Coarse-graining Markov state models with PCCA

Coarse-graining Markov state models

- *Coarse-graining Markov state models* here means finding a smaller transition matrix that does a similar job as the large original transition matrix.

- We have already seen one way of reducing the dimension of a transition matrix. Let’s take this as our starting point…

The truncated eigendecomposition

- The eigendecomposition of $\mathbf{P}(\tau)$ reads
$$\mathbf{P}(\tau) = \mathbf{R}\boldsymbol{\Lambda}(\tau)\mathbf{L}$$
- We have seen that for sufficiently large lag times τ, the majority of eigenvalues become almost zero.
- We can therefore truncate the matrix $\boldsymbol{\Lambda}(\tau)$.

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0.99 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Delete this and call the reduced matrix $\widetilde{\boldsymbol{\Lambda}}$. We can also ignore the corresponding eigenvectors in $\mathbf{R}$, $\mathbf{L}$ and call the reduced matrix $\widetilde{\mathbf{R}}$, $\tilde{\mathbf{L}}$.

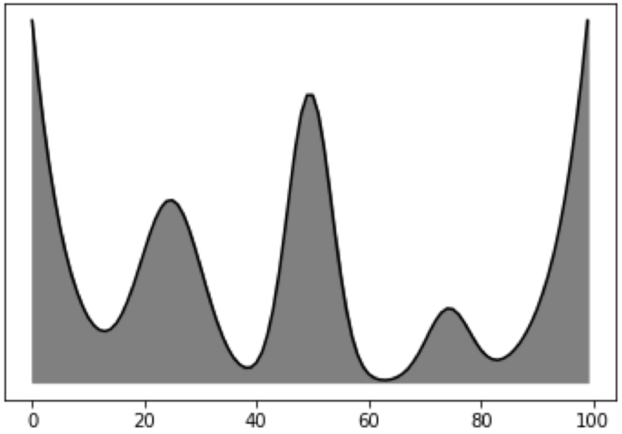
The truncated eigendecomposition

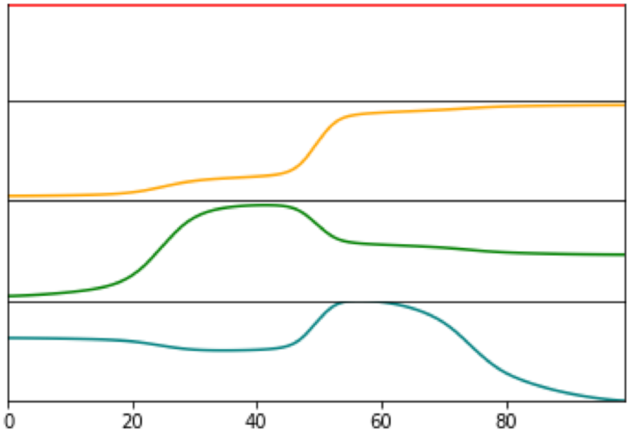
- We now have $\mathbf{P}(\tau) \approx \widetilde{\mathbf{R}}\widetilde{\mathbf{\Lambda}}(\tau)\tilde{\mathbf{L}}$.
- And also $\mathbf{P}(\tau)^N \approx \widetilde{\mathbf{R}}\widetilde{\mathbf{\Lambda}}^N(\tau)\tilde{\mathbf{L}}$ since $\tilde{\mathbf{L}}\widetilde{\mathbf{R}} = \mathrm{Id}$.

- So did we find what we wanted?
  - $\widetilde{\mathbf{\Lambda}}(\tau)$ replaces $\mathbf{P}$ for large τ ✓
  - $\widetilde{\mathbf{\Lambda}}(\tau)$ is a small matrix ✓
  - But $\widetilde{\mathbf{\Lambda}}(\tau)$ is not a transition matrix.
    e.g. $\widetilde{\mathbf{\Lambda}}\mathbf{1} \neq \mathbf{1}$

- Can we correct the last point?

