight.*

Proof. We define

$$X'_T: \Omega \rightarrow \text{GL}_d(\mathbb{R}), \quad \omega \mapsto \begin{cases} X_T(\omega), & \text{if } \det(X_T(\omega)) \neq 0, \\ X, & \text{if } \det(X_T(\omega)) = 0. \end{cases}$$

Since $X_T \xrightarrow{T \rightarrow \infty} X$ in probability, we get that $X'_T \xrightarrow{T \rightarrow \infty} X$ in probability. Let

$$X''_T := (X'_T)^{-1}.$$

Because matrix inversion and the determinant are continuous maps, it follows from the continuous mapping theorem (Theorem B.8 in the Appendix) that

$$\|X''_T\|_1 \xrightarrow{T \rightarrow \infty} \|X^{-1}\|_1 \quad \text{and} \quad \det(X_T) \xrightarrow{T \rightarrow \infty} \det(X) \neq 0 \quad \text{in probability.}$$

This means that for all $\delta > 0$ and $\epsilon > 0$, there is a $T(\delta, \epsilon) \in \mathbb{N}$ such that

$$\mathbb{P}(\|\|X''_T\|_1 - \|X^{-1}\|_1\| \geq \delta) < \frac{\epsilon}{4} \quad \text{and} \quad \mathbb{P}(\det(X_T) = 0) < \frac{\epsilon}{4} \quad \text{for all } T > T(\delta, \epsilon).$$

Fix $\delta > 0$ and $\epsilon > 0$. Then, with the event

$$E_T(\delta) := \left\{ \omega \in \Omega \mid \|\|X''_T(\omega)\|_1 - \|X^{-1}\|_1\| \geq \delta \right\} \cup \left\{ \omega \in \Omega \mid \det(X_T(\omega)) = 0 \right\}$$

we have that $\mathbb{P}(E_T(\delta)) < \epsilon/2$ for all $T > T(\delta, \epsilon)$. Furthermore, $X_T(\omega)$ is invertible with $X_T(\omega)^{-1} = X''_T(\omega)$ and $\|X''_T(\omega)\|_1 \leq \|X^{-1}\|_1 + \delta$ for all $\omega \in \Omega \setminus E_T(\delta)$. Since $(X_T Y_T)_{T \in \mathbb{N}}$ is uniformly tight, we have that for all $\tilde{\epsilon} > 0$, there is an $m(\tilde{\epsilon}) > 0$ such that

$$\sup_{T \in \mathbb{N}} \mathbb{P}(\|X_T Y_T\|_1 > m(\tilde{\epsilon})) \leq \frac{\tilde{\epsilon}}{2}.$$

Hence, for $\tilde{\epsilon} = \epsilon$ and for all $T > T(\delta, \epsilon)$, we have that

$$\begin{aligned} 1 - \frac{\epsilon}{2} &\leq \mathbb{P}(\|X_T Y_T\|_1 \leq m(\epsilon)) \\ &= \mathbb{P}(\|X_T Y_T\|_1 \leq m(\epsilon), E_T(\delta)) + \mathbb{P}(\|X_T Y_T\|_1 \leq m(\epsilon), \Omega \setminus E_T(\delta)) \\ &\leq \mathbb{P}(E_T(\delta)) + \mathbb{P}(\|X''_T\|_1 \|X_T Y_T\|_1 \leq \|X''_T\|_1 m(\epsilon), \Omega \setminus E_T(\delta)) \\ &\leq \frac{\epsilon}{2} + \mathbb{P}(\|X_T^{-1} X_T Y_T\|_1 \leq (\|X^{-1}\|_1 + \delta)m(\epsilon), \Omega \setminus E_T(\delta)) \\ &\leq \frac{\epsilon}{2} + \mathbb{P}(\|Y_T\|_1 \leq (\|X^{-1}\|_1 + \delta)m(\epsilon)), \end{aligned}$$

where we used Lemma 2.7 and the fact that $X''_T = X_T^{-1}$ on $\Omega \setminus E_T(\delta)$ for the third inequality. With $m^\delta(\epsilon) := (\|X^{-1}\|_1 + \delta)m(\epsilon)$, where we chose $\delta > 0$ arbitrarily, it follows that

$$\sup_{T > T(\delta, \epsilon)} \mathbb{P}(\|Y_T\|_1 > m^\delta(\epsilon)) \leq \epsilon.$$

