Drift estimation in sparse sequential dynamic imaging: with application to nanoscale fluorescence microscopy

Supplement: Proofs and Additional Material

Alexander Hartmann,
Institute for Mathematical Stochastics, Georg-August-University Göttingen, Germany

Stephan Huckemann,
Institute for Mathematical Stochastics, Georg-August-University Göttingen, Germany, Felix Bernstein Institute for Mathematical Statistics in the Biosciences, Georg-August-University Göttingen, Germany

Jörn Dannemann,
Institute for Mathematical Stochastics, Georg-August-University Göttingen, Germany

Oskar Laitenberger,
Laser Laboratory, Göttingen, Germany

Claudia Geisler,
Laser Laboratory, Göttingen, Germany

Alexander Egner,
Laser Laboratory, Göttingen, Germany

and Axel Munk[1]
Institute for Mathematical Stochastics, Georg-August-University Göttingen, Germany, Felix Bernstein Institute for Mathematical Statistics in the Biosciences, Georg-August-University Göttingen, Germany,
Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany

Key words and phrases: drift estimation, image registration, semiparametrics, M-estimation, nanoscale fluorescence microscopy, super resolution microscopy, asymptotic normality, sparsity, registration

AMS 2000 Subject Classification: Primary 62M10, 62M40, Secondary 92C05, 92C37

7 Appendix

Recall the notations defined in Subsection 2.2.

[1]*Address for correspondence*: Axel Munk, Institute for Mathematical Stochastics, Georg-August-University Göttingen, Goldschmidtstraße 7, 37077 Göttingen, Germany. E-mail: munk@math.uni-goettingen.de

7.1 Proof of Theorem 2.9

Plan of Proof. We start with a proof of (9), which follows a standard three step argument in M-estimation (e.g. (van der Vaart, 2000) and (Gamboa et al., 2007)), although the details are quite elaborate. First we show the uniqueness of the population contrast minimizer ϑ_0. In a second step we establish the continuity of $\vartheta \to \widetilde{M}(\vartheta)$. Thirdly, we verify that $\widetilde{M}_T(\vartheta) \to \widetilde{M}(\vartheta)$ a.s. uniformly over $\vartheta \in \Theta$ as $T, \xi_T \to \infty$, $\xi_T = o(\sqrt{T})$. In consequence, (van der Vaart, 2000, Theorem 5.7) (yielding weak consistency) can be adapted to obtain strong consistency. For convenience, here is the corresponding argument:

Since $\hat{\vartheta}_T$ is defined as a minimizer of $\widetilde{M}_T$ (hence $\widetilde{M}_T(\hat{\vartheta}_T) \leq \widetilde{M}_T(\vartheta_0)$) and $\widetilde{M}_T(\vartheta_0) \to \widetilde{M}(\vartheta_0)$ a.s., we have a.s. that

$$\limsup_{T\to\infty}\big(\widetilde{M}_T(\hat{\vartheta}_T) - \widetilde{M}(\vartheta_0)\big) = \limsup_{T\to\infty}\big(\widetilde{M}_T(\hat{\vartheta}_T) - \widetilde{M}_T(\vartheta_0)\big) + \lim_{T\to\infty}\big(\widetilde{M}_T(\vartheta_0) - \widetilde{M}(\vartheta_0)\big) \leq 0.$$

It follows that

$$\begin{aligned} \limsup_{T\to\infty} \widetilde{M}(\hat{\vartheta}_T) - \widetilde{M}(\vartheta_0) &\leq \limsup_{T\to\infty} \Big(\widetilde{M}(\hat{\vartheta}_T) - \widetilde{M}_T(\hat{\vartheta}_T)\Big) \\ &\leq \limsup_{T\to\infty} \sup_{\vartheta\in\Theta} \Big|\widetilde{M}(\vartheta) - \widetilde{M}_T(\vartheta)\Big| \;=\; 0 \text{ a.s.} \end{aligned} \tag{17}$$

Because of the uniqueness of the minimizer ϑ_0, the continuity of $\widetilde{M}$ and the compactness of Θ, we have that for every $\epsilon > 0$ there is $\eta_\epsilon > 0$ such that $\widetilde{M}(\vartheta) > \widetilde{M}(\vartheta_0) + \eta_\epsilon$ for all $\vartheta \in \Theta$ with $\|\vartheta - \vartheta_0\| \geq \epsilon$. Hence

$$\begin{aligned} P\Big(\limsup_{T\to\infty}\big\{\|\hat{\vartheta}_T - \vartheta_0\| \geq \epsilon\big\}\Big) &\leq P\Big(\limsup_{T\to\infty}\big\{\widetilde{M}(\hat{\vartheta}_T) > \widetilde{M}(\vartheta_0) + \eta_\epsilon\big\}\Big) \\ &\leq P\Big\{\limsup_{T\to\infty} \widetilde{M}(\hat{\vartheta}_T) \geq \widetilde{M}(\vartheta_0) + \eta_\epsilon\Big\} \;=\; 0\,, \end{aligned}$$

where the last equality follows from (17).

Step I: uniqueness of the contrast minimizer ϑ_0. First note that $\widetilde{M}(\vartheta) \geq -\sum_{k\in\mathbb{Z}^2} |f_k|^2$ for all ϑ with equality for $\vartheta = \vartheta_0$. If this minimum is attained for some ϑ then for each k with $|f_k|^2 > 0$

$$\left|\int_0^1 h_k(\delta_t^\vartheta - \delta_t^{\vartheta_0})\, dt\right|^2 = 1$$

since $|\int_0^1 h_k\, dt| \leq \int_0^1 |h_k|\, dt = 1$. This implies that $h_k(\delta_t^\vartheta - \delta_t^{\vartheta_0}) = 1$, i.e.

$$2\pi \left\langle k, \delta_t^\vartheta - \delta_t^{\vartheta_0}\right\rangle \equiv 0 \mod 2\pi$$

By Assumption 2.4 this holds for $k \in \{(k_1,k_2),(k_1',k_2')\}$ with $k_1k_2' - k_2k_1' \neq 0$. Hence, we can treat each dimension separately and obtain $\delta_t^\vartheta \equiv \delta_t^{\vartheta_0} \mod 2\pi$ a.e. Since this holds also for $k \in \{(k_1'', k_2''), (k_1''', k_2''')\}$ with $k_1''k_2''' - k_2''k_1''' \neq 0$, due to the part of the Assumption on non-common divisors we obtain $\delta_t^{\vartheta} = \delta_t^{\vartheta_0}$ a.e. and hence $\vartheta = \vartheta_0$.

Step II: continuity of $\widetilde{M}$. For $\vartheta, \vartheta' \in \Theta$ we have that

$$
\begin{aligned}
|\widetilde{M}(\vartheta) - \widetilde{M}(\vartheta')| &\leq \sum_{k \in \mathbb{Z}^2} |f_k|^2 \left| \left| \int_0^1 h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0})\, dt \right|^2 - \left| \int_0^1 h_k(\delta_t^{\vartheta'} - \delta_t^{\vartheta_0})\, dt \right|^2 \right| \\
&\leq 2 \sum_{k \in \mathbb{Z}^2} |f_k|^2 \left| \int_0^1 \left(e^{2\pi i \left\langle k, \delta_t^{\vartheta} - \delta_t^{\vartheta_0} \right\rangle} - e^{2\pi i \left\langle k, \delta_t^{\vartheta'} - \delta_t^{\vartheta_0} \right\rangle} \right) dt \right| \\
&\leq 2 \sum_{k \in \mathbb{Z}^2} |f_k|^2 \int_0^1 \left| 1 - e^{2\pi i \left\langle k, \delta_t^{\vartheta'} - \delta_t^{\vartheta} \right\rangle} \right| dt \\
&\leq 4\pi \sum_{k \in \mathbb{Z}^2} |k| |f_k|^2 \int_0^1 \left\| \delta_t^{\vartheta} - \delta_t^{\vartheta'} \right\| dt\,,
\end{aligned}
$$

where we use

$$
|a|^2 - |b|^2 \leq 2|a - b| \tag{18}
$$

for $a, b \in \mathbb{C}$ with $|a|, |b| < 1$ in the second inequality and $|1 - e^{ix}|^2 = 2 - 2\cos x \leq x^2$ in the fourth one. By Assumptions 2.4, 2.6, this implies the continuity of $\widetilde{M}(\vartheta)$.

