

Supplemental Materials

Molecular Biology of the Cell

Berro et al.

SUPPLEMENTAL FIGURE LEGENDS

SUPPLEMENTAL FIGURE S1. Score calculation in the proof of the continuous alignment method. Blue, reference dataset; red, dataset to align. $f(x)$, value of the original signal at time x ; t , unknown offset; t_{offset} , original offset between both datasets; $t_{sampling}$, time interval between two consecutive measurements of the signal; x_i , original time points where the reference dataset has been measured; x_{scorei} , time points where the score function is evaluated.

SUPPLEMENTAL FIGURE S2. Interface of the PatchTrackingTools used for patch tracking and quality control (toolbox on the right side). The specimen is a live wild type fission yeast cell (middle left) expressing Fim1p-mEGFP. The bottom left image is the sum projection of cells imaged with a spinning disk confocal microscope over 5 consecutive 360 nm-spaced z-slices. The row of images at the top left shows the five consecutive z-slices for the patch in the yellow box in the bottom left image. The bottom right image is a montage of fluorescence micrographs generated automatically where each row corresponds to a z-slice and each column to a different time point. This allows the user to check quickly that all the fluorescence of the patch of interest has been collected with no other patch interfering. Images on the left part of the figure are color-coded in function of pixel intensities using the ImageJ “Fire” lookup table (dark blue, lower intensities; bright orange, higher intensities). The bottom right montage has reversed contrast (white, lower intensities; black, higher intensities).

SUPPLEMENTAL FIGURE S3. (A) Timing variability for the number of molecules of fimbrin for different patches. Each dot is the size of the horizontal gray lines in Figure 3B, i.e. represents the standard deviation along the x-axis for the normalized number of molecules. (B) Comparison of the quality of three alignment methods: alignment on the peak value; alignment minimizing the variability between datasets at the sampling resolution (1 s); and continuous alignment (using a 100 ms resolution). The points are the average and the vertical lines are the standard deviation of the Root Mean Square Difference (RMSD) between each dataset and the averaged dataset for the 24 tracks used in Figure 3D.

SUPPLEMENTAL FIGURE S4. Z-test for the average displacements of endocytic patches plotted in Figure 5B. In each figure, we perform a z-test at 95% testing the null hypothesis that the displacement at each time point is the same significance as at a reference time point (A: 0 s; B: 3 s; C: 6 s; D: 8 s, green dots on the plots; red line: threshold z-score to reject the null hypothesis with 95% confidence). (A) The displacements between time 4 s and 8 s are significantly different from the displacement at time 0 s. (B) The displacements between time 5 s and 7 s are significantly different from the displacement at time 3 s. (C) The displacements between time 0 s and 4 s are significantly different from the displacement at time 6 s. (D) The quality of the data at time 8 s is too low to conclude that the displacement at time 8 s is significantly different than the displacements at different time points except for $t = 0$ s.

SUPPLEMENTAL FIGURE S5. Comparison of patches imaged at the bottom edge or in the middle plane of the cell. The resolution in z is smaller than the resolution in x and y in confocal microscopy. Therefore, data from patches from the middle of the cell give an accurate description of the movement perpendicular to the membrane and data from patches from the bottom of the cell give an accurate description of the movement parallel to the membrane. No significant difference is found between patches tracked in the middle or at the bottom of the cell. (A) Number of molecules; (B) Displacement; (C) Distance from origin. Olive, patches at the bottom of the cell; teal, patches in the middle of the cell; green, all patches pooled together. Dark colors, average values; light colors, standard deviations. Inset of (A) The relative standard deviation in the number of molecules is on average 20%.

SUPPLEMENTAL FIGURE S6. Temporal evolution of patch distributions in cells from Figure 6. (A) Number of patches. (B) Patch distribution in cells (red, densest tip; green: middle of the cell; blue, less dense tip). (C) Tip symmetry index. (D) OP_{50} index. In all panels, each subpanel corresponds to the cell with same color in Figure 8.

Supplemental materials

Proof of Continuous Alignment Method

Julien Berro

Let us consider 2 datasets of N_{sample} datapoints that represent two independent measurements of the same signal $f(x)$ at sampling intervals $t_{sampling}$ (Supplemental figure S1A). Consider these two samples have been measured at an unknown offset t_{offset} from each other. The continuous alignment method aims to find this offset by minimizing the square difference between the piecewise linear interpolation of the sampled data for different offset t . Here we prove that the value of the offset t that minimizes this score is an estimate of the original unknown offset t_{offset} .

We call $f_{Reference}(x)$, the piecewise linear interpolation of one datasets (blue curve, Supplemental figure S1) and $f_{ToAlign}(x, t)$, the piecewise linear interpolation of the other dataset, translated along the x axis with the offset t (blue curve, Supplemental figure S1).

