# **Dye/DNA Plates Documentation**

Release v2.2.0

Robert F. DeJaco

# **CONTENTS:**

1	Modules				
	1.1	Wells			
	1.2	Get Data	3		
	1.3	Quantifying Noise and Linearity	5		
	1.4	Parameter Extraction			
	1.5	Plotting	9		
		1.5.1 Raw Data	9		
		1.5.2 Quantifying Noise and Linearity	10		
		1.5.3 Parameters			
	1.6	Util	12		
2	Indices and tables				
Рy	thon ]	Module Index	17		

**CHAPTER** 

**ONE** 

# **MODULES**

# 1.1 Wells

 $src.wells.column\_row\_to\_well$  (ix: int, iy: int)  $\rightarrow$  str Convert indices to well

```
>>> column_row_to_well(0, 0)
'A1'
>>> column_row_to_well(0, 7)
'H1'
>>> column_row_to_well(11, 0)
'A12'
>>> column_row_to_well(11, 7)
'H12'
>>> well_to_row(column_row_to_well(11, 7))
7
>>> well_to_column(column_row_to_well(11, 7))
11
```

### **Parameters**

- ix (int) a index
- **iy** (*int*) y index

Returns well

Return type str

 ${\tt src.wells.number\_to\_column}$  (number: int)  $\to$  int

```
>>> number_to_column(1)
1
>>> number_to_column(0)
0
>>> number_to_column(11)
11
>>> number_to_column(95)
11
>>> number_to_column(84)
0
```

 ${\tt src.wells.number\_to\_row}\ (number:\ int)\ o {\tt int}$ 

```
>>> number_to_row(0)
0
>>> number_to_row(1)
0
>>> number_to_row(11)
0
>>> number_to_row(12)
1
>>> number_to_row(95)
7
>>> number_to_row(84)
7
```

 $\verb|src.wells.number_to_well| (\textit{number: int}) \rightarrow \textit{str}$ 

Return number of well

```
>>> number_to_well(0)
'A1'
>>> number_to_well(11)
'A12'
>>> number_to_well(95)
'H12'
>>> number_to_well(84)
'H1'
```

 $src.wells.well_to_column(well:str) \rightarrow int$ 

Convert well name to column

Parameters well (str) – name of well

**Returns** ix – a-index for well

Return type int

 $src.wells.well\_to\_number(well: str) \rightarrow int$  Return number of well

```
>>> well_to_number("A1")
0
>>> well_to_number("A12")
```

(continues on next page)

(continued from previous page)

```
11
>>> well_to_number("H12")
95
>>> well_to_number("H1")
84
```

 $src.wells.well_to_row(well: str) \rightarrow int$  Well to y index

```
>>> well_to_row('H100')
7
>>> well_to_row('H1')
7
>>> well_to_row('A1')
0
>>> well_to_row('B1')
1
>>> well_to_row('Z1')
Traceback (most recent call last):
...
ValueError: 'Z' is not in list
```

Parameters well (str) - well name

**Returns** iy - y-index of well

Return type int

# 1.2 Get Data

```
class src.get_data.CombinedData(*replicates)
```

Combined data from several replicate plates. See changes to F and C in  $Raw\ Data\ and\ Scaling\ portion$  of  $Results\ and\ Discussion$ .

### Variables

```
• \mathbf{F} (np.ndarray) – Fluorescence data, \mathbf{F}_{\mathbf{D}}^t in Equation (16a)
```

- C(np.ndarray) Dye concentrations,  $C_D^t$  in Equation (16a)
- D (float) DNA concentration, D in Equation (16a)
- **t** (str) Type of DNA, t, "SS" or "DS"
- N(int) number of nucleobases per single strand, N
- $\mathbf{M\_tls}$  (np.array)  $\mathbf{M}^{\mathrm{TLS}}$ , set externally, defaults to np.array([])
- C\_hat  $(np.array) \hat{C}$ , set externally, defaults to np.array([])
- $V_M(np.array) V(M)$  (see Section S1.2), defaults to np.array([])
- $V_C(np.array) V(C)$  (see Section S1.2), defaults to np.array([])
- M\_std  $(np.array) \sqrt{V(\mathbf{M})}$  (see Section S1.2), defaults to np.array([])
- C\_std  $(np.array) \sqrt{V(C)}$  (see Section S1.2), defaults to np.array([])

**\_\_init\_\_** (\*replicates)  $\rightarrow$  None Combine replicate F and C

1.2. Get Data 3

**Parameters replicates** (typing.List[Data]) – list of plates to gather together as replicates

### $make\_subset(F\_min)$

Make subset of data as described in Raw Data and Scaling portion of manuscript

**Note:** F and C are overwritten.

class  $src.get_data.RawData$  (fluorescence\_file\_name, B\_d, t, l, N,  $dye\_conc\_file\_name='dye\_conc\_uM.csv'$ )

Stores raw data.