Step III: $\widetilde{M}_T \to \widetilde{M}$ uniformly in ϑ a.s. Recall from model (4) that

$$
Y_k^t = h_k(-\delta_t^{\vartheta_0}) f_k + W_k^t
$$

with the true and unknown parameter $\vartheta_0 \in \Theta$. Hence with (7) we have that

$$
\widetilde{M}_T(\vartheta) = - \sum_{|k| < \xi_T} \left| \frac{1}{T} \sum_{t \in \mathbb{T}} \left(h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0}) f_k + h_k(\delta_t^{\vartheta}) W_k^t \right) \right|^2 = A_T(\vartheta) - B_T(\vartheta) - C_T(\vartheta)
$$

with

$$
\begin{aligned}
A_T(\vartheta) &:= - \sum_{|k| < \xi_T} \left| \frac{1}{T} \sum_{t \in \mathbb{T}} h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0}) f_k \right|^2, \\
B_T(\vartheta) &:= \sum_{|k| < \xi_T} 2\, \mathrm{Re} \left(\left(\frac{1}{T} \sum_{t \in \mathbb{T}} h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0}) f_k \right) \left(\frac{1}{T} \sum_{t' \in \mathbb{T}} h_k(-\delta_{t'}^{\vartheta}) \overline{W_k^{t'}} \right) \right), \\
C_T(\vartheta) &:= \sum_{|k| < \xi_T} \left| \frac{1}{T} \sum_{t \in \mathbb{T}} h_k(\delta_t^{\vartheta}) W_k^t \right|^2.
\end{aligned}
$$

To derive the desired uniform convergence we will show for the deterministic part that $A_T \to \widetilde{M}$ uniformly in ϑ while the random parts B_T and C_T converge to zero uniformly a.s. Considering

$$
\begin{aligned}
|A_T(\vartheta) - \widetilde{M}(\vartheta)| \quad &\leq \quad \sum_{|k|<\xi_T} |f_k|^2 \left| \left| \frac{1}{T} \sum_{t\in\mathbb{T}} h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0}) \right|^2 - \left| \int_0^1 h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0})\, dt \right|^2 \right| \\
&\quad + \sum_{|k|\geq\xi_T} |f_k|^2 \left| \int_0^1 h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0})\, dt \right|^2 ,
\end{aligned}
$$

and applying (18) again to the first sum while noting that the second is bounded by $\sum_{|k|\geq\xi_T} |f_k|^2 = o(1)$ ($\xi_T \to \infty$ by hypothesis and $\sum_k |f_k|^2 < \infty$ by Assumption 2.4) gives

$$
|A_T(\vartheta) - \widetilde{M}(\vartheta)| \quad \leq \quad \sum_{|k|<\xi_T} 2|f_k|^2 \left| \frac{1}{T} \sum_{t\in\mathbb{T}} h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0}) - \int_0^1 h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0})\, dt \right| + o(1).
$$

Since the total variation of $t \mapsto h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0})$ is bounded by a constant times $|k|$ uniformly in ϑ (Assumption 2.6), we have for some constant C that

$$
\left| \frac{1}{T} \sum_{t\in\mathbb{T}} h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0}) - \int_0^1 h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0})\, dt \right| < \frac{|k|\, C}{T}.
$$

In consequence of $\sum_k |k||f_k|^2 < \infty$ (Assumption 2.4) this implies that

$$
|A_T(\vartheta) - \widetilde{M}(\vartheta)| = O(1/T)\, ,
$$

uniformly in ϑ as desired. Next, we show

$$
\sup_{\vartheta\in\Theta} C_T(\vartheta) \quad = \quad \sup_{\vartheta\in\Theta} \sum_{|k|<\xi_T} \left| \frac{1}{T} \sum_{t\in\mathbb{T}} h_k(\delta_t^{\vartheta}) W_k^t \right|^2 = o\left(\frac{\xi_T^2}{T} \right) \text{ a.s.} \tag{19}
$$

Since $h_k(\delta_t^{\vartheta})$ acts as a rotation, $h_k(\delta_t^{\vartheta}) W_k^t =: U_k^t + iV_k^t$ ($t \in \mathbb{T}, |k| < \xi_T$) are again independently complex normally distributed; in particular, every $U_k^t = \mathrm{Re}(h_k(\delta_t^{\vartheta}) W_k^t)$ is independent of $V_k^t = \mathrm{Im}(h_k(\delta_t^{\vartheta}) W_k^t)$. Let

$$
\bar{U}_{k,T} = \frac{1}{\sqrt{T}} \sum_{t\in\mathbb{T}} U_k^t, \quad \bar{V}_{k,T} = \frac{1}{\sqrt{T}} \sum_{t\in\mathbb{T}} V_k^t.
$$

Because of $E(\epsilon_{j,t}^4) = 3$ and Assumption 2.7 we have

$$\begin{aligned}
\mathrm{Var}(\bar{U}_{k,T}^2) &\leq E(\bar{U}_{k,T}^4) \\
&= \frac{3}{T^2}\sum_{t\in\mathbb{T}}\frac{1}{n_t^2}\sum_{j\in J_t}\sigma_{j,t}^4\cos(-2\pi\langle k, x_{j,t}-\delta_t^{\vartheta}\rangle)^4 \\
&\quad +\frac{3}{T^2}\sum_{t\neq t'}\frac{1}{n_t n_{t'}}\sum_{j\in J_t}\sum_{j'\in J_{t'}}\sigma_{j,t}^2\sigma_{j',t'}^2\cos(-2\pi\langle k, x_{j,t}-\delta_t^{\vartheta}\rangle)^2\cos(-2\pi\langle k, x_{j',t'}-\delta_{t'}^{\vartheta}\rangle)^2 \\
&\leq 3\sigma_{\max}^4\left(\frac{1}{T^2}\sum_{t\in\mathbb{T}}\frac{1}{n_t}+1\right)\leq 6\sigma_{\max}^4,
\end{aligned}$$

and similarly $\mathrm{Var}(\bar{V}_{k,T}^2)\leq 6\sigma_{\max}^4$. Again by Assumption 2.7,

$$\begin{aligned}
E(\bar{U}_{k,T}^2+\bar{V}_{k,T}^2) &= \frac{1}{T}\sum_{t\in\mathbb{T}}\frac{1}{n_t}\sum_{j\in J_t}\sigma_{j,t}^2\left(\cos(-2\pi\langle k, x_{j,t}-\delta_t^{\vartheta}\rangle)^2+\sin(-2\pi\langle k, x_{j,t}-\delta_t^{\vartheta}\rangle)^2\right) \\
&= \frac{1}{T}\sum_{t\in\mathbb{T}}\frac{1}{n_t}\sum_{j\in J_t}\sigma_{j,t}^2\leq\sigma_{\max}^2.
\end{aligned}$$