The score is evaluated every $t_{sampling}$ at $x_{score_i} = x_{score_0} + i \times t_{sampling}$ (gray arrows, Supplemental figure S1). Let us assume without loss of generality $x_i + t \leq x_{score_i} \leq x_{i+1}$.

For the following theorem, we also assume that the signal $f(x)$ is at least one time differentiable and that the higher order derivatives and the sampling interval are such that $f(x+t) \approx f(x) + t \times f'(x)$ for $0 \leq t \leq t_{sampling}$. This condition occurs asymptotically when $t_{sampling}$ tends to 0.

Theorem 1. *The score function of the continuous alignment method*

$$score(t) = \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}-1} \|f_{Reference}(x_{score_i}) - f_{ToAlign}(x_{score_i}, t)\|^2$$

is minimum for $t = t_{offset}$

Proof:

$$\begin{aligned} score(t) &= \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}-1} \|f_{Reference}(x_{score_i}) - f_{ToAlign}(x_{score_i}, t)\|^2 \\ &= \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}} \left\| \left\{ f(x_i) + \frac{f(x_{i+1}) - f(x_i)}{t_{sampling}} (x_{score_i} - x_i) \right\} \right. \\ &\quad \left. - \left\{ f(x_i + t_{offset}) + \frac{f(x_{i+1} + t_{offset}) - f(x_i + t_{offset})}{t_{sampling}} (x_{score_i} - x_i - t) \right\} \right\|^2 \\ &= \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}} \left\| \left\{ f(x_i) + \left(\sum_{k=1}^{N_{diff}} f^{(k)}(x_i) \frac{t_{sampling}^k}{k!} + o(t_{sampling}^{N_{diff}}) \right) \frac{(x_{score_i} - x_i)}{t_{sampling}} \right\} \right. \\ &\quad \left. - \left\{ f(x_i) + \sum_{k=1}^{N_{diff}} f^{(k)}(x_i) \frac{t_{offset}^k}{k!} + o(t_{offset}^k) \right. \right. \\ &\quad \left. \left. + \left(\sum_{k=1}^{N_{diff}} f^{(k)}(x_i) \frac{(t_{sampling} + t_{offset})^k - t_{offset}^k}{k!} + o(t_{sampling}^{N_{diff}}) \right) \frac{(x_{score_i} - x_i - t)}{t_{sampling}} \right\} \right\|^2 \end{aligned}$$

$$= \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}} \left\| \sum_{k=1}^{N_{diff}} \frac{f^{(k)}(x_i)}{k!} \left(-t_{offset}^k + t_{sampling}^k \frac{(x_{score_i} - x_i)}{t_{sampling}} \right) + \left((t_{sampling} + t_{offset})^k - t_{offset}^k \right) \frac{(x_{score_i} - x_i - t)}{t_{sampling}} + o(t_{sampling}^{N_{diff}}) \right\|^2$$

Keeping the first differentiate only ($N_{diff}=1$) (see below)

$$score(t) \approx \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}} (f'(x_i)(t - t_{offset}))^2 = (t - t_{offset})^2 \left\{ \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}} f'(x_i)^2 \right\}$$

which is minimum (actually 0) when $t = t_{offset}$.

Note: the continuous alignment method becomes less efficient when the original signal has large second order derivative between 2 sampling points, that is the signal is far from a straight line between 2 consecutive sampling points. A higher sampling frequency (by decreasing $t_{sampling}$) will therefore improve the continuous alignment. Note also that this formula shows that the continuous alignment method gives better results for signals with large first order derivatives. Indeed, if the first order derivative is close to 0 (typically close to a straight line parallel to the x-axis), the continuous alignment method will not give good results, just like the regular ‘‘manual’’ alignment method, as expected intuitively.

Presence of noise in the measured data

Now, let us assume that each measurement contains noise. We note ε_x the measurement noise of the reference dataset at time x . Let assume that each measurement noise is independent from each other and is distributed around the same distribution with same average et same variance σ^2 .

Noting $\delta = x_{score_i} - x_i$, we now have

$$f_{Reference_WithNoise}(x_{score_i}) = f_{Reference_NoNoise}(x_{score_i}) + \varepsilon_{x_i} + \frac{\varepsilon_{x_i+t_{sampling}} - \varepsilon_{x_i}}{t_{sampling}} \delta$$

$$f_{ToAlign_WithNoise}(x_{score_i}, t) = f_{ToAlign_NoNoise}(x_{score_i}, t) + \varepsilon_{x_i+t_{offset}} + \frac{\varepsilon_{x_i+t_{offset}+t_{sampling}} - \varepsilon_{x_i+t_{offset}}}{t_{sampling}} (\delta - t)$$

We will now show that the score function of the continuous alignment method is still minimum for $t = t_{offset}$ in the presence of ‘reasonable’ noise in the measured data.