Since this holds for all $\epsilon > 0$ with $m^\delta(\epsilon)$ not depending on T , we get that $(Y_T)_{T > T(\delta, \epsilon)}$ is uniformly tight. Because a single random variable is always tight, there is an $m_T(\epsilon) > 0$ for each $T \in \{1, \dots, T(\delta, \epsilon)\}$ and all $\epsilon > 0$ such that $\mathbb{P}(\|Y_T\|_1 > m_T(\epsilon)) \leq \epsilon$. We conclude that for all $\epsilon > 0$, there is an

$$\tilde{m}^\delta(\epsilon) := \max\{m_1(\epsilon), \dots, m_{T(\delta, \epsilon)}(\epsilon), m^\delta(\epsilon)\} > 0,$$

such that

$$\sup_{T \in \mathbb{N}} \mathbb{P}(\|Y_T\|_1 > \tilde{m}^\delta(\epsilon)) \leq \epsilon,$$

which means that $(Y_T)_{T \in \mathbb{N}}$ is uniformly tight. \square

Appendix B

Well-known theorems

For the reader's convenience, we present some theorems from the literature which are applied in the proofs in this work.

The following theorem can be found in Heuser, H. (1995, page 101).

Theorem B.1 (Continuity of parameter integrals). *Let X a metric space, E a Banach space, and $(\Omega, \mathcal{A}, \mu)$ a measure space. Let $g: X \times \Omega \rightarrow E$ with*

- (i) $\omega \mapsto g(x, \omega)$ is μ -measurable for all $x \in X$,
- (ii) $x \mapsto g(x, \omega)$ is continuous μ -a.e. on Ω ,
- (iii) there is an $h \in \mathcal{L}^1(\Omega, \mu, E)$ such that $|g(x, \omega)| \leq h(\omega)$ for all $(x, \omega) \in X \times \Omega$.

Then, $G: X \rightarrow E$, $x \mapsto \int_{\Omega} g(x, \omega) \mu(d\omega)$ is well-defined and continuous.

The following theorem can be found in Heuser, H. (1995, page 101).

Theorem B.2 (Differentiability of parameter integrals). *Let $d \in \mathbb{N}$, $U \subseteq \mathbb{R}^d$ open, E a Banach space, and $(\Omega, \mathcal{A}, \mu)$ a measure space. Let $g: U \times \Omega \rightarrow E$ with*

- (i) $\omega \mapsto g(x, \omega)$ is in $\mathcal{L}^1(\Omega, \mu, E)$ for all $x \in U$,
- (ii) $x \mapsto g(x, \omega)$ is continuously differentiable μ -a.e. on Ω ,
- (iii) there is an $h \in \mathcal{L}^1(\Omega, \mu, E)$ such that for all $(x, \omega) \in U \times \Omega$, $i \in \{1, \dots, d\}$,

$$\left| \frac{\partial g}{\partial x_i}(x, \omega) \right| \leq h(\omega).$$

Then, $G: U \rightarrow E$, $x \mapsto \int_{\Omega} g(x, \omega) \mu(d\omega)$ is well-defined and continuously differentiable such that for all $(x, \omega) \in U \times \Omega$, $i \in \{1, \dots, d\}$,

$$\frac{\partial G}{\partial x_i}(x) = \int_{\Omega} \frac{\partial g}{\partial x_j}(x, \omega) \mu(d\omega).$$

The following theorem can be found in Forster, O. (2004, page 133).

Theorem B.3. *For $x, y \in \mathbb{R}$ we have $\cos(x \pm y) = \cos(x) \cos(y) \mp \sin(x) \sin(y)$.*

The following theorem can be found in van der Vaart (2000, Theorem 2.13).