In consequence, Kolmogorov's strong law (see e.g. (Sen and Singer, 1993, Theorem 2.3.10)) yields that

$$\begin{aligned}
&\left|\frac{1}{\#\{|k|<\xi_T\}}\sum_{|k|<\xi_T}\left|\frac{1}{\sqrt{T}}\sum_{t\in\mathbb{T}}h_k(\delta_t^{\vartheta})W_k^t\right|^2-\frac{1}{T}\sum_{t\in\mathbb{T}}\frac{1}{n_t}\sum_{j\in J_t}\sigma_{j,t}^2\right| \\
=&\left|\frac{1}{\#\{|k|<\xi_T\}}\sum_{|k|<\xi_T}(\bar{U}_{k,T}^2+\bar{V}_{k,T}^2)-\frac{1}{T}\sum_{t\in\mathbb{T}}\frac{1}{n_t}\sum_{j\in J_t}\sigma_{j,t}^2\right| \\
\to&\ 0\text{ a.s.},\quad T\to\infty.
\end{aligned}$$

Since $\#\{|k|<\xi_T\}=O(\xi_T^2)$ this yields (19). Finally,

$$\sup_{\vartheta}|B_T(\vartheta)|^2=o(1)\text{ a.s.}$$

follows at once from $|A_T(\vartheta)|\leq\sum_k|f_k|^2$ by definition, (19) and the observation that $|B_T(\vartheta)|^2\leq 2|A_T(\vartheta)|\,|C_T(\vartheta)|$. This concludes the proof of Step III.

The proof of (10). Observe that, using the Plancherel equality, we have

$$\begin{aligned}
\left\|\hat{f}_T-f\right\|_2^2 &= \sum_{|k|<\xi_T}\left|\frac{1}{T}\sum_{t\in\mathbb{T}}h_k(\delta_t^{\hat{\vartheta}_T})Y_k^t-f_k\right|^2+\sum_{|k|\geq\xi_T}|f_k|^2 \\
&= \sum_{|k|<\xi_T}\left|\frac{1}{T}\sum_{t\in\mathbb{T}}\left(h_k(\delta_t^{\hat{\vartheta}_T}-\delta_t^{\vartheta_0})f_k+h_k(\delta_t^{\hat{\vartheta}_T})W_k^t\right)-f_k\right|^2+o(1)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{|k|<\xi_T} |f_k|^2 \frac{1}{T^2} \sum_{t,t'\in\mathbb{T}} \left(h_k(\delta_t^{\hat\vartheta_T} - \delta_t^{\vartheta_0}) - 1\right)\left(h_k(-\delta_{t'}^{\hat\vartheta_T} + \delta_{t'}^{\vartheta_0}) - 1\right) \\
&\quad + \sum_{|k|<\xi_T} \left| \frac{1}{T} \sum_{t\in\mathbb{T}} h_k(\delta_t^{\hat\vartheta_T}) W_k^t \right|^2 \\
&\quad\quad + 2 \sum_{|k|<\xi_T} \frac{1}{T^2} \sum_{t,t'\in\mathbb{T}} \left(h_k(\delta_t^{\hat\vartheta_T} - \delta_t^{\vartheta_0}) - 1\right) f_k h_k(-\delta_{t'}^{\hat\vartheta_T}) \overline{W_k^{t'}} + o(1) \\
&\leq 4\pi L \|\hat\vartheta_T - \vartheta_0\| \sum_{|k|<\xi_T} \left(|f_k|^2\,|k| + |f_k|\,|k|\,\frac{1}{\sqrt{T}} |G_k^T| \right) + o(1) \text{ a.s.} \qquad (20)
\end{aligned}
$$

with G_k^T defined below, by (19), since $|h_k(\delta_t^{\hat\vartheta_T} - \delta_t^{\vartheta_0}) - 1| \leq 2$ as well as (recalling the argument following display (18))

$$
\left| h_k(\delta_t^{\hat\vartheta_T} - \delta_t^{\vartheta_0}) - 1 \right| \leq 2\pi |k| \|\delta_t^{\hat\vartheta_T} - \delta_t^{\vartheta_0}\| \leq 2\pi L |k| \|\hat\vartheta_T - \vartheta_0\|
$$

with the constant $L > 0$ from Assumption 2.8 and the following argument. Setting

$$
G_k^T := \frac{1}{\sqrt{T}} \sum_{t'\in\mathbb{T}} h_k(-\delta_{t'}^{\hat\vartheta_T}) \overline{W_k^{t'}}\,,
$$

we obtain complex normal deviates independent in k with the property

$$
\frac{1}{T^2} \sum_{t,t'\in\mathbb{T}} \left(h_k(\delta_t^{\hat\vartheta_T} - \delta_t^{\vartheta_0}) - 1\right) f_k h_k(-\delta_{t'}^{\hat\vartheta_T}) \overline{W_k^{t'}} \;=\; \frac{f_k}{\sqrt{T}} \left(\frac{1}{T} \sum_{t\in\mathbb{T}} \left(h_k(\delta_t^{\hat\vartheta_T} - \delta_t^{\vartheta_0}) - 1\right) \right) G_k^T .
$$

Now (20) yields indeed $\|\hat f_T - f\|_2^2 \to 0$ a.s. if $\xi_T^2/\sqrt{T} \to 0$ since $\|\hat\vartheta_T - \vartheta_0\| \to 0$ a.s. as shown in the proof of the first part of Theorem 2.9, $\sup_{k\in\mathbb{Z}} |f_k|\,|k| < \infty$ by Remark 2.5 and $\sum_{|k|<\xi_T} |f_k|^2\,|k| < \infty$ by Assumption 2.4. The same argument that led to (19) shows that the variance of

$$
\frac{1}{\sqrt{T}} \sum_{|k|<\xi_T} |f_k|\,|k|\,|G_k^T|
$$

is of order $o(1)$ in case of $\xi_T/\sqrt{T} \to 0$, which gives convergence of $\|\hat f_T - f\|_2 \to 0$ in probability, completing the proof. □

7.2 Proof of (i) of Theorem 2.13

With the d-dimensional real vector $\mathbf{a}_{k,t}^{\vartheta} := 2\pi \operatorname{grad}_\vartheta \langle k, \delta_t^{\vartheta} \rangle$ verify that

$$
\begin{aligned}
\operatorname{grad}_\vartheta \left(\sum_{t\in\mathbb{T}} h_k(\delta_t^{\vartheta}) Y_k^t \sum_{t'\in\mathbb{T}} \overline{h_k(\delta_{t'}^{\vartheta}) Y_k^{t'}} \right) &= 2\,\mathrm{Re} \left(\sum_{t,t'\in\mathbb{T}} \operatorname{grad}_\vartheta \left(h_k(\delta_t^{\vartheta}) Y_k^t \right) \overline{h_k(\delta_{t'}^{\vartheta}) Y_k^{t'}} \right) \\
&= -2\,\mathrm{Im} \left(\sum_{t,t'\in\mathbb{T}} \mathbf{a}_{k,t}^{\vartheta} h_k(\delta_t^{\vartheta}) Y_k^t\, \overline{h_k(\delta_{t'}^{\vartheta}) Y_k^{t'}} \right) . \qquad (21)
\end{aligned}
$$

Moreover, with the true parameter $\vartheta_0 \in \Theta$ and arbitray $\vartheta \in \Theta$ recall from (2) that

$$h_k(\delta_t^{\vartheta})Y_k^t = h_k(\delta_t^{\vartheta} - \delta_t^{\vartheta_0})f_k + h_k(\delta_t^{\vartheta})W_k^t.$$