Indeed,

$$score_{WithNoise}(t) = \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}} \left\| score_{NoNoise_i}(t) + \varepsilon_{x_i} - \varepsilon_{x_i+t_{offset}} + \frac{\varepsilon_{x_i+t_{sampling}} - \varepsilon_{x_i}}{t_{sampling}} \delta - \frac{\varepsilon_{x_i+t_{offset}+t_{sampling}} - \varepsilon_{x_i+t_{offset}}}{t_{sampling}} (\delta - t) \right\|^2$$

$$= \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}} \left(score_{NoNoise_i}(t)^2 + \left(\varepsilon_{x_i} - \varepsilon_{x_i+t_{offset}} + \frac{\varepsilon_{x_i+t_{sampling}} - \varepsilon_{x_i}}{t_{sampling}} \delta - \frac{\varepsilon_{x_i+t_{offset}+t_{sampling}} - \varepsilon_{x_i+t_{offset}}}{t_{sampling}} (\delta - t) \right)^2 \right. \\ \left. + 2 score_{NoNoise_i}(t) \left(\varepsilon_{x_i} - \varepsilon_{x_i+t_{offset}} + \frac{\varepsilon_{x_i+t_{sampling}} - \varepsilon_{x_i}}{t_{sampling}} \delta - \frac{\varepsilon_{x_i+t_{offset}+t_{sampling}} - \varepsilon_{x_i+t_{offset}}}{t_{sampling}} (\delta - t) \right) \right)$$

$$= score_{NoNoise}(t) + \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}} \left(\varepsilon_{x_i} - \varepsilon_{x_i+t_{offset}} + \frac{\varepsilon_{x_i+t_{sampling}} - \varepsilon_{x_i}}{t_{sampling}} \delta - \frac{\varepsilon_{x_i+t_{offset}+t_{sampling}} - \varepsilon_{x_i+t_{offset}}}{t_{sampling}} (\delta - t) \right)^2 \\ + \frac{1}{N_{sample} - 1} \sum_{i=1}^{N_{sample}} 2 score_{NoNoise_i}(t) \left(\varepsilon_{x_i} - \varepsilon_{x_i+t_{offset}} + \frac{\varepsilon_{x_i+t_{sampling}} - \varepsilon_{x_i}}{t_{sampling}} \delta - \frac{\varepsilon_{x_i+t_{offset}+t_{sampling}} - \varepsilon_{x_i+t_{offset}}}{t_{sampling}} (\delta - t) \right)$$

The third term of the right-hand side is smaller than

$$2 \text{score}_{\text{NoNoise}}(t) \frac{1}{N_{\text{sample}} - 1} \sum_{i=1}^{N_{\text{sample}}} \left(\varepsilon_{x_i} - \varepsilon_{x_i+t_{\text{offset}}} + \frac{\varepsilon_{x_i+t_{\text{sampling}}} - \varepsilon_{x_i}}{t_{\text{sampling}}} \delta - \frac{\varepsilon_{x_i+t_{\text{offset}}+t_{\text{sampling}}} - \varepsilon_{x_i+t_{\text{offset}}}}{t_{\text{sampling}}} (\delta - t) \right)$$

The sum will be close to zero for large N_{sample} , since it is an estimator of the difference of the average error at x_i and at $x_i + t_{\text{offset}}$

Since the noise measurements are independent, and since $\delta/t_{\text{sampling}} < 1$ (it is equal to 0 when the score is evaluated at x_i), the second term of the right-hand side is an estimator of

$$2\sigma^2 \left(1 + \frac{\delta^2 + (\delta - t)^2}{t_{\text{sampling}}^2} \right) \approx 2\sigma^2$$

and

$$\text{score}_{\text{WithNoise}}(t) \approx \frac{1}{N_{\text{sample}} - 1} \sum_{i=1}^{N_{\text{sample}}} \|f'(x_i)(t - t_{\text{offset}})\|^2 + 2\sigma^2$$

Therefore, the continuous alignment method still works even in the presence of noise.

Note 1: large noise decreases the efficiency of the method, especially when the variance becomes of the same order of magnitude as the second order term of the Taylor polynomial or when the estimation of the mean and the variance become less accurate for small numbers of sampling points.