### **Variables**

- **F** (np.ndarray) Fluorescence data,  $F_d^{t\ell}$  (see Equation 11 of main text)
- **C** (np.ndarray) Dye concentrations, C (see Equation 10 of main text)
- **B** (float) DNA concentration,  $B_d$  in mol/L
- t(str) Type of DNA, t, "SS" or "DS" or "None".
- 1 (str) Replicate name,  $\ell$  is A, B, or C
- N (int) number of nucleobases per single strand

\_\_init\_\_ (fluorescence\_file\_name, B\_d, t, l, N, dye\_conc\_file\_name='dye\_conc\_uM.csv')

Scale the data before interpolating/solving optimization problem.

#### **Parameters**

- **fluorescence\_file\_name** (str) name of fluorescence file within data folder
- **B\_d** (float) Total concentration of nucleobases in mol/L  $B_d$
- **dye\_conc\_file\_name** (*str*, *optional*) name of dye concentration file name within data folder, defaults to "dye\_conc\_uM.csv"
- **t** (str) Type of DNA, "SS", "DS", or "None".
- 1 (str) Replicate name,  $\ell$  is A, B, or C
- N (int) number of nucleobases per single strand

```
src.get_data.excel_to_data(f_name: str, channel='GREEN')
```

Convert "Raw Data" sheet of excel file to pandas dataframe. Uses Equation (9) of manuscript to calculate temperature associated with each cycle.

#### **Parameters**

- f name (str) name of excel file
- channel (str, optional) name of channel to investigate, defaults to "GREEN"

**Returns** Formatted data frame, with wells sorted from A1, A2 ... H11, H12

Return type pd.DataFrame

```
src.get_data.get_C (file_name)
```

Get total dye concentration associated with each well.

**Parameters file\_name** (*str*) – CSV file formatted like a 96-well plate. The top left corner looks like

Row	1	2
A		
В		
С		

The values are concentrations of dye in units of mol/L

**Returns** Mapping of well name ("A1",...) to dye concentration [units of mol/L]

Return type Dictionary

# 1.3 Quantifying Noise and Linearity

```
src.noise_removal.compute_M_LS(F, C)
```

Calculate M by least-squares approximation

**Returns**  $M^{LS}$ , see Equation (17)

Return type np.array

 $\verb|src.noise_removal.compute_M_plus|(F, c\_plus)|$ 

Get updated guess for M

### **Parameters**

- **F** (np.ndarray) Fluorescence matrix **F**
- **c\_plus** (np.array) Concentration matrix updated  $c_+$

**Returns**  $M_+$  by Equation (S3b)

Return type np.array

 $src.noise\_removal.compute\_c\_plus(F, C, M\_minus, rho\_squared)$ Compute updated guess of concentrations,  $c_+$ 

#### **Parameters**

- **F** (np.ndarray) Fluorescence **F**
- **C** (np.array) Dye Concentration **C**
- M\_minus (np.array) Guess for M, M\_
- rho\_squared (float) Weight,  $\rho^2$

**Returns** c<sub>+</sub> by Equation (S3a)

Return type np.ndarray

 $src.noise\_removal.predictor\_corrector(F, C, rho\_squared, maxiter=100000, print\_iter=True)$ 

Solve Equation (18) with predictor-corrector algorithm

### **Parameters**

- **F** (np.ndarray) Fluorescence data **F**
- **C** (np.ndarray) Dye concentration data **C**
- rho\_squared (float) Weighting factor for concentrations,  $\rho^2$  in Equation (18a)
- maxiter (int, optional) maximum iterations allowed, by default 100000

• **print\_iter** (bool, optional) – whether or not to print the total number of iterations performed, by default True