Theorem B.4 (Continuity theorem of Lévy and Cramér). *Let $(X_T)_{T \in \mathbb{N}}$ a sequence of random variables in \mathbb{R} , X a random variable in \mathbb{R} , and $\phi_{X_T}(y) := \mathbb{E}(e^{iyX_T})$ and $\phi_X(y) := \mathbb{E}(e^{iyX})$ their respective characteristic functions. Then, $X_T \xrightarrow{\mathcal{D}} X$ if and only if $\phi_{X_T}(y) \xrightarrow{T \rightarrow \infty} \phi_X(y)$ for all $y \in \mathbb{R}$.*

The following theorem can be found in van der Vaart (2000, page 16).

Corollary B.5 (Cramér-Wold Device). *Let $d \in \mathbb{N}$, $(X_T)_{T \in \mathbb{N}}$ a sequence of random vectors in \mathbb{R}^d , and X a random vector in \mathbb{R}^d . Then, $X_T \xrightarrow{\mathcal{D}} X$ if and only if $\langle \xi, X_T \rangle \xrightarrow{\mathcal{D}} \langle \xi, X \rangle$ for all $\xi \in \mathbb{R}^d$.*

The following theorem can be found in van der Vaart (2000, Theorem 5.7).

Theorem B.6 (van der Vaart). *Let $\Theta \subseteq \mathbb{R}^d$, $\theta_0 \in \Theta$, and $M: \Theta \rightarrow \mathbb{R}$ a function. Furthermore, let $(Y_T)_{T \in \mathbb{N}}$ a sequence of random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in a measure space (Ω', \mathcal{A}') and let $m: \Theta \times \Omega' \rightarrow \mathbb{R}$ a function such that $y \mapsto m(\theta, y)$ is measurable for all $\theta \in \Theta$ and $\theta \mapsto m(\theta, y)$ is continuous for all $y \in \Omega'$. For all $T \in \mathbb{N}$, define*

$$M_T: \Theta \rightarrow \mathbb{R}, \quad \theta \mapsto m(\theta, Y_T).$$

Let $(\hat{\theta}_T)_{T \in \mathbb{N}}$ a sequence of estimators for θ . Assume that

$$\sup_{\theta \in \Theta} |M_T(\theta) - M(\theta)| \xrightarrow{T \rightarrow \infty} 0 \quad \text{in probability,} \quad (\text{B.1})$$

$$\inf \{M(\theta) \mid \theta \in \Theta, \|\theta - \theta_0\| \geq \epsilon\} > M(\theta_0) \quad \text{for all } \epsilon > 0, \quad (\text{B.2})$$

$$\limsup_{T \rightarrow \infty} (M_T(\hat{\theta}_T) - M_T(\theta_0)) \leq 0. \quad (\text{B.3})$$

Then, $\hat{\theta}_T \xrightarrow{T \rightarrow \infty} \theta_0$ in probability.

The following theorem can be found in van der Vaart (2000, Theorem 3.8).

Theorem B.7 (Delta method). *Let $(\hat{\mu}_T)_{T \in \mathbb{N}}$ a sequence of random vectors in \mathbb{R}^d , $\mu_0 \in \mathbb{R}^d$, and $\Sigma \in \mathbb{R}^{d \times d}$, such that*

$$\sqrt{T}(\hat{\mu}_T - \mu_0) \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, \Sigma) \quad \text{in distribution.}$$

Let $k \in \mathbb{N}$ and let $g: \mathbb{R}^d \rightarrow \mathbb{R}^k$ continuously differentiable. Then

$$\sqrt{T}(g(\hat{\mu}_T) - g(\mu_0)) \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, J_g(\mu_0)\Sigma J_g(\mu_0)^\top) \quad \text{in distribution,}$$

where $J_g(\mu_0) \in \mathbb{R}^{k \times d}$ is the Jacobi matrix of g at μ_0 .

The following theorem can be found in van der Vaart (2000, Theorem 2.3).