Note 2: in specific cases the continuous alignment method can be seen as a variant of the least-square method. In general it is not, since both signals to align contain independent noise. However, if the reference signal does not contain noise, the continuous alignment method is equivalent to the least-square method. This is the case when the continuous alignment method is run iteratively, using the average of the tracks aligned in the previous run as the reference function. The least-square method has been shown empirically, and formally in some cases, to give the best estimator of the parameter t , i.e. an unbiased estimator with minimum variance.

Average of aligned datasets

With the canonical discrete alignment of the datasets, a systematic error is made when averaging all the different datasets. Indeed, if the real offset was t_k , the value at position x ($x_i \leq x \leq x_i + t_{\text{sampling}}$) of the linearly interpolated data is:

$$f_{\text{DiscreteAlignment}}(x) = f(x_i + t_k) + \frac{f(x_i + t_k + t_{\text{sampling}}) - f(x_i + t_k)}{t_{\text{sampling}}} (x - x_i)$$

Noting that $\frac{f(x_i+t_k+t_{\text{sampling}})-f(x_i+t_k)}{t_{\text{sampling}}} = f'(x_i + t_k) + O(t_{\text{sampling}})$ and $(x - x_i) = O(t_{\text{sampling}})$, we can write

$$f_{\text{DiscreteAlignment}}(x) = f(x_i + t_k) + f'(x_i + t_k)(x - x_i) + o(t_{\text{sampling}})$$

Noting that $(x - x_i) = (x - x_i - t_k) + t_k$, we can now write

$$f_{\text{DiscreteAlignment}}(x) = f(x_i + t_k) + f'(x_i + t_k)(x - x_i - t_k) + f'(x_i + t_k)t_k + o(t_{\text{sampling}})$$

Therefore

$$f_{\text{DiscreteAlignment}}(x) = f(x) + f'(x_i + t_k)t_k + o(t_{\text{sampling}})$$

This shows that when aligning the datasets on the discrete sampling time points, one makes a systematic error on the estimation of the function, roughly equal to $f'(x_i + t_k)t_k$.

Theorem 2. *The average of datasets aligned with the canonical discrete alignment method contains a systematic error roughly equal to*

$$\frac{1}{N_{\text{Datasets}}} \sum_{k=1}^{N_{\text{Datasets}}} f'(x_i + t_k)t_k$$

at position x , $x_i \leq x \leq x_i + t_{\text{sampling}}$

Proof:

The average value of all the N_{Datasets} interpolated data at position x is:

$$\begin{aligned} \text{Average}_{\text{DiscreteAlignment}}(x) &= \frac{1}{N_{\text{Datasets}}} \sum_{k=1}^{N_{\text{Datasets}}} f_{\text{DiscreteAlignment}}(x) \\ &= \frac{1}{N_{\text{Datasets}}} \sum_{k=1}^{N_{\text{Datasets}}} f(x) + f'(x_i + t_k)t_k + o(t_{\text{sampling}}) \\ &\approx f(x) + \frac{1}{N_{\text{Datasets}}} \sum_{k=1}^{N_{\text{Datasets}}} f'(x_i + t_k)t_k \end{aligned}$$

Theorem 3. *The average of datasets aligned with the continuous alignment method at position x converges to the value of the original signal $f(x)$*

Proof:

For x such as $x_i + t_k \leq x \leq x_i + t_k + t_{\text{sampling}}$

$$\begin{aligned} f_{\text{ContinuousAlignment}}(x) &= f(x_i + t_k) + \frac{f(x_i + t_k + t_{\text{sampling}}) - f(x_i + t_k)}{t_{\text{sampling}}}(x - x_i - t_k) \\ &= f(x_i + t_k) + f'(x_i + t_k)(x - x_i - t_k) + o(t_{\text{sampling}}) \\ &= f(x) + o(t_{\text{sampling}}) \end{aligned}$$

As a consequence,

$$\text{Average}_{\text{ContinuousAlignment}}(x) = f(x) + o(t_{\text{sampling}})$$

Presence of noise in the measured data

Corollary2. *Theorem 2 and 3 are still valid in the presence of noise in the measured signal.*

Proof:

Let now assume that each measurement at position x contains independent noise, noted ε_x . We now have:

$$f_{\text{ContinuousAlignment}}(x) \approx f(x) + \varepsilon_{x_i+t_k} + \frac{\varepsilon_{x_i+t_k+t_{\text{sampling}}} - \varepsilon_{x_i+t_k}}{t_{\text{sampling}}}(x - x_i - t_k) \text{ for } x_i + t_k \leq x \leq x_i + t_k + t_{\text{sampling}}/2$$

and

$$f_{\text{ContinuousAlignment}}(x) \approx f(x) + \varepsilon_{x_i+t_k} + \frac{\varepsilon_{x_i+t_k} - \varepsilon_{x_i+t_k-t_{\text{sampling}}}}{t_{\text{sampling}}}(x - x_i - t_k) \text{ for } x_i + t_k - t_{\text{sampling}}/2 \leq x \leq x_i + t_k$$