**Returns**  $(\mathbf{M}, \mathbf{c})$  – solution,  $(\mathbf{M}^{\text{TLS}}, \widehat{\mathbf{C}})$ 

**Return type** tuple(np.array, np.array)

# 1.4 Parameter Extraction

Stores multiple instances of CombinedData for one DNA type

#### Variables

- N(int) number of nucleobases, N
- M1 (np.array)  $\mathbf{M}^{TLS}$  associated with  $\mathbf{D} = 1$
- M2 (np.array)  $\mathbf{M}^{\mathrm{TLS}}$  associated with  $\mathbf{D}=2$ . Several high temperatures are removed to reflect smaller temperature range associated with  $\mathbf{D}=1$
- C1  $(np.array) \widehat{\mathbf{C}}$  associated with  $\mathbf{D} = 1$
- C2  $(np.array) \widehat{\mathbf{C}}$  associated with  $\mathbf{D} = 2$
- $\mathbf{r}$  (np.array) r, as defined in Equation (S7).
- **V\_C1**  $(np.array) V(\mathbf{C})$  associated with  $\mathbf{D} = 1$
- V\_C2 (np.array) V(C) associated with D = 2
- **V\_M1**  $(np.array) V(\mathbf{M})$  associated with  $\mathbf{D} = 1$
- V\_M2 (np.array) V(mathbfM) associated with D=2. Several high temperatures are removed to reflect smaller temperature range associated with D=1
- dT (float) Change in temperature from one cycle to next,  $\Delta T$
- \_\_\_init\_\_\_(cls1: src.get\_data.CombinedData, cls2: src.get\_data.CombinedData)
  Initialize data

**Note:** Since different temperature ranges for each, need to make subset of dataset that has more temperatures. Dataset with lower DNA concentration cls1 always has less temperatures.

### **Parameters**

- cls1 (CombinedData) Data of DNA type at  $\mathbf{D}=1$
- cls2 (CombinedData) Data of DNA type at  $\mathbf{D}=2$

### $\textbf{get}\_\textbf{K} \ ( \ ) \ \rightarrow numpy.array$

Get K from vectorized version of Equation (21)

Returns K

Return type np.array

 $\texttt{get}\_\texttt{K\_std}\,(\,) \, \to \text{numpy.array}$ 

Get standard deviation estimate of **K** 

Returns

$$\sqrt{\Delta \mathbf{M}^2 \left(4V(\mathbf{M}_1) + 2V(\mathbf{M}_2)\right) + \frac{\Delta H^2}{8\Delta \mathbf{M}^4} \left(V(\mathbf{M}_1) + V(\mathbf{M}_2)\right)}$$

where 
$$\Delta H := 2\mathbf{M}_1 - \mathbf{M}_2$$
,  $\Delta \mathbf{M} := \mathbf{M}_2 - \mathbf{M}_1$ 

Return type np.array

 $\mathtt{get\_dg}() \rightarrow \mathrm{numpy.array}$ 

Get free energy of dye binding,  $\Delta g$ , vectorized version of Equation (25).

Returns

Return type np.array

 $\texttt{get\_dg\_std}\,(\,) \, \to numpy.array$ 

Get estimate of standard deviation in  $\Delta g$ .

**Returns** 

$$\frac{RT_j}{2\Delta\mathbf{M}_j\left(2\mathbf{M}_{j1}-\mathbf{M}_{j2}\right)}\sqrt{\mathbf{M}_{j2}^2V(\mathbf{M}_{j1})+\mathbf{M}_{j1}^2V(\mathbf{M}_{j2})}$$

where 
$$\Delta \mathbf{M}_{j} = \mathbf{M}_{j2} - \mathbf{M}_{j1}$$
.

Return type np.array

 $\textbf{get\_dh} \ (\ ) \ \rightarrow numpy.array$ 

Get differential enthalpy of binding,  $\Delta h$  as the vectorized version of Equation (26).