At $\vartheta = \vartheta_0$ the right hand side is just $f_k + h_k(\delta_t^{\vartheta_0})W_k^t$. In consequence we have for $\widetilde{M}_T$ from (7) that

$$\mathrm{grad}_{\vartheta}\widetilde{M}_T(\vartheta_0) = \sum_{|k|\leq\xi_T} H_k^T \tag{22}$$

where $\mathbf{a}_k^t = \mathbf{a}_{k,t}^{\vartheta_0}$, $f_k = e_k + ig_k$, $h_k(\delta_t^{\vartheta_0})W_k^t = \tau_k^t A_k^t + i\omega_k^t B_k^t$ with standard deviations

$$\tau_k^t := \sqrt{\frac{1}{n_t}\sum_{j\in J_t}\sigma_{j,t}^2\cos(-2\pi\langle k, x_{j,t} - \delta_t^{\vartheta_0}\rangle)^2},$$
$$\omega_k^t := \sqrt{\frac{1}{n_t}\sum_{j\in J_t}\sigma_{j,t}^2\sin(-2\pi\langle k, x_{j,t} - \delta_t^{\vartheta_0}\rangle)^2},$$

and

$$\begin{aligned} H_k^T &:= \frac{2}{T^2}\,\mathrm{Im}\left(\sum_{t,t'\in\mathbb{T}}\mathbf{a}_k^t\Big(|f_k|^2 + f_k\overline{h_k(\delta_{t'}^{\vartheta_0})W_k^{t'}} + h_k(\delta_t^{\vartheta_0})W_k^t\overline{f_k} + h_k(\delta_t^{\vartheta_0})W_k^t\,\overline{h_k(\delta_{t'}^{\vartheta_0})W_k^{t'}}\Big)\right) \\ &= \frac{2}{T^2}\sum_{t,t'\in\mathbb{T}}\mathbf{a}_k^t\Big(g_k\tau_k^{t'}A_k^{t'} - e_k\omega_k^{t'}B_k^{t'} + e_k\omega_k^t B_k^t - g_k\tau_k^t A_k^t + \tau_k^{t'}\omega_k^t A_k^{t'}B_k^t - \tau_k^t\omega_k^{t'}A_k^t B_k^{t'}\Big). \end{aligned}$$

Note that $A_k^t, B_k^t \sim \mathcal{N}(0,1)$ ($k \in \mathbb{Z}^2$, $t \in \mathbb{T}$) are all mutually independent, and for $k = (0,0)$ we have $\omega_{(0,0)}^t \equiv 0$.

To determine the limit distribution of $\sqrt{T}\mathrm{grad}_{\vartheta}M_T(\vartheta)$ we look at its projections $\sqrt{T}\langle x, \mathrm{grad}_{\vartheta}M_T(\vartheta)\rangle$ with arbitrary but fixed $0 \neq x = (x_1,\ldots,x_d) \in \mathbb{R}^d$. To this end denote by $H_k^T(j)$ and $\mathbf{a}_k^t(j)$ the j-th component of H_k^T and $\mathbf{a}_k^t$, respectively, $j \in \{1,\ldots,d\}$, and set

$$G_k^T := \sum_{j=1}^d x_j H_k^T(j), \quad a_k^t := \sum_{j=1}^d x_j\mathbf{a}_k^t(j). \tag{23}$$

Introducing the independent normal vectors $A_k := (\tau_k^t A_k^t/\bar{\tau}_k^T)_{t\in\mathbb{T}}$, $B_k := (\omega_k^t B_k^t/\bar{\omega}_k^T)_{t\in\mathbb{T}}$ with (cf. Assumption 2.12)

$$\bar{\tau}_k^T = \sqrt{\frac{1}{T}\sum_{s\in\mathbb{T}}(\tau_k^s)^2} > 0, \quad \bar{\omega}_k^T = \sqrt{\frac{1}{T}\sum_{s\in\mathbb{T}}(\omega_k^s)^2} > 0,$$

each with independent components as well as the unit vector $e := (1)_{t\in\mathbb{T}}/\sqrt{T}$ and the vector $a_k = (a_k^t)_{t\in\mathbb{T}}$ and denoting the transpose of a_k by a_k' etc., we obtain

$$\begin{aligned} G_k^T &= \frac{2\bar{\tau}_k^T\bar{\omega}_k^T}{T^{3/2}}\Big(a_k'B_kA_k'e - e'B_kA_k'a_k\Big) \\ &\quad +\frac{2}{T}\Big(\bar{\tau}_k^T g_k a_k' ee'A_k - \bar{\omega}_k^T e_k a_k' ee' B_k + \bar{\omega}_k^T e_k a_k' B_k - \bar{\tau}_k^T g_k a_k' A_k\Big)\,. \end{aligned}$$

To tackle the first term introduce a unit vector b_k orthogonal to e such that $a_k = \alpha_k e + \beta_k b_k$, $\alpha_k, \beta_k \in \mathbb{R}$ and define a matrix $U = U_k \in SO(T)$ having e and b_k as the first two columns. Then, with the independent normal vectors $\widetilde{A}_k = U'A_k$, $\widetilde{B}_k = U'B_k$ with independent components, each with zero mean,

$$\begin{aligned} a_k'B_kA_k'e - e'B_kA_k'a_k &= A_k'(ea_k' - a_ke')B_k \\ &= A_k'UU'(ea_k' - a_ke')UU'B_k \\ &= A_k'U(e,b_k,*)'\Big(e(\alpha_k e + \beta_k b_k)' - (\alpha_k e + \beta_k b_k)e'\Big)(e,b_k,*)U'B_k \\ &= \widetilde{A}_k'\Big((1,0,\ldots,0)'(\alpha_k,\beta_k,0,\ldots,0) - (\alpha_k,\beta_k,0,\ldots,0)'(1,0,\ldots,0)\Big)\widetilde{B}_k \\ &= \widetilde{A}_k'\,\beta_k \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \widetilde{B}_k\,. \end{aligned}$$

In consequence, with the first components $\widetilde{A}_k^{(1)}$, $\widetilde{B}_k^{(1)}$ and second components $\widetilde{A}_k^{(2)}$, $\widetilde{B}_k^{(2)}$ of $\widetilde{A}_k$ and $\widetilde{B}_k$,

$$\begin{aligned} G_k^T &= \frac{2\bar{\tau}_k^T\bar{\omega}_k^T\beta_k}{T^{3/2}}\Big(\widetilde{A}_k^{(1)}\widetilde{B}_k^{(2)} - \widetilde{A}_k^{(2)}\widetilde{B}_k^{(1)}\Big) \\ &\quad +\frac{2}{T}\Big(\bar{\tau}_k^T g_k\alpha_k\widetilde{A}_k^{(1)} - \bar{\omega}_k^T e_k\alpha_k\widetilde{B}_k^{(1)} + \bar{\omega}_k^T e_k(\alpha_k\widetilde{B}_k^{(1)} + \beta_k\widetilde{B}_k^{(2)}) - \bar{\tau}_k^T g_k(\alpha_k\widetilde{A}_k^{(1)} + \beta_k\widetilde{A}_k^{(2)})\Big)\,. \end{aligned}$$