We note

$$\Delta_\varepsilon(x, k) = \begin{cases} \varepsilon_{x_i+t_k} - \varepsilon_{x_i+t_k-t_{\text{sampling}}}, & \text{if } x_i + t_k \leq x \leq x_i + t_k + t_{\text{sampling}}/2 \\ \varepsilon_{x_i+t_k+t_{\text{sampling}}} - \varepsilon_{x_i+t_k}, & \text{if } x_i + t_k \leq x \leq x_i + t_k + t_{\text{sampling}}/2 \end{cases}$$

and $\tau_k = (x - x_i - t_k)/t_{\text{sampling}}$. Since t_k is uniformly distributed in $] - t_{\text{sampling}}/2, t_{\text{sampling}}/2 [$, τ_k is uniformly distributed in $] - 1/2, 1/2 [$.

Note that $\Delta_\varepsilon(x, k)$ is the difference of two independent random variables and therefore its mean converges to 0. We now have,

$Average_{Continuous_WithNoise}(x)$

$$= Average_{Continuous_NoNoise}(x) + \frac{1}{N_{Datasets}} \sum_{k=1}^{N_{Datasets}} \varepsilon_{x_i+t_k} + \frac{1}{N_{Datasets}} \sum_{k=1}^{N_{Datasets}} \Delta_\varepsilon(x, k)\tau_k$$

If all the noise measurements follow the same distribution with mean 0, then for large number of datasets both sums will converge to 0 and

$$Average_{Continuous_WithNoise}(x) \approx Average_{Continuous_NoNoise}(x)$$

A similar property can be shown for the average of data aligned with the canonical discrete alignment method.

Theorem 4. *The variance of the sampled data aligned with the continuous alignment method is*

$$Variance_{Continuous_WithNoise}(x) \approx \frac{7}{6}\sigma^2$$

where σ^2 is the variance of the measurement noise

Proof:

$$\begin{aligned} Variance_{Continuous_WithNoise}(x) &= \frac{1}{N_{Datasets}} \sum_{k=1}^{N_{Datasets}} \|(f(x) + \varepsilon_{x_i+t_k} + \Delta_\varepsilon(x, k)\tau_k + o(t_{sampling})) - (f(x) + o(t_{sampling}))\|^2 \\ &= \frac{1}{N_{Datasets}} \sum_{k=1}^{N_{Datasets}} \|\varepsilon_{x_i+t_k} + \Delta_\varepsilon(x, k)\tau_k + o(t_{sampling})\|^2 \\ &= \frac{1}{N_{Datasets}} \sum_{k=1}^{N_{Datasets}} \varepsilon_{x_i+t_k}^2 + \Delta_\varepsilon(x, k)^2\tau_k^2 + 2\varepsilon_{x_i+t_k}\Delta_\varepsilon(x, k)\tau_k + o(t_{sampling}) \end{aligned}$$

Therefore, $Variance_{Continuous_WithNoise}(x)$ converges to

$var(\varepsilon_{x_i+t_k}) + var(\Delta_\varepsilon(x, k)\tau_k) + 2cov(\varepsilon_{x_i+t_k}\Delta_\varepsilon(x, k), \tau_k)$ for large $N_{Datasets}$.

$\varepsilon_{x_i+t_k}\Delta_\varepsilon(x, k)$ and τ_k being independent variates, their covariance will converge to 0 for large $N_{Datasets}$.

In addition, $\Delta_\varepsilon(x, k)$ and τ_k are independent, therefore

$$var(\Delta_\varepsilon(x, k)) \times var(\tau_k) = 2 var(\varepsilon_{x_i+t_k}) \times 1/12 = \sigma^2/6$$

And $Variance_{Continuous_WithNoise}(x) = \frac{7}{6}\sigma^2$

Theorem 5. *The variance of the sampled data aligned with the canonical discrete alignment method is*

$$Variance_{Discrete_WithNoise}(x) \approx \frac{7}{6}\sigma^2 + var(f'(x_i + t)t)$$

where σ^2 is the variance of the measurement noise and

$$var(f'(x_i + t)t) = \frac{1}{N_{Datasets}} \sum_{k=1}^{N_{Datasets}} (f'(x_i + t_k)t_k - Av(f'(x_i + t)t))^2$$

and $Av(f'(x_i + t)t) = \frac{1}{N_{Datasets}} \sum_{k=1}^{N_{Datasets}} f'(x_i + t_k)t_k$