Returns

Return type np.array

 $\mathtt{get\_dh\_std}() \rightarrow \mathtt{numpy.array}$ 

Get estimate of error in  $\Delta h_j$ 

Returns

Return type np.array

 $\mathtt{get}_{\mathtt{f}}() \rightarrow \mathrm{numpy.array}$ 

Get f from vectorized version of Equation (23)

Returns f

Return type np.array

 $\mathtt{get\_f\_std}() \rightarrow \mathtt{numpy.array}$ 

Get standard deviation estimate of f

Returns

$$\sqrt{V(\mathbf{M}_1) + V(\mathbf{M}_2) + \frac{E_B^2 V_A + E_A^2 V_B}{E_B^4}}$$

where

$$\begin{split} V_A := (V_B + H_+^2)(V(\mathbf{M}_1) + V(\mathbf{M}_2) + \Delta \mathbf{M}^2) - E_A^2 \\ V_B := 2V(\mathbf{M}_1) + V(\mathbf{M}_2) \\ E_A := \Delta \mathbf{M} H_+ \\ E_B := 2\mathbf{M}_1 - \mathbf{M}_2 \\ \Delta \mathbf{M} := \mathbf{M}_2 - \mathbf{M}_1 \\ H_+ := 2\mathbf{M}_1 + \mathbf{M}_2 \end{split}$$

### Return type np.array

```
get_theta_b_all_1(n)
```

Returns  $\theta_{b,j,1}$  in Equation (S6a) pointwise in j The index b is the total number of wells.

### Returns

### Return type np.array

 $src.parameter\_extraction.calculate\_relative\_brightness (f\_SS, f\_DS)$  Calculate relative brightness, Equation (24).

### **Parameters**

- **f\_SS** (np.array) Molar fluorescence of single-stranded DNA, **f**<sup>SS</sup>.
- $f_DS(np.array)$  Molar fluorescence of double-stranded DNA,  $f^{DS}$ .

**Returns** Relative brightness,  $\mathbf{f}_{i}^{\mathrm{DS}}/\mathbf{f}_{i}^{\mathrm{SS}}$  for each j associated with SS.

### Return type np.array

src.parameter\_extraction.calculate\_relative\_brightness\_err (SS\_M1, 
$$SS_M2$$
,  $DS_M1$ ,  $DS_M2$ ,  $SS_V_M1$ ,  $SS_V_M2$ ,  $SS_V_M1$ ,  $SS_V_M2$ ,  $DS_V_M1$ ,  $DS_V_M2$ )  $\rightarrow$  numpy.array

Estimate error in relative brightness, Equation (24).

### **Parameters**

- SS\_M1  $(np.array) \mathbf{M}_1^{SS}$
- SS\_M2  $(np.array) M_2^{SS}$
- DS\_M1  $(np.array) \mathbf{M}_1^{DS}$
- DS\_M2  $(np.array) \mathbf{M}_2^{\mathrm{DS}}$
- SS\_V\_M1  $(np.array) V(\mathbf{M}_1^{SS})$  \_description\_
- $SS_V_M2$  (np.array)  $V(M_2^{SS})$
- DS\_V\_M1  $(np.array) V(\mathbf{M}_1^{DS})$
- DS\_V\_M2  $(np.array) V(\mathbf{M}_2^{DS})$

**Returns** Estimate of error in relative brightness

Return type np.array

# 1.5 Plotting

### 1.5.1 Raw Data

text)

```
SS B 1:
src.plot_raw_data.make_figure_2 (SS_A_1:
                                                              src.get_data.RawData,
                                           src.get data.RawData,
                                                                  SS C 1:
                                                                              src.get data.RawData,
                                           SS A 2:
                                                               src.get_data.RawData,
                                                                                           SS B 2:
                                           src.get_data.RawData,
                                                                DS_A_1:
                                                                              src.get_data.RawData,
                                           DS_B_1:
                                                               src.get_data.RawData,
                                                                                           DS_A_2:
                                          src.get_data.RawData,
                                                                 DS_B_2:
                                                                              src.get_data.RawData,
                                          A_1: src.get_data.RawData)
     Makes figure 2
          Parameters
                • SS_A_1 (RawData) – Data associated with (t, \ell, d) = (SS, A, 1) (see Table 1 of main
                • SS B 1 (RawData) – Data associated with (t, \ell, d) = (SS, B, 1) (see Table 1 of main
```