At this point we note that

$$\beta_k^2 \;=\; \|a_k - \alpha_k e\|^2 = \sum_{t\in\mathbb{T}}\left(a_k^t - \frac{1}{T}\sum_{t'\in\mathbb{T}} a_k^{t'}\right)^2 \;=\; \sum_{t\in\mathbb{T}}(a_k^t)^2 - \frac{1}{T}\left(\sum_{t\in\mathbb{T}} a_k^t\right)^2 \qquad (24)$$

whence $\beta_k = O(|k|\sqrt{T})$ from the definition of a_k^t and Assumption 2.11. Furthermore, by Assumption 2.12, $\bar{\tau}_k^T \to \sigma_{A,k}$ and $\bar{\omega}_k^T \to \sigma_{B,k}$ uniformly in k as $T\to\infty$. Hence, the variance of the first term of G_k^T scales with $|k|^2/T^2$, thus

$$(G^T)_1 := \sum_{|k|\le\xi_T}\frac{2\bar{\tau}_k^T\bar{\omega}_k^T\beta_k}{T^{3/2}}\Big(\widetilde{B}_k^{(1)}\widetilde{A}_k^{(2)} - \widetilde{B}_k^{(2)}\widetilde{A}_k^{(1)}\Big) = O_p\left(\sqrt{\sum_{|k|<\xi_T}\frac{|k|^2}{T^2}}\right) = O_p(\xi_T^2/T) \qquad (25)$$

i.e. with the hypothesis $\xi_T^4/T \to 0$, we obtain

$$\sqrt{T}\,(G^T)_1 \;\to\; 0 \text{ in probability.} \tag{26}$$

Let us further note at this point for future use in case of $\xi_T \to \infty$ with $\xi_T^4/T \to 0$ due to $\beta_k \le C|k|\sqrt{T}$ with a suitable constant $C > 0$, we have also that

$$|(G^T)_1| \;\le\; \xi_T^2 \frac{1}{\xi_T^2} \sum_{|k|<\xi_T} \frac{2\bar{\tau}_k^T \bar{\omega}_k^T C \xi_T}{T} \left| \widetilde{B}_k^{(1)} \widetilde{A}_k^{(2)} - \widetilde{B}_k^{(2)} \widetilde{A}_k^{(1)} \right| \;\to\; 0 \text{ a.s.} \tag{27}$$

The second term of G_k^T reduces to

$$\frac{2}{T}\Big(\bar{\omega}_k^T e_k \beta_k \widetilde{B}_k^{(2)} - \bar{\tau}_k^T g_k \beta_k \widetilde{A}_k^{(2)}\Big)$$

which is normally distributed with zero mean and variance

$$\begin{aligned} &\frac{4}{T^2}\beta_k^2\big((\bar{\tau}_k^T g_k)^2 + (\bar{\omega}_k^T e_k)^2\big) \\ &= \frac{16\pi^2\big((\bar{\tau}_k^T g_k)^2 + (\bar{\omega}_k^T e_k)^2\big)}{T^2} \left[\sum_{t\in\mathbb{T}} \left\langle k, \sum_{j=1}^d x_j \partial_{\vartheta_j} \delta_t^\vartheta \right\rangle^2 - \frac{1}{T}\left(\sum_{t\in\mathbb{T}} \left\langle k, \sum_{j=1}^d x_j \partial_{\vartheta_j} \delta_t^\vartheta) \right\rangle \right)^2 \right], \end{aligned}$$

for $\vartheta = \vartheta_0$ cf. (24). Since the normal random deviates in

$$(G^T)_2 := \sum_{|k|<\xi_T} \frac{2}{T}\Big(\bar{\omega}_k^T e_k \beta_k \widetilde{B}_k^{(2)} - \bar{\tau}_k^T g_k \beta_k \widetilde{A}_k^{(2)}\Big)$$

are independent in k, we have that $\sqrt{T}\,(G^T)_2$ is normally distributed with zero mean and variance converging to

$$\begin{aligned} &16\pi^2 \sum_{k\in\mathbb{Z}^2} \big((\sigma_{A,k} g_k)^2 + (\sigma_{B,k} e_k)^2\big) \left[\int_0^1 \Big\langle k, (\mathrm{grad}_\vartheta \delta_t^{\vartheta_0})' x \Big\rangle^2 dt - \Big\langle k, \int_0^1 (\mathrm{grad}_\vartheta \delta_t^{\vartheta_0})' x \, dt \Big\rangle^2 \right] \\ &=: \sigma_x^2 < \infty \end{aligned} \tag{28}$$

if $f \in H^1\big([0,1]\big)$. Recalling the notation of (22), (23) and $\sum_{|k|<\xi_T} G_k^T = (G^T)_1 + (G^T)_2 = \langle x, \mathrm{grad}_\vartheta M_T(\vartheta)\rangle$ as well as collecting the results of (26) and (28) we have thus shown that for any $0 \ne x \in \mathbb{R}^d$

$$\sqrt{T}\langle x, \mathrm{grad}_\vartheta M_T(\vartheta)\rangle \to \mathcal{N}(0, \sigma_x^2)$$

whenever $T, \xi_T \to \infty$ with ξ_T of rate $o(T^{1/4})$. Since this holds true for every x, the joint distribution of $\sqrt{T}\mathrm{grad}_\vartheta M_T(\vartheta)$ at $\vartheta = \vartheta_0$ is asymptotically multivariate normal with covariance matrix as asserted in Theorem 2.13.

In view of use below we note here that we obtain with suitable constants $C, C' > 0$ (C' due to Remark 2.5), $\sigma_{\max}$ from Assumption 2.7 and independent standard normal C_k

$(k \in \mathbb{Z})$ that

$$|(G^T)_2| = \left| \frac{2}{T} \sum_{|k|<\xi_T} \beta_k \sqrt{(\bar\tau_k^T g_k)^2 + (\bar\omega_k^T e_k)^2} C_k \right| \le \frac{2\sigma_{\max} C}{\sqrt{T}} \sum_{|k|<\xi_T} |f_k||k||C_k|$$
$$\le \frac{2\sigma_{\max} C C' \xi_T^2}{\sqrt{T}} \frac{1}{\xi_T^2} \sum_{|k|<\xi_T} |C_k| \to 0 \ \text{ a.s. if } \xi_T \to \infty \text{ and } \xi_T^4/T = O(1). \quad (29)$$

□

Remark 7.1. *As shown above, asymptotic normality of the second part* $\sqrt{T}\,(G^T)_2$ *of* $\sqrt{T}\,\mathrm{grad}_\vartheta \widetilde{M}_T(\vartheta_0)$ *holds regardless of the rate of* ξ_T*. If we relax* $\xi_T^4/T \to 0$ *to* $C_1 T^{1/4} \le \xi_T \le C_2 T^{1/4}$ *with suitable constants* $C_1, C_2 > 0$*, the first part* $\sqrt{T}\,(G^T)_1$ *will no longer converge to zero but will be tight, cf. (25). Since then also* $\hat\vartheta \to \vartheta_0$ *by Theorem 2.9, although the* $(G^T)_1$ *and* $(G^T)_2$ *will be dependent for this rate of* ξ_T*, we expect that asymptotic normality still holds. The corresponding covariance matrix, however, will have a more complicated structure than being a multiple of* $\tilde\Sigma$*.*

7.3 Proof of (ii) of Theorem 2.13

Here we build on the proof of (i) of Theorem 2.13 within the preceding section and use the notation there. In addition let $\mathbf{b}_{k,t}^\vartheta := 2\pi\,\mathrm{Hess}_\vartheta \langle k, \delta_t^\vartheta \rangle$. Then we obtain at once from (21)

$$\mathrm{Hess}_\vartheta \left(\sum_{t\in\mathbb{T}} h_k(\delta_t^\vartheta) Y_k^t \sum_{t'\in\mathbb{T}} \overline{h_k(\delta_{t'}^\vartheta) Y_k^{t'}} \right) = D_k^T + F_k^T$$

with

$$D_k^T := -2\,\mathrm{Im} \left(\sum_{t,t'\in\mathbb{T}} \mathbf{b}_{k,t}^\vartheta h_k(\delta_t^\vartheta) Y_k^t\, \overline{h_k(\delta_{t'}^\vartheta) Y_k^{t'}} \right)$$
$$F_k^T := -2\,\mathrm{Re} \left(\sum_{t,t'\in\mathbb{T}} \mathbf{a}_{k,t}^\vartheta (\mathbf{a}_{k,t}^\vartheta - \mathbf{a}_{k,t'}^\vartheta)' h_k(\delta_t^\vartheta) Y_k^t\, \overline{h_k(\delta_{t'}^\vartheta) Y_k^{t'}} \right).$$

