27. Fachtagung „Experimentelle Strömungsmechanik“
3. – 5. September 2019, Erlangen

Anhang zu Korrektur dynamischer Fehler für direkte Spektralschätzung von ungleichmäßig abgetasteten Daten inklusive korrelierter Abtastintervalle

Appendix to Bias Correction for Direct Spectral Estimation from Irregularly Sampled Data Including Sampling Schemes with Correlation

H. Nobach, N. Damaschke und V. Kühn

A Herleitungen der Korrekturfaktoren für verschiedene Abtastschemen

Bei unabhängiger Abtastung mit stetiger Verteilung der Abtastintervalle sind alle Kreuzprodukte unabhängig voneinander und somit ist $\beta'_k = 1$ für alle $k \neq 0$. Für $k = 0$ ergeben n Abtastungen in einem bestimmten Zeitintervall insgesamt n^2 Selbst- und Kreuzprodukte. Die Anzahl n der Abtastungen in einem Zeitintervall ist Poisson-verteilt mit der Wahrscheinlichkeit

$$p(n) = \frac{\alpha'^n}{n!} \mathrm{e}^{-\alpha'}, \tag{1}$$

mit der mittleren Anzahl α'. Der Erwartungswert der n^2 Selbst- und Kreuzprodukte für $k = 0$ ist dann

$$\sum_{n=0}^{\infty} n^2 p(n) = \alpha' + \alpha'^2. \tag{2}$$

Nach der Normierung auf α'^2 erhält man schließlich $\beta'_0 = 1 + {}^1\!/_{\alpha'}$.
Bei diskreter Verteilung der Abtastintervalle mit unabhängigen Fehlstellen sind alle Kreuzprodukte unabhängig voneinander und somit ist $\beta'_k = 1$ für alle $k \neq 0$. Im Gegensatz zum vorherigen Fall, bei dem mehrere Abtastungen in einem Zeitschritt erfolgen können, kann bei nominell äquidistanter Abtastung pro Zeitschritt nur entweder genau eine oder keine gültige Abtastung erfolgen. Deshalb treten in $R(0)$ nur Selbstprodukte auf, die ihrerseits mit einer mittleren Anzahl von α' pro Zeitschritt auftreten. Nach der Normierung auf α'^2 erhält man schließlich $\beta'_0 = {}^1\!/_{\alpha'}$.
Bei korrelierter Abtastung mit der Korrelation γ zwischen zwei Zeitschritten, in denen n_1 und n_2 Abtastungen erfolgen, besteht der Zusammenhang

$$n_2 = \gamma n_1 + (1 - \gamma) n'_2, \tag{3}$$

wobei γn_1 der korrelierte Teil in n_2 und $(1 - \gamma) n'_2$ der von n_1 unabhängige Teil ist, wobei n_1 und n'_2 beide den Mittelwert α' haben. Das Produkt $n_1 n_2$ lässt sich dann zu

$$n_1 n_2 = \gamma n_1^2 + (1 - \gamma) n_1 n'_2. \tag{4}$$

umformen.
Bei stetiger Verteilung der Abtastintervalle hat n_1^2 den Erwartungswert $\alpha' + \alpha'^2$, bei diskreter Verteilung der Abtastintervalle hat n_1^2 dagegen den Erwartungswert α' (s. Herleitungen zuvor für unabhängige Abtastung). Weil n_1 und n'_2 unabhängig sind, hat ihr Produkt in beiden Fällen den Erwartungswert α'^2. Für die stetige Verteilung der Abtastintervalle folgt schließlich

$$\mu_{n_1 n_2} = \gamma(\alpha' + \alpha'^2) + (1 - \gamma)\alpha'^2 = \gamma\alpha' + \alpha'^2 \tag{5}$$

und nach Normierung auf α'^2 folgt

$$\beta'_k = \frac{\mu_{n_1 n_2}}{\alpha'^2} = 1 + \frac{\gamma_k}{\alpha'}. \tag{6}$$

Für die diskrete Verteilung der Abtastintervalle folgt

$$\mu_{n_1 n_2} = \gamma\alpha' + (1-\gamma)\alpha'^2 \tag{7}$$

und nach Normierung auf α'^2

$$\beta'_k = \frac{\mu_{n_1 n_2}}{\alpha'^2} = 1 - \gamma_k + \frac{\gamma_k}{\alpha'}. \tag{8}$$