- SS\_C\_1 (RawData) Data associated with  $(t,\ell,d)=(SS,C,1)$  (see Table 1 of main text)
- SS\_A\_2 (RawData) Data associated with  $(t,\ell,d)=(SS,A,2)$  (see Table 1 of main text)
- SS\_B\_2 (RawData) Data associated with  $(t,\ell,d)=(SS,B,2)$  (see Table 1 of main text)
- DS\_A\_1 (RawData) Data associated with  $(t,\ell,d)=(DS,A,1)$  (see Table 1 of main text)
- DS\_B\_1 (RawData) Data associated with  $(t,\ell,d)=(DS,B,1)$  (see Table 1 of main text)
- DS\_A\_2 (RawData) Data associated with  $(t,\ell,d)=(DS,A,2)$  (see Table 1 of main text)
- DS\_B\_2 (RawData) Data associated with  $(t,\ell,d)=(DS,B,2)$  (see Table 1 of main text)
- **A\_1** (RawData) Data associated with  $(\ell, d) = (A, 1)$  (without DNA, see Table 1 of main text)

get\_ticks=False, ordered\_by\_row=True)

ax,

colorbar=False,

```
src.plot_raw_data.make_figure_S1()
    Makes Figure S1.
src.plot_raw_data.make_figure_S3()
    Makes Figure S3.
src.plot_raw_data.plot_linemap(cls: src.get_data.RawData,
```

Plot F vs C for various temperatures (colors)

## Parameters

- cls (Data) Instance of data (i.e., a dataset)
- ax (axis) Matplotlib axis to plot on

1.5. Plotting 9

- **colorbar** (bool, optional) whether or not to plot colorbar, in which case the axis is colorbar axis, by default False
- get\_ticks (bool, optional) whether or not to return list of ticks, by default False
- ordered\_by\_row (bool, optional) whether or not well concentrations are ordered by row, by default True

**Returns** Only returns list of get\_ticks=True.

Return type None or list

# 1.5.2 Quantifying Noise and Linearity

```
src.plot_noise_removal.plot_Chat_vs_C (Cs, C_hats, C_stds, fname)
Plot comparison between C and \widehat{C}. Used to plot figure 5
```

### **Parameters**

- **Cs** (tuple of np.arrays) each element of the tuple corresponds to **C** for a specific **D** and type.
- **C\_hats** (tuple of np.arrays) each element of the tuple corresponds to  $\widehat{\mathbf{C}}$  for a specific  $\mathbf{D}$  and type.
- **C\_stds** (tuple of np.arrays) each element of the tuple corresponds to the estimate for the standard deviation in  $\hat{\mathbf{C}}$  for a specific  $\mathbf{D}$  and type.
- **fname** (str) file name to save as figure (relative path).

### **Parameters**

- **Fs** (tuple of np.array) each element of the tuple corresponds to **F** for a specific **D** and type.
- **Fhats** (tuple of np.array) each element of the tuple corresponds to **F** for a specific **D** and type.
- **Ts** (tuple of np.array) each element of the tuple corresponds to the temperatures associated with a specific **D** and type.
- **fname** (str) name of figure to plot (relative path)
- **sname** (str) name of symbol in latex

```
src.plot_noise_removal.plot_error_C (ax, C, Chat, Chat_std=None, **kwargs)
Plot Errors in Dye concentration
```

### **Parameters**

- **ax** (matplotlib.pyplot.axis) plots on this axis
- C (np.array) Nominal dye concentrations determined experimentally (dimensionless)
- Chat (np.array) Dye concentrations determined from total least squares
- **Chat\_std** (np.array) Estimate of standard deviations in dye concentrations determined from total least squares
- **kwargs** Used to determine color if not provided. Also used when calling ax.plot.

```
src.plot_noise_removal.plot_error_F (ax, F_k_tl, F_hat_k_tl, T)

Plot the predicted fluorescence \hat{\mathbf{F}}_{ijk}^{t\ell} against \mathbf{F}_{ijk}^{t\ell} for i=1,\ldots,n and j=1,\ldots,n.
```

#### **Parameters**

- ax (matplotlib.pyplot.axes) axes to plot on
- **F\_kl** (np.ndarray) Experimentally measured fluorescence  $\mathbf{F}_k^{t\ell}$ .
- **F\_hat\_kl** (np.ndarray) Predicted fluorescence  $\widehat{\mathbf{F}}_k^{t\ell}$ .
- **T** (np.array) Temperatures in K