In particular, in consequence of (7)

$$\mathrm{Hess}_\vartheta \widetilde{M}_T(\vartheta) = -\frac{1}{T^2} \sum_{|k|<\xi_T} (D_k^T + F_k^T)\,. \quad (30)$$

Note that $E(D_k^T) = 0$. Setting $\vartheta = \vartheta_0$ observe that the argument of the previous section (using the matrices $\mathbf{b}_{k,t}^\vartheta$ instead of the vectors $\mathbf{a}_{k,t}^\vartheta$) that led to (27) and (29) gives at once

$$\frac{1}{T^2} \sum_{|k|<\xi_T} D_k^T \to 0 \text{ a.s. if } T, \xi_T \to \infty \text{ and } \xi_T^4/T \to 0\,. \quad (31)$$

Likewise, the same follows for the random part of F_k^T. More precisely for $\vartheta = \vartheta_0$:

$$\begin{aligned} F_k^T &= -2 \sum_{t,t' \in \mathbb{T}} \mathbf{a}_{k,t}^{\vartheta_0} (\mathbf{a}_{k,t}^{\vartheta_0} - \mathbf{a}_{k,t'}^{\vartheta_0})' \\ &\quad \mathrm{Re}\Big(|f_k|^2 + f_k \overline{h_k(\delta_{t'}^{\vartheta_0}) W_k^{t'}} + h_k(\delta_t^{\vartheta_0}) W_k^t \overline{f_k} + h_k(\delta_t^{\vartheta_0}) W_k^t \, \overline{h_k(\delta_{t'}^{\vartheta_0}) W_k^{t'}} \Big) \\ &= -2 \sum_{t,t' \in \mathbb{T}} |f_k|^2 \mathbf{a}_{k,t}^{\vartheta_0} (\mathbf{a}_{k,t}^{\vartheta_0} - \mathbf{a}_{k,t'}^{\vartheta_0})' + \widetilde{F}_k^T \end{aligned}$$

with

$$\begin{aligned} &\widetilde{F}_k^T : \\ &= -2 \sum_{t,t' \in \mathbb{T}} \mathbf{a}_{k,t}^{\vartheta_0} (\mathbf{a}_{k,t}^{\vartheta_0} - \mathbf{a}_{k,t'}^{\vartheta_0})' \, \mathrm{Re}\Big(f_k \overline{h_k(\delta_{t'}^{\vartheta_0}) W_k^{t'}} + h_k(\delta_t^{\vartheta_0}) W_k^t \overline{f_k} + h_k(\delta_t^{\vartheta_0}) W_k^t \, \overline{h_k(\delta_{t'}^{\vartheta_0}) W_k^{t'}} \Big) \end{aligned}$$

yields

$$E(\widetilde{F}_k^T) = 0 \text{ and } \frac{1}{T^2} \sum_{|k| < \xi_T} \widetilde{F}_k^T \rightarrow 0 \text{ a.s. if } T, \xi_T \to \infty \text{ and } \xi_T^4 / T \to \infty . \qquad (32)$$

Since we have the deterministic limit

$$\sum_{|k| < \xi_T} \frac{2}{T^2} \sum_{t,t' \in \mathbb{T}} |f_k|^2 \mathbf{a}_{k,t}^{\vartheta_0} (\mathbf{a}_{k,t}^{\vartheta_0} - \mathbf{a}_{k,t'}^{\vartheta_0})' \rightarrow 2 \sum_{k \in \mathbb{Z}^2} |f_k|^2 \iint_{[0,1]^2} \mathbf{a}_{k,t}^{\vartheta_0} (\mathbf{a}_{k,t}^{\vartheta_0} - \mathbf{a}_{k,t'}^{\vartheta_0})' \, dt dt'$$

as $T, \xi_T \to \infty$ due to Assumption 2.11 on bounded total variation of first ϑ-derivatives, in conjunction with (30), (31) and (32) the definition of $\mathbf{a}_{k,t}^{\vartheta_0}$ yields the assertion (ii) of Theorem 2.13.

7.4 Proof of Theorem 2.14

Under Assumption 2.11, standard expansion arguments from M-estimation can be used as follows. Since $M_T(\vartheta)$ is twice continuously differentiable for ϑ near ϑ_0 and $\hat{\vartheta}_T$ converges a.s. to ϑ_0, we have that

$$\begin{aligned} 0 &= \mathrm{grad}_\vartheta M_T(\hat{\vartheta}_T) \\ &= \mathrm{grad}_\vartheta M_T(\vartheta_0) + \mathrm{Hess}_\vartheta M_T(\vartheta_0)(\hat{\vartheta}_T - \vartheta_0) + \Big(\mathrm{Hess}_\vartheta M_T(\hat{\vartheta}^*) - \mathrm{Hess}_\vartheta M_T(\vartheta_0) \Big)(\hat{\vartheta}_T - \vartheta_0) \end{aligned}$$

where $\hat{\vartheta}^*$ lies between ϑ_0 and $\hat{\vartheta}_T$. The continuity of the second derivatives gives that $\hat{\vartheta}_T - \vartheta_0$ and $\mathrm{grad}_\vartheta M_T(\vartheta_0)$ are of the same asymptotic order since $\mathrm{Hess}_\vartheta M_T(\vartheta_0) \to 8\pi^2 \Sigma$ a.s. holds by (ii) of Theorem 2.13. Hence

$$8\pi^2 \Sigma (\hat{\vartheta}_T - \vartheta_0) = -\,\mathrm{grad}_\vartheta M_T(\vartheta_0) + o_P(\|\hat{\vartheta}_T - \vartheta_0\|)$$

which in conjunction with (i) of Theorem 2.13, yields both asymptotic assertions.

7.5 Ad Example 2.15

Lemma 7.2. *In the situation of Example 2.15,* $\det(\Sigma) = 0$ *iff there is* $x \in \mathbb{R}^2 \setminus \{0\}$ *s.t.*

$$f(y + rx) = f(y) \quad \textit{for all } y \in \mathbb{R}^2, r \in \mathbb{R}, \tag{33}$$

where f *is* $[0,1]^2$*-periodic.*

Proof. Since for $x \in \mathbb{R}^2 \setminus \{0\}$ we have

$$x'\Sigma x = \frac{1}{12} \sum_{k \in \mathbb{Z}^2} |f_k|^2 \langle k, x\rangle^2 \geq 0,$$

the matrix Σ is positive semidefinite. Hence, $\det(\Sigma) = 0$ iff there is an $x \in \mathbb{R}^2 \setminus \{0\}$ s.t. $x'\Sigma x = 0$. This is the case iff

$$|f_k|^2 \neq 0 \text{ implies } \langle k, x\rangle^2 = 0 \quad \text{for all } k \in \mathbb{Z}^2. \tag{34}$$

If this implication holds, we have for all $y \in \mathbb{R}^2$ and $r \in \mathbb{R}$ that

$$f(y + rx) = \sum_{k \in \mathbb{Z}^2} f_k e^{2\pi i \langle k, y+rx\rangle} = \sum_{k \in \mathbb{Z}^2} f_k e^{2\pi i \langle k, y\rangle} e^{2\pi i r \langle k, x\rangle} = \sum_{k \in \mathbb{Z}^2} f_k e^{2\pi i \langle k, y\rangle} = f(y),$$

i.e. (33). If, on the other hand, (33) holds, then the two functions f and $f^{rx}(\,\cdot\,) := f(\,\cdot\, + rx)$ are identical. Subsequently, their respective Fourier coefficients f_k and $f_k^{rx} = e^{2\pi i r\langle k, x\rangle} f_k$ are also the same, i.e. (34) holds. □