Theorem B.8 (Continuous mapping). *Let $d, k \in \mathbb{N}$ and let $g: \mathbb{R}^d \rightarrow \mathbb{R}^k$ continuous a.e. on \mathbb{R}^d . Furthermore, let $(X_T)_{T \in \mathbb{N}}$ a sequence of random vectors in \mathbb{R}^d and X a random vector in \mathbb{R}^d .*

1. *If $X_T \xrightarrow{T \rightarrow \infty} X$ in distribution, then $g(X_T) \xrightarrow{T \rightarrow \infty} g(X)$ in distribution.*
2. *If $X_T \xrightarrow{T \rightarrow \infty} X$ in probability, then $g(X_T) \xrightarrow{T \rightarrow \infty} g(X)$ in probability.*
3. *If $X_T \xrightarrow{T \rightarrow \infty} X$ almost surely, then $g(X_T) \xrightarrow{T \rightarrow \infty} g(X)$ almost surely.*

The following theorem can be found in van der Vaart (2000, Theorem 2.8).

Theorem B.9 (Slutzky's Lemma). *For $d \in \mathbb{N}$, let $(X_T)_{T \in \mathbb{N}}$ and $(Y_T)_{T \in \mathbb{N}}$ sequences of random vectors in \mathbb{R}^d , $(Z_T)_{T \in \mathbb{N}}$ a sequence of random variables in \mathbb{R} , X a random vector in \mathbb{R}^d , $c \in \mathbb{R}^d$, and $c' \in \mathbb{R}$. Assume that $X_T \xrightarrow{T \rightarrow \infty} X$ in distribution, $Y_T \xrightarrow{T \rightarrow \infty} c$ in probability, and $Z_T \xrightarrow{T \rightarrow \infty} c'$ in probability. Then,*

1. $X_T + Y_T \xrightarrow{T \rightarrow \infty} X + c$ in distribution,
2. $Z_T X_T \xrightarrow{T \rightarrow \infty} c' X$ in distribution,
3. if $Z_T \neq 0$ almost surely and $c' \neq 0$, $Z_T^{-1} X_T \xrightarrow{T \rightarrow \infty} (c')^{-1} X$ in distribution.

The following theorem can be found in van der Vaart (2000, Example 2.6).

Theorem B.10 (Markov's inequality). *Let $d \in \mathbb{N}$, $m, p > 0$, and let $(X_T)_{T \in \mathbb{N}}$ a sequence of random vectors in \mathbb{R}^d . Then,*

$$\mathbb{P}(\|X_T\| > m) \leq \frac{\mathbb{E}(\|X_T\|^p)}{m^p}.$$

The following theorem can be found in van der Vaart (2000, Theorem 2.4).

Theorem B.11 (Prohorov). *Let $d \in \mathbb{N}$, $(X_T)_{T \in \mathbb{N}}$ a sequence of random vectors in \mathbb{R}^d .*

1. *If $X_T \xrightarrow{T \rightarrow \infty} X$ in distribution for some random vector X in \mathbb{R}^d , then $(X_T)_{T \in \mathbb{N}}$ is uniformly tight.*
2. *If $(X_T)_{T \in \mathbb{N}}$ is uniformly tight, then there exists a subsequence $(T_k)_{k \in \mathbb{N}}$ such that $X_{T_k} \xrightarrow{k \rightarrow \infty} X$ in distribution for some random vector X in \mathbb{R}^d .*