B Herleitungen der Korrelationskoeffizienten für verwendete Abtastschemen mit Korrelation

Für das Abtastschema b mit stetiger Verteilung der Abtastintervalle mit einem Mindestabtastintervall zwischen aufeinanderfolgenden Werten wird der Korrelationskoeffizient γ_0 aus der Verteilung der Anzahl der Abtastungen n in einer bestimmten Zeitschritt abgeleitet. Für n Abtastungen in einem Zeitschritt beinhaltet der Korrelationskoeffizient n^2 Selbst- und Kreuzprodukte, was zu

$$\gamma_k = \sum_{n=0}^{\lfloor 1/d' \rfloor + 1} n^2 P_0(n) \tag{9}$$

führt, mit dem normierten Mindestintervall zwischen den Abtastungen $d' = d/\Delta\tau$ und der Wahrscheinlichkeit $P_0(n)$, dass es n Abtastungen innerhalb eines Zeitschritts $\Delta\tau$ gibt. Um die Wahrscheinlichkeit $P_0(n)$ zu erhalten, müssen auch Abtastungen vor dem untersuchten Zeitintervall berücksichtigt werden, da ihre Totzeit die Wahrscheinlichkeit darauffolgender Abtastungen innerhalb des betrachteten Intervalls beeinflussen. Die Wahrscheinlichkeit, dass keine vorhergehende Abtastung das betrachtete Zeitintervall beeinflusst, ist $P_e = 1 - \alpha' d'$ (e - empty). Die Wahrscheinlichkeit, dass das Zeitintervall durch die Totzeit einer vorangehenden Abtastung vollständig abgedeckt wird, ist $P_f = \alpha' \max\{0, d'-1\}$ (f - full). In diesem Fall können innerhalb des untersuchten Zeitintervalls keine weiteren Abtastungen erfolgen. Dieser Fall kann nur dann eintreten, wenn $d' > 1$. Die Wahrscheinlichkeit, dass das betrachtete Zeitintervall zum Teil durch die Totzeit einer vorangehenden Abtastung blockiert wird, ist $P_p = \alpha' \min\{1, d'\}$ (p - partial). Die Wahrscheinlichkeit, genau n Abtastungen in einem Zeitschritt $\Delta\tau$ zu haben, besteht aus der Summe dieser drei Fälle, also

$$P_0(n) = P_e(n) + P_f(n) + P_p(n), \tag{10}$$

wobei $P_e(n)$, $P_f(n)$ und $P_p(n)$ die Wahrscheinlichkeiten sind, n Abtastungen innerhalb des betrachteten Zeitintervalls zu erhalten, entweder ohne vorangehende Abtastung, mit vollständiger oder mit teilweiser Blockierung des Intervalls. Die Wahrscheinlichkeiten dieser drei Teile sind

$$P_e(n) = \left(1 - \alpha' d'\right)\left[P'\left(n, 1 - d' \max\{0, n-1\}\right) - P'\left(n+1, 1 - nd'\right)\right] \tag{11}$$

$$P_f(n) = \begin{cases} \alpha' \max\{0, d'-1\} & \text{für } n = 0 \\ 0 & \text{sonst} \end{cases} \tag{12}$$

$$P_p(n) = \int_0^{\min\{1,d'\}} \alpha' \left[P'\left(n, x - d' \max\{0, n-1\}\right) - P'\left(n+1, x - nd'\right)\right] \mathrm{d}x \tag{13}$$

wobei $P'(n,x)$ die Wahrscheinlichkeit ist, mindestens n Abtastungen innerhalb des (normierten) Teils x eines Zeitschrittes zu erhalten, nämlich

$$P'(n,x) = 1 - \sum_{j=0}^{n-1} \frac{(\alpha'' x)^j}{j!} \mathrm{e}^{-\alpha'' x} \tag{14}$$

mit der normierten mittleren Anzahl von Abtastungen α'' in der Restzeit zwischen den jeweiligen Totzeiten der angenommenen Abtastungen