7.6 Motion Blur Measure

To evaluate our drift correction we use a version of the motion blur measure m_2 proposed in (Xu et al., 2013) which is based on the work of (Chen et al., 2010). It is defined as

$$m_2 := \log\left(\frac{J(\varphi_{\max})}{J(\varphi_{\min})}\right). \tag{35}$$

Here, $J(\varphi) := \sum_{j=1}^{N^2} \Big(\Delta I\big((x_j)_1, (x_j)_2\big)_\varphi\Big)^2$ is the average squared directional derivative of an image I in direction $\big(\cos(\varphi), \sin(\varphi)\big)'$, $\varphi \in [0, 2\pi)$, $\varphi_{\min}$ is the motion direction, and $\varphi_{\max}$ is the direction perpendicular to $\varphi_{\min}$. Note, that $J(\varphi) = 0$ iff I is constant in direction φ. An advantage of m_2 is that it does not depend on the scale of the image. In (Chen et al., 2010), $\varphi_{\min}$ is selected as a minimizer of the functional J. The idea is that the image is blurred in the direction of the motion and thus the image intensity f

changes little in this direction (on average), while it varies much more in the perpendicular direction. The minimizer is obtained as follows:

Rewrite $J(\varphi) = \big(\cos(\varphi), \sin(\varphi)\big) D \big(\cos(\varphi), \sin(\varphi)\big)'$, where

$$D = \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix}, \quad d_{rs} := \sum_{j=1}^{N^2} \frac{\partial I}{\partial (x)_r}\big((x_j)_1, (x_j)_2\big) \cdot \frac{\partial I}{\partial (x)_s}\big((x_s)_j, (x_j)_2\big).$$

Then, $J(\varphi) = d_{11} + d_{12}\sin(2\varphi) + (d_{22} - d_{11})\big(\sin(\varphi)\big)^2$. We get the minimum value of J by setting $dJ(\varphi)/d\varphi = d_1 2\cos(2\varphi) + (d_{22} - d_{11})\sin(2\varphi) = 0$, which yields $\varphi = \varphi_m + (r\pi)/2$, $r \in \mathbb{Z}$, with $\varphi_m = \arctan\big(2d_{12}/(d_{11} - d_{22})\big)/2$. The motion direction is then determined by

$$\varphi_{\min} := \begin{cases} \varphi_m & \text{if } J(\varphi_m) \leq J(\varphi_m + \pi/2), \\ \varphi_m + \pi/2 & \text{if } J(\varphi_m) > J(\varphi_m + \pi/2). \end{cases}$$

The $J(\varphi_{\max})$ 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). In our simulation study, since we already know the true drift $\delta_t(\vartheta)$, we choose the average drift direction $\int_0^1 \partial\delta_t(\vartheta)/\partial t \, dt = \delta_1(\vartheta)$ as the motion direction (after normalization). Hence, in our context (where I is either $\hat{f}_T$ or the superimposed image, see Table 4) we get the motion blur measure

$$\tilde{m}_2 := \log\left(\frac{\sum_{j=1}^{N^2}\big\langle \operatorname{grad}_x I\big((x_j)_1, (x_j)_2\big), \operatorname{Rot}_{\pi/2}\delta_1(\vartheta)/||\delta_1(\vartheta)||_2\big\rangle^2}{\sum_{j=1}^{N^2}\big\langle \operatorname{grad}_x I\big((x_j)_1, (x_j)_2\big), \delta_1(\vartheta)/||\delta_1(\vartheta)||_2\big\rangle^2}\right), \tag{36}$$

where $||\cdot||_2$ is the Euclidean norm and

$$\operatorname{Rot}_{\pi/2} := \begin{pmatrix} \cos(\pi/2) & -\sin(\pi/2) \\ \sin(\pi/2) & \cos(\pi/2) \end{pmatrix}$$

is the rotation through $\pi/2$. Note that the average drift direction used to determine the motion blur (36) in the case of a drift function with jump is (before normalization)

$$t_0\delta_{t_0}(\vartheta) + (1 - t_0)\big(\delta_1(\vartheta) - \lim_{t \searrow t_0} \delta_t(\vartheta)\big)$$

instead of just $\delta_1(\vartheta)$, where t_0 is the time at which the jump occurs. We calculated an approximation of $\operatorname{grad}_x I$ as follows (see e.g. (Gonzalez, R.C. and Woods, R.E., 2002)).

Let I be a pixel image of size $M \times N$. For every pixel location (i, j), $i \in \{1, \dots, M\}$, $j \in \{1, \dots, N\}$, the gradient of I is defined as $\nabla I(i, j) := \big(G_x(i, j), G_y(i, j)\big)'$ with

$$G_x(i,j) := \sum_{i',j'=-1}^{1} S_x(i'{+}2, j'{+}2) I(i{+}i', j{+}j'), \quad G_y(i,j) := \sum_{i',j'=-1}^{1} S_y(i'{+}2, j'{+}2) I(i{+}i', j{+}j'),$$

		linear drift	quadratic drift	cubic drift	drift with jump
true parameter	ϑ_0	$(0.195, 0.117)$	$(0.195, 0.039, 0, 0.078)$	$(0.195, 0, 0.039, 0, 0.039, 0.195)$	$(0.312, 0.312, 0.156, 0.312, 0.156, 0.234, 0.5)$
error type	T	$\hat{\vartheta}_T$	$\hat{\vartheta}_T$	$\hat{\vartheta}_T$	$\hat{\vartheta}_T$
Gaussian	20	$(0.191, 0.115)$	$(0.179, 0.053, 0.022, 0.054)$	$(0.135, 0.108, -0.021, -0.018, 0.091, 0.153)$	$(0.377, 0.277, 0.183, 0.285, 0.059, 0.241, 0.5)$
	50	$(0.195, 0.121)$	$(0.201, 0.039, 0.001, 0.083)$	$(0.191, 0.015, 0.027, -0.006, 0.056, 0.184)$	$(0.329, 0.326, 0.16, 0.337, 0.228, 0.224, 0.53)$
	100	$(0.2, 0.119)$	$(0.188, 0.049, 0.011, 0.069)$	$(0.225, -0.064, 0.085, 0.003, 0.022, 0.204)$	$(0.32, 0.285, 0.162, 0.313, 0.2, 0.222, 0.48)$
t-distr.	20	$(0.189, 0.119)$	$(0.169, 0.062, 0.053, 0.016)$	$(0.144, 0.084, 0.002, -0.018, 0.097, 0.145)$	$(0.302, 0.376, 0.147, 0.286, 0.023, 0.266, 0.54)$
	50	$(0.193, 0.123)$	$(0.186, 0.046, 0.029, 0.056)$	$(0.194, 0.034, 0.004, 0.015, 0.083, 0.134)$	$(0.321, 0.304, 0.159, 0.313, 0.152, 0.241, 0.51)$
	100	$(0.203, 0.114)$	$(0.168, 0.072, 0.052, 0.022)$	$(0.205, -0.041, 0.078, 0.017, 0.07, 0.146)$	$(0.342, 0.384, 0.141, 0.3, 0.152, 0.234, 0.5)$
Poisson	20	$(0.183, 0.127)$	$(0.197, 0.03, 0.016, 0.075)$	$(0.116, 0.157, -0.054, 0.01, 0.086, 0.141)$	$(0.268, 0.352, 0.148, 0.337, 0.094, 0.263, 0.54)$
	50	$(0.203, 0.11)$	$(0.172, 0.062, 0.002, 0.076)$	$(0.181, 0.04, 0.013, -0.004, 0.02, 0.212)$	$(0.361, 0.293, 0.182, 0.318, 0.109, 0.245, 0.53)$
	100	$(0.193, 0.124)$	$(0.151, 0.081, 0.031, 0.047)$	$(0.147, 0.071, 0.009, -0.006, 0.061, 0.179)$	$(0.285, 0.317, 0.155, 0.325, 0.192, 0.226, 0.5)$