Bibliography

- Allasonnière, S., Amit, Y., and Trouvé, A. (2007). Towards a coherent statistical framework for dense deformable template estimation. *J. R. Statist. Soc. B*, 69(1):3–29.
- Antoniadis, A. and Bigot, J. (2006). Poisson inverse problems. *Ann. Statist.*, 34(5):2132–2158.
- Aspelmeier, T., Egner, A., and Munk, A. (2015). Modern statistical challenges in high-resolution fluorescence microscopy. *Annu. Rev. Statist. Appl.*, 2:163–202.
- Babcock, H., Sigal, Y. M., and Zhuang, X. (2012). A high-density 3D localization algorithm for stochastic optical reconstruction microscopy. *Opt. Nanoscopy*, 1(6).
- Berning, S., Willig, K. I., Steffens, H., Dibaj, P., and Hell, S. W. (2012). Nanoscopy in a living mouse brain. *Science*, 335(6068):551.
- Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645.
- Bigot, J., Gadat, S., Klein, T., and Marteau, C. (2013). Intensity estimation of non-homogeneous Poisson processes from shifted trajectories. *Electron. J. Statist.*, 7(2013):881–931.
- Bigot, J., Gamboa, F., and Vimond, M. (2009). Estimation of translation, rotation, and scaling between noisy images using the Fourier-Mellin transform. *SIAM J. Imaging Sci.*, 2(2):614–645.
- Bissantz, N., Claeskens, G., Holzmann, H., and Munk, A. (2009). Testing for lack of fit in inverse regression with applications to biophotonic imaging. *J. R. Statist. Soc. B*, 71(1):25–48.
- Bruhn, A., Weickert, J., and Schnörr, C. (2005). Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. *Int. J. Comput. Vis.*, 61(3):211–231.
- Cavalier, L. and Koo, J. Y. (2002). Poisson intensity estimation for tomographic data using a wavelet shrinkage approach. *IEEE Trans. Inf. Theory*, 48(10):2794–2802.
- Chen, X., Yang, J., Wu, Q., and Zhao, J. (2010). Motion blur detection based on lowest directional high-frequency energy. In *Proc. 2010 IEEE 17th Int. Conf. Image Process.*, pages 2533–2536.
- Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex fourier series. *Math. Comput.*, 19:297–301.

- Cox, S., Rosten, E., Monypenny, J., Jovanovic-Talisman, T., Burnette, D. T., Lippincott-Schwarz, J., Jones, G. E., and Heintzmann, R. (2012). Bayesian localization microscopy reveals nanoscale podosome dynamics. *Nat. Methods*, 9(2):195–200.
- Cuzol, A., Hellier, P., and Memin, E. (2007). A low dimensional fluid motion estimator. *Int. J. Comput. Vis.*, 75(3):329–349.
- Derrode, S. and Ghorbel, F. (2004). Shape analysis and symmetry detection in gray-level objects using the analytical fourier-mellin representation. *Signal Process.*, 84:25–39.
- Deschout, H., Znacchi, F. C., Mlodzianoski, M., Diaspro, A., Bewersdorf, J., Hess, S. T., and Braeckmans, K. (2014). Precisely and accurately localizing single emitters in fluorescence microscopy. *Nat. Methods*, 11(3):253–266.
- Egner, A., Geisler, C., von Middendorff, C., Bock, H., Wenzel, D., Medda, R., Andresen, M., Stiel, A., Jakobs, S., Eggeling, C., Schoenle, A., and Hell, S. W. (2007). Fluorescence nanoscopy in whole cells by asynchronous localization of photoswitching emitters. *Biophys. J.*, 93(9):3285–3290.
- Evans, L. C. (1998). *Partial Differential Equations (Graduate Studies in Mathematics vol 19)(Providence, RI: Am. Math. Soc.)*. Oxford University Press.
- Fleet, D. J. and Weiss, Y. (2006). *Optical Flow Estimation*, pages 237–257. Springer.
- Foroosh, H., Zerubia, J., and Berthod, M. (2002). Extension of phase correlation to subpixel registration. *IEEE Trans. Image Process.*, 11(3):188–200.
- Forster, O. (2004). *Analysis 1, Differential- und Integralrechnung einer Veränderlichen*. Vieweg, 7th edition.
- Frick, K., Marnitz, M., and Munk, A. (2013). Statistical multiresolution estimation for variational imaging: With an application in Poisson-biophotonics. *J. Math. Imaging Vis.*, 46(3):370–387.
- Gamboa, F., Loubes, J.-M., and Maza, E. (2007). Semi-parametric estimation of shifts. *Electron. J. Statist.*, 1:616–640.
- Gauthier, J. P., Bornard, G., and Silbermann, M. (1991). Motion and pattern analysis: Harmonic analysis on motion groups and their homogeneous spaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21:159–171.
- Geisler, C., Hotz, T., Schoenle, A., Hell, S. W., Munk, A., and Egner, A. (2012). Drift estimation for single marker switching based imaging schemes. *Opt. Express*, 20(7):7274–7289.
- Geisler, C., Schoenle, A., von Middendorff, C., Bock, H., Eggeling, C., Egner, A., and Hell, S. W. (2007). Resolution of $\lambda/10$ in fluorescence microscopy using fast single molecule photo-switching. *Appl. Phys. A*, 88(2):223–226.
- Ghorbel, F. (1994). A complete invariant description for gray-level images by the harmonic analysis approach. *Pattern Recogn. Lett.*, 15(10):1043–1051.
- Gonzalez, R.C. and Woods, R.E. (2002). *Digital Image Processing*. Prentice Hall, 2nd edition.

