# Documentation for BNEW analysis error quantification

E. F. Koslover, C. K. Chan, J. A. Theriot

Last updated June 28, 2015

This package provides code for approximating errors in the fit parameters obtained via BNEW analysis. The large covariance matrices for the rescaled time-averaged $\widehat{\mathrm{MSD}}$ are calculated using Fortran 90 code. This code reads in a parameter file `param.arg` where `arg` is a command-line argument. The parameter file contains keyword and argument pairs on each line, with the possible keywords described NEEDREF. A required keyword is ACTION, which can be set to GETCOVARMAT in order to calculate the components that go into the covariance matrix $\mathbf{M}$ (Eq.S33) for a single exponentially decorrelating drift velocity, or to GETDRIFTCOUPLEMAT to calculate the coupling matrix for two drift velocities with different correlation times ($\sum_q \mathbf{F}^{(u,1)}\mathbf{F}^{(u,2)}$; see Eq.36).

The resulting matrices are parsed and error estimates are calculated using Matlab scripts.

## 1 Quick-start Example

Compile the code by typing the following in the `source` directory. You will need the `gfortran` compiler and the BLAS and LAPACK libraries installed (present by default in MacOS).

```
make
```

Run a calculation for drift velocity with correlation time $\tau = 30$.

```
./waveletCov.exe example1
```

An output file `covar.t30.txt` will be generated, which should match (within numerical error) to the provided file `covar.t30.txt_backup`.

Run a calculation for drift velocity with correlation time $\tau = 10$.

```
./waveletCov.exe example2
```

Run a calculation to get the coupling matrix for drift velocities with $\tau = 10$ and $\tau = 30$.

```
./waveletCov.exe example_couple
```

Parse the resulting matrices and calculate the estimated fit parameter errors as a function of drift velocity magnitude $\gamma$ using the matlab script `BNEWerrors_example.m`. This script will generate plots as shown in Fig. 1.

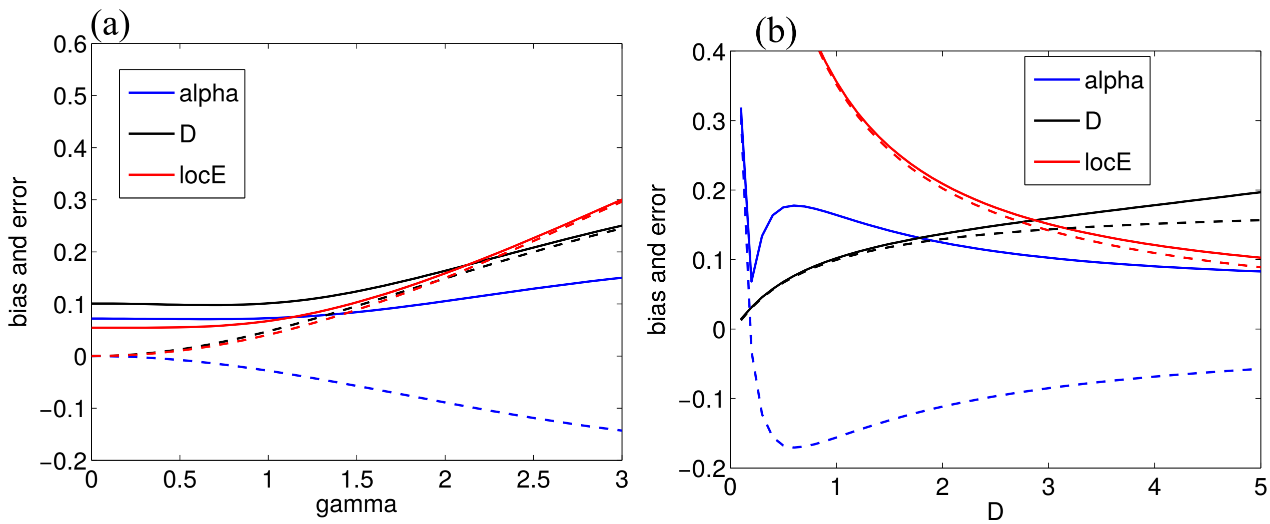


Figure 1: Error estimates for fitted parameters obtained with BNEW analysis. (a) Relative bias (dashed lines) and root mean square error (solid lines) for parameters as a function of drift velocity magnitude $\tilde{\gamma}$, assuming $\tilde{D} = 3, \tilde{\tau} = 30$. (b) Relative bias (dashed) and root mean square error (solid), plotted as a function of dimensionless diffusion constant $\tilde{D}$, for the case where the drift velocity is composed of two independent persistent random walks with $\tilde{\gamma}_1 = 1, \tilde{\tau}_1 = 30$ and $\tilde{\gamma}_2 = 1, \tilde{\tau}_2 = 10$. Errors are calculated assuming 100 independent tracks of length 200 datapoints each. SG3 wavelets with spans $2 \leq n \leq 13$ and time separations $1 \leq k \leq \lfloor 0.74n \rfloor$ were used for the analysis.