Table 1: *Displaying the estimated $\hat{\vartheta}_T$ 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.*

		linear drift	quadratic drift	cubic drift	drift with jump
true parameter	ϑ_0	$(0.195, 0.117)$	$(0.195, 0.039, 0, 0.078)$	$(0.195, 0, 0.039, 0, 0.039, 0.195)$	$(0.312, 0.312, 0.156, 0.312, 0.156, 0.234, 0.5)$
error type	T	mean of est's	mean of est's	mean of est's	mean of est's
Gaussian	20	$(0.196, 0.116)$	$(0.179, 0.056, 0.027, 0.051)$	$(0.151, 0.081, -0.001, 0.003, 0.064, 0.166)$	$(0.311, 0.316, 0.161, 0.314, 0.162, 0.235, 0.522)$
	50	$(0.195, 0.117)$	$(0.182, 0.052, 0.019, 0.06)$	$(0.177, 0.037, 0.017, -0.004, 0.074, 0.162)$	$(0.314, 0.311, 0.16, 0.316, 0.164, 0.234, 0.51)$
	100	$(0.195, 0.117)$	$(0.178, 0.056, 0.015, 0.064)$	$(0.168, 0.037, 0.026, -0.001, 0.07, 0.164)$	$(0.321, 0.31, 0.159, 0.305, 0.16, 0.231, 0.5)$
t-distr.	20	$(0.195, 0.114)$	$(0.177, 0.056, 0.028, 0.05)$	$(0.154, 0.071, 0.005, -0.011, 0.085, 0.155)$	$(0.305, 0.311, 0.171, 0.303, 0.166, 0.236, 0.517)$
	50	$(0.195, 0.117)$	$(0.182, 0.052, 0.022, 0.056)$	$(0.177, 0.034, 0.021, -0.001, 0.07, 0.163)$	$(0.312, 0.312, 0.16, 0.313, 0.161, 0.234, 0.509)$
	100	$(0.196, 0.116)$	$(0.176, 0.058, 0.016, 0.063)$	$(0.167, 0.046, 0.018, -0.001, 0.068, 0.166)$	$(0.311, 0.309, 0.159, 0.316, 0.157, 0.235, 0.506)$
Poisson	20	$(0.196, 0.116)$	$(0.174, 0.06, 0.021, 0.057)$	$(0.157, 0.063, 0.012, 0.001, 0.075, 0.156)$	$(0.311, 0.317, 0.162, 0.314, 0.155, 0.237, 0.524)$
	50	$(0.195, 0.117)$	$(0.174, 0.06, 0.021, 0.057)$	$(0.171, 0.045, 0.017, 0, 0.077, 0.154)$	$(0.322, 0.31, 0.164, 0.313, 0.156, 0.235, 0.514)$
	100	$(0.196, 0.117)$	$(0.176, 0.058, 0.024, 0.055)$	$(0.172, 0.033, 0.028, -0.006, 0.082, 0.155)$	$(0.312, 0.317, 0.157, 0.314, 0.159, 0.233, 0.506)$

Table 2: *Setting as in Table 1. Displaying the means of the estimators $\hat{\vartheta}_T$ 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	6e-3	5e-3	5e-3	26e-3	6e-3	8e-3	9e-3	8e-3	7e-3
Quadratic drift	63e-3	48e-3	44e-3	66e-3	54e-3	55e-3	65e-3	59e-3	61e-3
Cubic drift	138e-3	121e-3	133e-3	172e-3	130e-3	175e-3	142e-3	141e-3	144e-3
Drift with jump	79e-3	71e-3	67e-3	174e-3	80e-3	83e-3	87e-3	90e-3	86e-3

Table 3: *Root of the mean squared error* $\mathbb{E}||\hat{\vartheta}_T - \vartheta_0||^2$ *of the estimators* $\hat{\vartheta}_T$ *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 drift	0.067	0.050	0.006	−0.009	−0.009	−0.013	0.011	0.012	−0.053
	Quadratic drift	−0.005	0.011	−0.019	0.032	−0.009	−0.039	−0.031	−0.003	−0.108
	Cubic drift	−0.024	0.015	−0.073	−0.001	0.002	0.008	−0.016	0.048	−0.041
	Drift with jump	0.013	−0.034	0.029	0.007	−0.015	−0.015	0.016	0.031	−0.055
$\hat{f}_T$	Linear drift	−0.679	−0.842	−0.707	−0.205	−0.102	−0.192	−0.387	−0.318	−0.338
	Quadratic drift	−0.411	−0.447	−0.432	−0.147	−0.060	−0.128	−0.205	−0.188	−0.179
	Cubic drift	−0.686	−1.045	−0.710	−0.215	−0.112	−0.218	−0.375	−0.358	−0.514
	Drift with jump	−0.201	−0.326	−0.582	−0.096	−0.217	−0.072	−0.078	−0.123	−0.289

Table 4: *Blur measure values of the superimposed images (SI) and the estimated images* $\hat{f}_T$*. The corresponding estimators* $\hat{\vartheta}_T$ *are reported in Table 1. The images for cubic drift, drift with jump and* $T \in \{20, 50\}$ *are shown in Figures 3 and 4.*

where we extend the image periodically, i.e. $I(0,j) := I(M,j)$, $I(M+1,j) := I(1,j)$, $I(i,0) := I(i,N)$, 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, e.g. 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 are very noisy, we repeat that procedure once more.

7.7 Simulations: Tables

In this subsection we display the simulation results from Section 3.

Table 1 summarizes the drift parameter estimators $\hat{\vartheta}_T$ for one simulation in different drift models, error models, and for different image sequence lengths T. Table 2 displays the means of those estimators $\hat{\vartheta}_T$ from 100 simulations each. Table 3 shows the roots of the mean squared errors of the same estimators $\hat{\vartheta}_T$. Finally, Table 4 lists the blur measure values of the superimposed images and the estimated images $\hat{f}_T$ corresponding to the parameters in Table 1.

We used images of size 256×256 pixels and the image sequences had lengths $T \in \{20, 50, 100\}$. The drift functions were polynomials in t (time) of degree 1, 2, or 3, or piecewise linear with a jump. The true drift parameters θ_0 are shown in Tables 1 and 2. For the Gaussian and t_2-distributed errors we chose a noise level of $\sigma = 0.1$. The Fourier cutoff was set to $\xi_T = \sqrt{T}$ and the start value for the minimization algorithm was $0 \in \mathbb{R}^d$, where d is the dimension of the drift parameter ϑ_0. For each scenario (drift model, error model, number of frames T), we performed 100 simulations.

References

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.

Gamboa, F., Loubes, J.-M., and Maza, E. (2007). Semi-parametric estimation of shifts. *Electron. J. Statist.*, 1:616–640.

Gonzalez, R.C. and Woods, R.E. (2002). *Digital Image Processing.* Prentice Hall, 2 edition.

Sen, P. K. and Singer, J. M. (1993). *Large Sample Methods in Statistics.* Chapman & Hall.

van der Vaart, A. W. (2000). *Asymptotic Statistics.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

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.