- Gower, J. C. (1975). Generalized Procrustes analysis. *Psychometrika*, 40(1):33–51.
- Gustafsson, M. G. L. (2005). Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Natl. Acad. Sci. U.S.A.*, 102(37):13081–13086.
- Hafi, N., Grunwald, M., van den Heuvel, L. S., Aspelmeier, T., Chen, J.-H., Zagrebelsky, M., Schütte, O. M., Steinem, C., Korte, M., Munk, A., and Walla, P. J. (2014). Fluorescence nanoscopy by polarization modulation and polarization angle narrowing. *Nat. Methods*, 11(5):579–584.
- Hall, P. and Pittelkow, Y. E. (1990). Simultaneous bootstrap confidence bands in regression. *J. Statist. Comput. Simul.*, 37(1–2):99–113.
- Hartmann, A., Huckemann, S., Dannemann, J., Laitenberger, O., Geisler, C., Egner, A., and Munk, A. (2016). Drift estimation in sparse sequential dynamic imaging, with application to nanoscale fluorescence microscopy. *J. R. Statist. Soc. B*, 78(3):563–587.
- Heintzmann, R., Jovin, T. M., and Cremer, C. (2002). Saturated patterned excitation microscopy—a concept for optical resolution improvement. *J. Opt. Soc. Am. A, Opt. Image Sci. Vis.*, 19(8):1599–1609.
- Hell, S. W. (2003). Toward fluorescence nanoscopy. *Nat. Biotechnol.*, 21(11):1347–1355.
- Hell, S. W. (2007). Far-field optical nanoscopy. *Science*, 316(5828):1153–1158.
- Hell, S. W. (2009). Microscopy and its focal switch. *Nat. Methods*, 6(1):24–32.
- Hell, S. W. and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.*, 19(11):780–782.
- Hess, S. T., Girirajan, T. P. K., and Mason, M. D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.*, 91(11):4258–4272.
- Heuser, H. (1995). *Analysis 2*. Teubner, 9th edition.
- Hofmann, M., Eggeling, C., Jakobs, S., and Hell, S. W. (2005). Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photo-switchable proteins. *Proc. Natl. Acad. Sci. U.S.A.*, 102(49):17565–17569.
- Holden, S. J., Uphoff, S., and Kapanides, A. N. (2011). Daostorm: An algorithm for high-density super-resolution microscopy. *Nat. Methods*, 8(4):279–280.
- Huang, F., Hartwich, T. M. P., Rivera-Molina, F. E., Lin, Y., Duim, W. C., Long, J. J., Uchil, P. D., Myers, J. R., Baird, M. A., Mothes, W., Davidson, M. W., Toomre, D., and Bewersdorf, J. (2013). Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nat. Methods*, 10(7):653–658.
- Huang, J.-Z., Tan, T.-N., Ma, L., and Wang, Y.-H. (2005). Phase correlation based iris image registration model. *J. Comput. Sci. Technol.*, 20(3):419–425.
- Jones, S. A., Shim, S.-H., He, J., and Zhuang, X. (2011). Fast, three-dimensional super-resolution imaging of live cells. *Nat. Methods*, 8(6):499–508.