# 2 Calculation Details

## 2.1 ACTION GETCOVARMAT

For a single persistent random walk drift , the code calculates matrices used for obtaining the covariance of the adjusted (not rescaled) MSD datapoints. These calculations depend only on the correlation time $\tilde{\tau}$ (TAU keyword) and the length of each track (TRACKLEN keyword), and not on the other model parameters $\tilde{D}, \tilde{\gamma}$ or the number of tracks. The output file (name given by the OUTFILE keyword) begins with one line giving the value of $n_{\max}$ and one line listing the maximal $k$ value used for each span $n$. The calculated matrices are then written in the order $\mathbf{H}^{(u)}, \mathbf{H}^{(v)}, \mathbf{H}^{(\xi)}, \mathbf{F}^{(u,v)}, \mathbf{F}^{(u,\xi)}, \mathbf{F}^{(v,\xi)}$, where the matrices correspond to individual terms in Eq. S27, summed over the time separations $q$. Finally, the file contains the values of the rescaling functions $A_k^{(n)}, B_k^{(n)}$ and the second-order correlation for the drift velocity (last term of equation 15).

The individual matrices for each value of $q$ can be optionally saved into binary files (SAVEHMATFILE keyword) and can then be read in directly from those files (HMATFROMFILE keyword) to enable rapid calculation with different values of the track length, so long as the correlation time $\tilde{\tau}$ of the drift stays constant.

Keep in mind that the number of matrix elements scales as $n_{\max}^4$, so large values of $n_{\max}$ may use up all available RAM. Code has only been tested with $n_{\max} \leq 30$.

## 2.2 GETDRIFTCOVARMAT

This calculations finds the coupling matrix between two different drift velocities (last term in equation S36), summed over the time separations $q$. The time constants for the two drift velocities are given by the TAU keyword. Optionally, precalculated matrices for individual drift velocities can be loaded from binary files (HMATFROMFILE keyword) to make this calculation very rapid.

# 3 Keyword Parameters

The parameter file read in by the code has on each line a keyword followed by one or more values, separated by spaces. Lines beginning with # are treated as comments. Blank lines are ignored. The possible keywords are listed below.

- *ACTION*
  - value: 1 string (case insensitive)
  - Specify which calculation to run. Accepted values: GETCOVARMAT, GETDRIFT-COUPLEMAT
- *DEG*
  - value: 1 integer
  - Degree of wavelet. Used for Savitzky-golay wavelets only.
- *HMATFROMFILE*
  - value: 1 logical, 2 optional strings
  - defaults: F `Hmat_Q#.bin`
  - Read in precalculated matrices from binary files
  - Strings are file names, where the last instance of # is replaced with the value of Q.
  - Only first file name is used for GETCOVARMAT calculation. For the GETDRIFT-COUPLEMAT calculation, the two file names correspond to the two different drift velocities.
- *KSCL*
  - value: 1 float; default: 1.0
  - Maximal value of $k$ to use in the calculation. For each wavespan $n$, the values $1 \leq k \leq \lfloor k_{\mathrm{scl}} n \rfloor$ are used.
- *NMAX*
  - value: 1 integer; default 5
  - Maximal wavelet span to use in calculation.
- *OUTFILE*
  - value: 1 string; default: `*.txt`
  - Output file for calculations. If "*" is present in the filename, the last instance will be replaced by the command line argument `arg` used in calling the code.
- *SAVEHMATFILE*

- value: 1 logical, 1 optional strings
- defaults: F
- Save matrices for each value of $q$ into binary files.
- Used for GETCOVARMAT calculation only.
- Optionally, supply name of file to save to, with last instanceo # replaced with the value of Q.
- Default file name is the value specified by HMATFROMFILE.

- *TAU*
  - value: 1 or 2 floats
  - Correlation time(s) for drift velocity. Only first value is used in the GETCOVARMAT calculation; both values are required for GETDRIFTCOUPLEMAT calculation.

- *TRACKLEN*
  - value: 1 integer
  - Length of each track ($N + 1$ in the notation used in the paper)

- *WAVETYPE*
  - value: 1 string, case insensitive
  - Shape of wavelet to use. Options are: SVG (Savitzky-Golay wavelets), Haar, MEAN (sliding mean wavelet), and MSD (null wavelet for plain MSD)