Returns image for colorbar

### Return type img

src.plot\_noise\_removal.plot\_figure7 (M\_mean, M\_std, T, color\_D1='C0', color\_D2='C1')
Plot figure 7

### **Parameters**

- M\_mean (tuple of np.array) each element of the tuple corresponds to M<sup>TLS</sup> for a specific D and type.
- **M\_std** ( $tuple\ of\ np.array$ ) each element of the tuple corresponds to standard deviations in  $\mathbf{M}^{TLS}$  for a specific  $\mathbf{D}$  and type.
- **T** (tuple of np.array) each element of the tuple corresponds to the temperatures associated with a specific **D** and type.
- color\_D1 (str) name of color for data associated with D=1
- $color_D2$  name of color for data associated with D=2

### 1.5.3 Parameters

 $\label{eq:src.plot_params.plot_figure8} $$ (T_SS, f_SS, K_SS, f_std_SS, K_std_SS, T_DS, f_DS, K_DS, f_std_DS, K_std_DS)$$ Plot figure 8$ 

### Parameters

- **T\_SS** (np.array) temperatures for single-stranded DNA
- **f SS** (np.array) molar fluorescence for single-stranded DNA
- K\_SS (np.array) partition coefficients for single-stranded DNA
- **f\_std\_ss** (np.array) standard deviation in molar fluorescence of single-stranded DNA
- K\_std\_SS (np.array) standard deviation in partition coefficients for single-stranded DNA
- **T\_DS** (np.array) temperatures for double-stranded DNA
- **f\_DS** (np.array) molar fluorescence for double-stranded dna
- **K\_DS** (np.array) partition coefficients for single-stranded DNA
- f\_std\_DS (np.array) standard deviation in molar fluorescence of double-stranded DNA
- K\_std\_DS (np.array) standard deviation in partition coefficients for double-stranded DNA

1.5. Plotting

#### **Parameters**

- SS (Parameters) object containing all single-stranded DNA input and methods to perform calculations
- **DS** (Parameters) object containing all single-stranded DNA input and methods to perform calculations

```
src.plot\_params.\textbf{plot\_figure\_S2} (SS\_1: src.get\_data.CombinedData, SS\_2: src.get\_data.CombinedData, DS\_1: src.get\_data.CombinedData, DS\_2: src.get\_data.CombinedData)
```

Plot figure S2

### **Parameters**

- SS\_1 (CombinedData) object containing all replicate plates for single-stranded DNA with D = 1
- SS\_2 (CombinedData) object containing all replicate plates for single-stranded DNA with  ${\bf D}=2$
- DS\_1 (CombinedData) object containing all replicate plates for double-stranded DNA with  ${\bf D}=1$
- DS\_2 (CombinedData) object containing all replicate plates for double-stranded DNA with  ${\bf D}=2$

```
 src.plot\_params. \textbf{plot\_figure\_S4} (SS\_data: src.parameter\_extraction. Parameters, \\ src.parameter\_extraction. Parameters) \\ DS\_data: src.parameter\_extraction. Parameters)
```

Plot Figure S4

### **Parameters**

- SS\_data (Parameters) parameters class that has method to calculate  $\theta_{b,j,1}$  pointwise in j
- DS\_data (Parameters) parameters class that has method to calculate  $\theta_{b,j,1}$  pointwise in j

```
src.plot_params.plot_figure_S5 (T, rb, d_rb)
Plot figure S5, relative brightness
```

### **Parameters**

- **T** (np.array) array of temperatures
- **rb** (np.array) Relative brightness,  $\mathbf{f}_j^{\mathrm{DS}}/\mathbf{f}_j^{\mathrm{SS}}$
- **d\_rb** (np.array) standard deviation in relative brightness

# **1.6 Util**

```
\mbox{src.util.figure\_name\_to\_abspath} \ (\mbox{\it fname: str}) \ \to \mbox{\it str} Figure name to absolute path
```

**Parameters** fname (str) – name of figure

**Returns** absolute path to name of figure

Return type str

1.6. Util 13

# **CHAPTER**

# TWO

# **INDICES AND TABLES**

- genindex
- modindex
- search

# **PYTHON MODULE INDEX**

### S

```
src.get_data,3
src.noise_removal,5
src.parameter_extraction,6
src.plot_noise_removal,10
src.plot_params,11
src.plot_raw_data,9
src.util,12
src.wells,1
```