- Klar, T. A., Jakobs, S., Dyba, M., Egner, A., and Hell, S. W. (2000). Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc. Natl. Acad. Sci. U.S.A.*, 97(15):8206–8210.
- Lenz, R. (1990). *Group Theoretical Methods in Image Processing*. Lecture Notes in Comput. Sci. 413. Springer-Verlag, Berlin.
- Li, H., Haltmeier, M., Zhang, S., Frahm, J., and Munk, A. (2014). Aggregated motion estimation for real-time MRI reconstruction. *Magn. Reson. Med.*, 72:1039–1048.
- Liu, R. Y. (1988). Bootstrap procedures under some non-i.i.d. models. *Ann. Statist.*, 16(4):1696–1708.
- Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Ann. Statist.*, 21(1):255–285.
- Nowak, R. D. and Kolaczyk, E. D. (2000). A statistical multiscale framework for Poisson inverse problems. *IEEE Trans. Inf. Theory*, 46(5):1811–1825.
- Papenberg, N., Bruhn, A., Brox, T., Didas, S., and Weickert, J. (2006). Highly accurate optic flow computation with theoretically justified warping. *Int. J. Comput. Vis.*, 67(2):141–158.
- Quan, T., Zhu, H., Liu, X., Liu, Y., Ding, J., Zeng, S., and Huang, Z.-L. (2011). High-density localization of active molecules using structured sparse model and Bayesian information criterion. *Opt. Express*, 19(18):16963–16974.
- Reddy, B. and Chatterji, B. (1996). An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans. Image Process.*, 5(8):1266–1271.
- Rudin, W. (1990). *Fourier Analysis on Groups*. John Wiley & Sons, New York. (Reprint of the 1962 original).
- Rust, M. J., Bates, M., and Zhuang, X. W. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods*, 3(10):793–795.
- Schmidt, R., Wurm, C. A., Jakobs, S., Engelhardt, J., Egner, A., and Hell, S. W. (2008). Spherical nanosized focal spot unravels the interior of cells. *Nat. Methods*, 5(6):539–544.
- Segman, J. S., Rubinstein, J., and Zeevi, Y. Y. (1992). The canonical coordinates method for pattern deformation: Theoretical and computational considerations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14:519–523.
- Silverman, B. W., Jones, M. C., Wilson, J. D., and Nychka, D. W. (1990). A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography. *J. R. Statist. Soc. B*, 52(2):271–324.
- Thompson, R., Larson, D., and Webb, W. (2002). Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.*, 82(5):2775–2783.
- van der Vaart, A. W. (2000). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.
- Vardi, Y., Shepp, L. A., and Kaufman, L. (1985). A statistical model for positron emission tomography. *J. Am. Statist. Assoc.*, 80(389):8–20.

- Weickert, J. and Schnörr, C. (2001). A theoretical framework for convex regularizers in PDE-based computation of image motion. *Int. J. Comput. Vis.*, 45(3):245–264.
- Westphal, V., Rizzoli, S. O., Lauterbach, M. A., Kamin, D., Jahn, R., and Hell, S. W. (2008). Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. *Science*, 320(5873):246–249.
- Wu, C. F. J. (1986). Jackknife, bootstrap, and other resampling methods in regression analysis (with discussion). *Ann. Statist.*, 14(4):1261–1295.
- Xu, W., Mulligan, J., Xu, D., and Chen, X. (2013). Detecting and classifying blurred image regions. In *Proc. 2013 IEEE Int. Conf. Multimed. Expo*, pages 1–6.
- Zhang, B., Fadili, J. M., and Starck, J. L. (2008). Wavelets, ridgelets, and curvelets for Poisson noise removal. *IEEE Trans. Image Process.*, 17(7):1093–1108.
- Zhu, L., Zhang, W., Elnatan, D., and Huang, B. (2012). Faster STORM using compressed sensing. *Nat. Methods*, 9(7):721–726.