Proof:

We note $Av(f'(x_i + t)t) = \frac{1}{N_{Datasets}} \sum_{k=1}^{N_{Datasets}} f'(x_i + t_k)t_k$

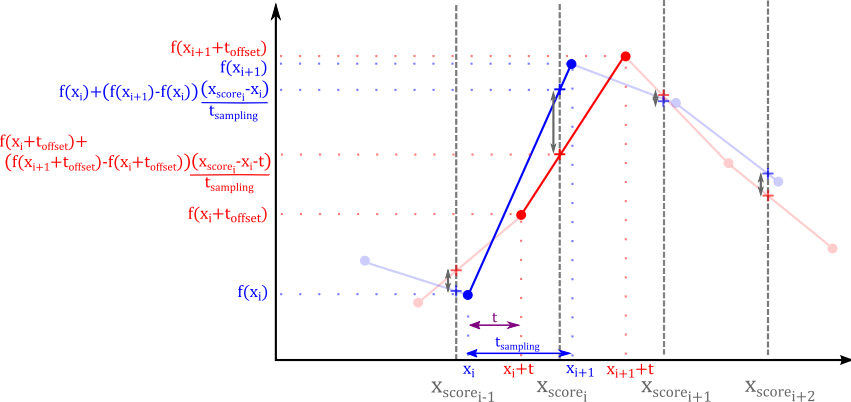
$$\begin{aligned}
\text{Variance}_{\text{DiscreteWithNoise}}(x) &= \frac{1}{N_{\text{Datasets}}} \sum_{k=1}^{N_{\text{Datasets}}} \left\| \left(f(x) + f'(x_i + t_k)t_k + \varepsilon_{x_i+t_k} + \Delta_\varepsilon(x, k)\tau_k + o(t_{\text{sampling}}) \right) \right. \\
&\quad \left. - \left(f(x) + Av(f'(x_i + t)t) + o(t_{\text{sampling}}) \right) \right\|^2 \\
&= \frac{1}{N_{\text{Datasets}}} \sum_{k=1}^{N_{\text{Datasets}}} \left\| \varepsilon_{x_i+t_k} + \Delta_\varepsilon(x, k)\tau_k + (f'(x_i + t_k)t_k - Av(f'(x_i + t)t)) + o(t_{\text{sampling}}) \right\|^2 \\
&= \frac{1}{N_{\text{Datasets}}} \sum_{k=1}^{N_{\text{Datasets}}} \varepsilon_{x_i+t_k}^2 + \Delta_\varepsilon(x, k)^2 \tau_k^2 + (f'(x_i + t_k)t_k - Av(f'(x_i + t)t))^2 \\
&\quad + 2(f'(x_i + t_k)t_k - Av(f'(x_i + t)t))(\varepsilon_{x_i+t_k} + \Delta_\varepsilon(x, k)\tau_k) + 2\varepsilon_{x_i+t_k}\Delta_\varepsilon(x, k)\tau_k \\
&\quad + o(t_{\text{sampling}}) \\
&\approx \text{var}(\varepsilon_{x_i+t_k}) + \text{var}(\Delta_\varepsilon(x, k)\tau_k) + \text{var}(f'(x_i + t)t) + 2\text{cov}(\varepsilon_{x_i+t_k}\Delta_\varepsilon(x, k), \tau_k) \\
&\quad + \frac{1}{N_{\text{Datasets}}} \sum_{k=1}^{N_{\text{Datasets}}} 2f'(x_i + t_k)t_k\Delta_\varepsilon(x, k)\tau_k \text{ for large } N_{\text{Datasets}}
\end{aligned}$$