$$\alpha'' = \frac{\alpha'}{1 - \alpha' d'} \tag{15}$$

Für die Korrelationskoeffizienten γ_k mit $k \neq 0$, können nur Kreuzprodukte zwischen verschiedenen Abtastungen auftreten. Dafür wird eine Abtastung zum Zeitpunkt t_1 innerhalb eines Zeitschritts angenommen und eine zweite Abtastung zum Zeitpunkt t_2 innerhalb eines anderen Zeitschritts, die k Zeitschritte voneinander entfernt sind. Weil die Autokorrelation symmetrisch ist, werden nur die Fälle mit $k > 0$ betrachtet. Für $k < 0$ können die Betrachtungen durch Verwendung von $|k|$ übertragen werden. Die erste Abtastung kann zu jedem Zeitpunkt t_1 innerhalb eines Zeitschritts mit der mittleren Rate α erfolgen. Die Normierung auf einen Zeitschritt $\Delta\tau$ ergibt die normierte Zeiten $t_1' = {}^{t_1}\!/_{\Delta\tau}$ bzw. $t_2' = {}^{t_2}\!/_{\Delta\tau}$. Demnach erfolgt die erste Abtastung mit einer mittleren Anzahl pro Zeitschritt von α'. Die Abtastung zum zweiten Zeitpunkt t_2' hängt von der Anzahl n weiterer Abtastungen zwischen t_1' und t_2' ab. Betrachtet man auch hier nur die Intervalle zwischen den durch die Totzeit blockierten Zeiten, dann treten in der verbleibenden Zeit weitere Abtastungen mit der Abtastrate α'' auf, genau wie zuvor für $k = 0$. Die Anzahl n weiterer Abtastungen zwischen t_1' und t_2' und ihre Totzeiten nd' und zusätzlich die Totzeit der ersten Abtastung bei t_1', führt schließlich dazu, dass die Abtastungen bei t_1' und t_2' nicht unabhängig voneinander erfolgen. Ohne Einschränkung der Allgemeinheit, liege t_1' im Zeitschritt mit der Nummer null. Dann gilt für den Korrelationskoeffizient

$$\gamma_k = \int\limits_{-0.5}^{0.5} \int\limits_{|k|-0.5}^{|k|+0.5} \mathrm{d}t_1' \mathrm{d}t_2' \alpha' \alpha'' \sum_{n=0}^{\left\lfloor {}^{(t_2'-t_1')}\!/_{d'} - 1 \right\rfloor} P(n, t_2' - t_1' - (n+1)d') \tag{16}$$

wobei $P(n, x)$ die Wahrscheinlichkeit ist, dass n weitere Abtastungen innerhalb des (normierten) Zeitintervalls x auftreten, nämlich

$$P(n, x) = \frac{(\alpha'' x)^n}{n!} \mathrm{e}^{-\alpha'' x} \tag{17}$$

Für das Abtastschema d mit diskreter Verteilung der Abtastintervalle mit korrelierten Datenlücken wird die Anzahl n der Wechsel von gültigen zu ungültigen Daten bzw. umgekehrt betrachtet. Zwischen zwei Zeitpunkten, die $|k|$ Zeitschritte auseinander liegen, können bis zu $|k|$ solcher Wechsel erfolgen, die in der durchgeführten Simulation unabhängig voneinander mit der mittleren Anzahl pro Zeitschritt von $c' = 0.2$ erfolgen. Ein solcher Wechsel tritt zu einem bestimmten Zeitschritt mit der Wahrscheinlichkeit c' auf, kein Wechsel entsprechend mit der Wahrscheinlichkeit $1 - c'$. Eine gerade Anzahl solcher Wechsel trägt positiv zur Korrelation γ_k bei, während eine ungerade Anzahl von Wechseln die Korrelation entsprechend verringert. Die Wechsel können zu jedem Zeitpunkt zwischen $\Delta\tau$ und $|k|\,\Delta\tau$ erfolgen, wobei nur die Anzahl von Wechseln von Bedeutung ist, ihre Reihenfolge dagegen nicht. Die Korrelation von Datenlücken bzw. den gültigen Abtastungen wird damit zu

$$\gamma_k = \sum_{n=0}^{|k|} (-1)^n \binom{|k|}{n} \left(1 - c'\right)^{|k|-n} c^n = \left(1 - 2c'\right)^{|k|} . \tag{18}$$