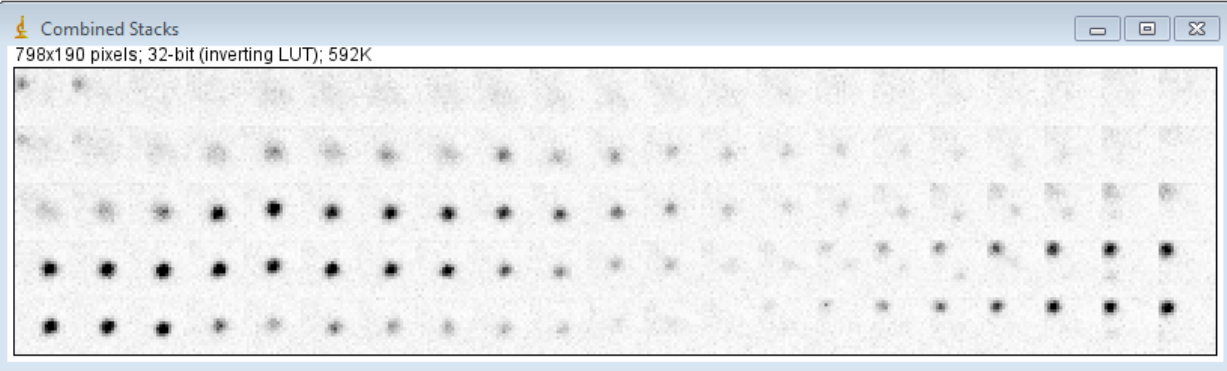
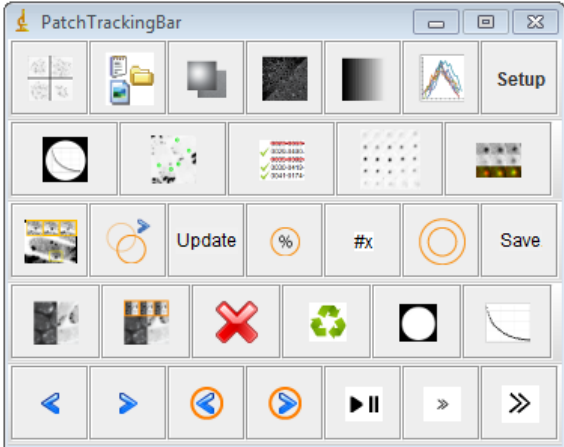
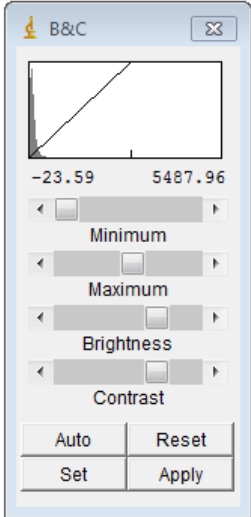
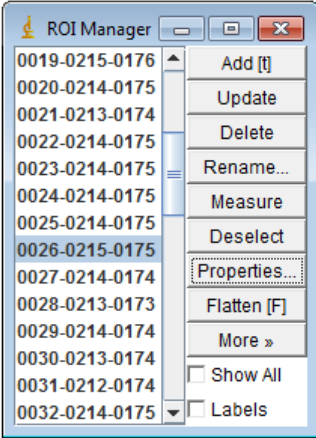
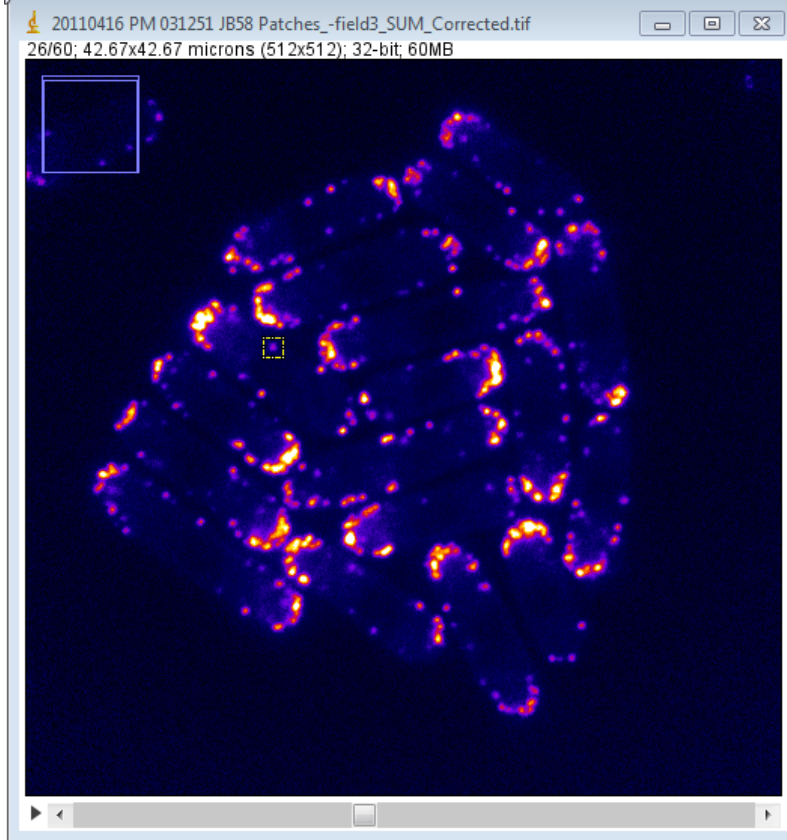
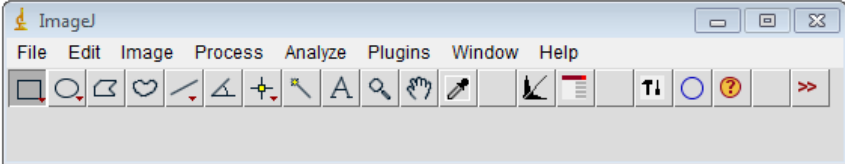
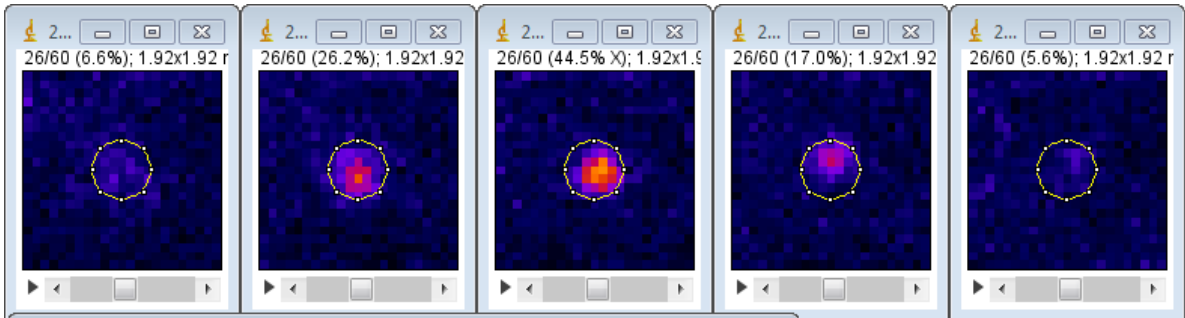
The covariance and the last sum actually converge to 0 for large N_{Datasets} .

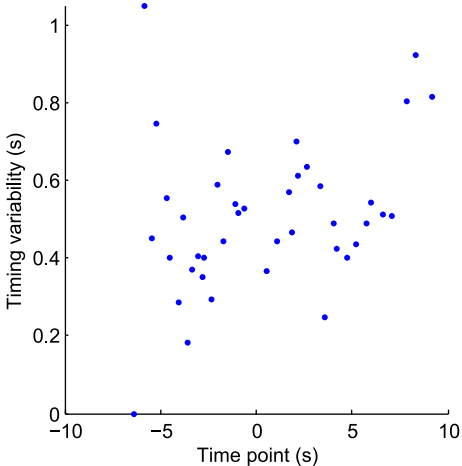
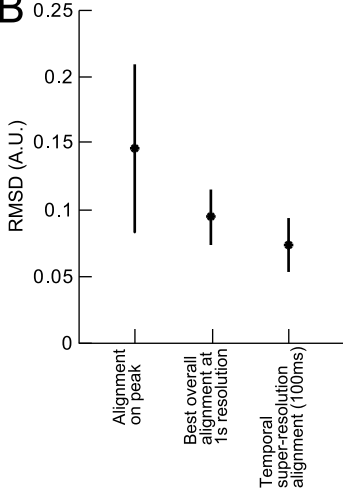
Therefore

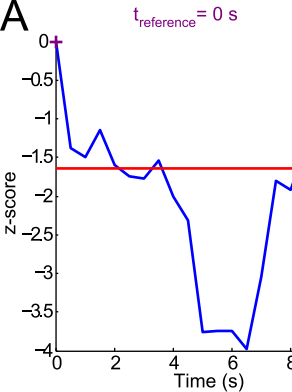
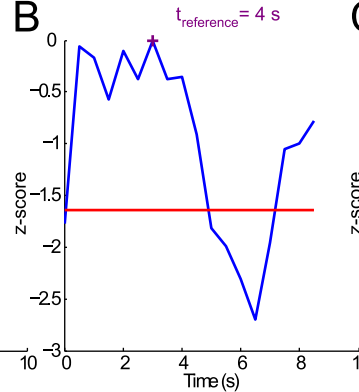
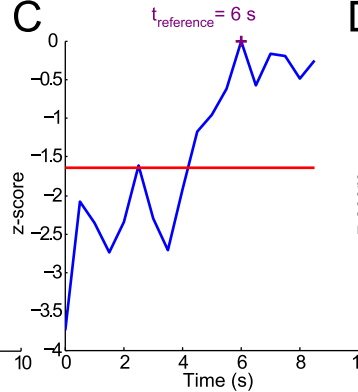
$$\text{Variance}_{\text{DiscreteWithNoise}}(x) = \frac{7}{6}\sigma^2 + \text{var}(f'(x_i + t)t).$$

In conclusion, the canonical discrete alignment misestimates the mean and the variance of the original signal.





A**B**

A**B****C****D**