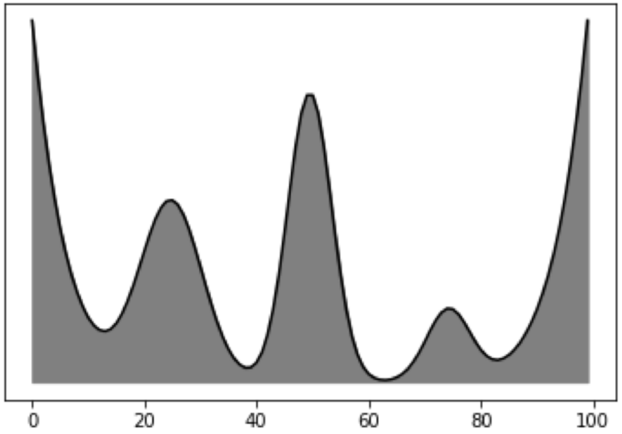
A closer look at the eigenvectors

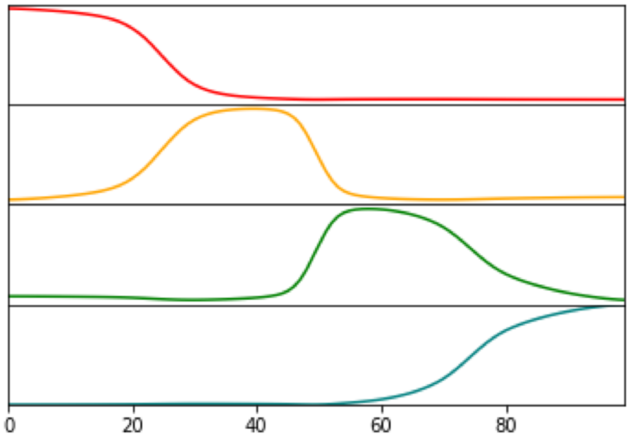




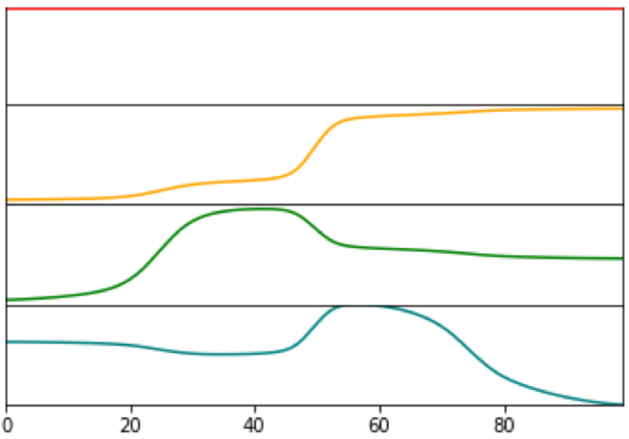
A closer look at the eigenvectors

- The dominant eigenvectors can be linearly transformed into a indicator vectors for the metastable states.
- These indicators are called **memberships**.





$$= \begin{pmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{11} & a_{22} & a_{33} & a_{42} \\ a_{13} & a_{23} & a_{33} & a_{43} \\ a_{14} & a_{24} & a_{33} & a_{44} \end{pmatrix}$$



$$\boldsymbol{\chi}^{\mathrm{T}} \quad = \quad \mathbf{A}^{\mathrm{T}} \quad \widetilde{\mathbf{R}}^{\mathrm{T}}$$

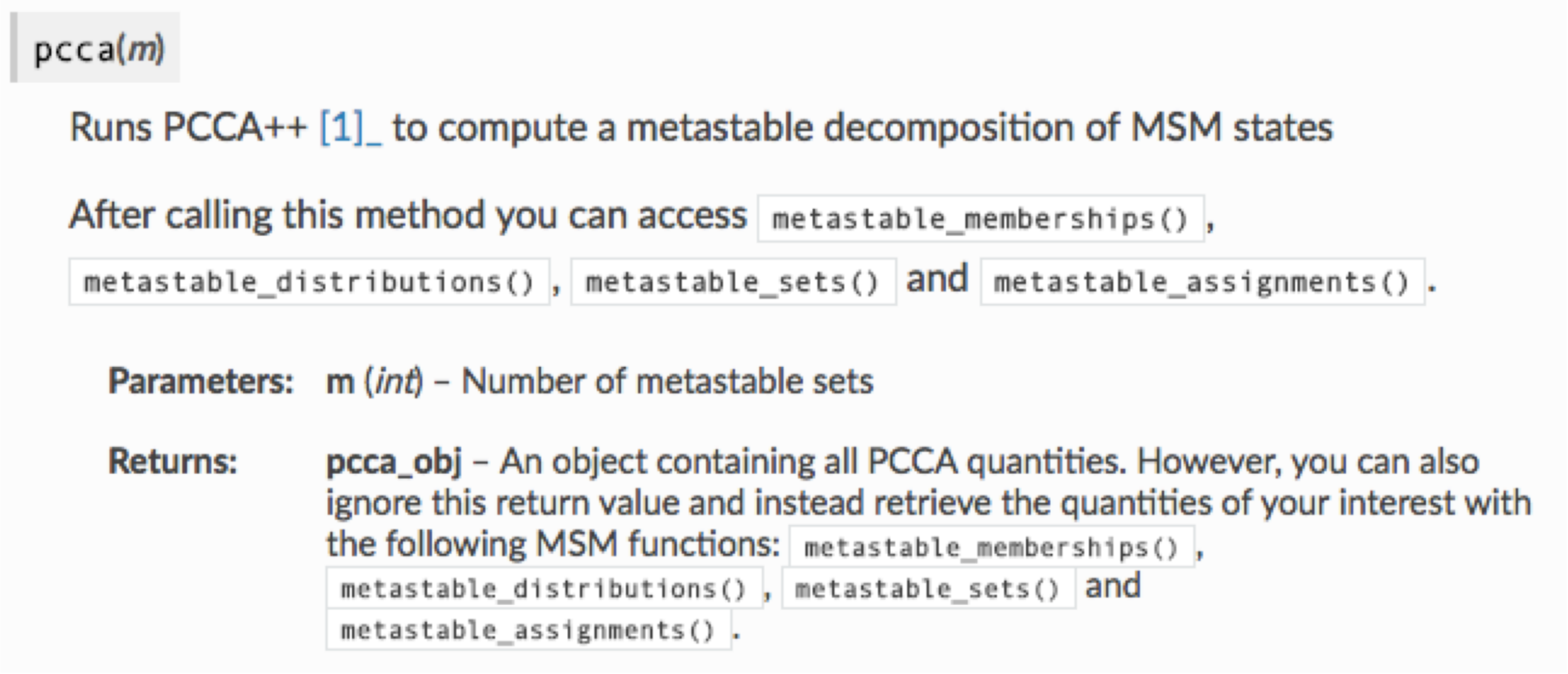
Coarse-graining with PCCA

- Use eigendecomposition and insert $\mathbf{A}\mathbf{A}^{-1}$:

$$\mathbf{P} = \widetilde{\mathbf{R}}\boldsymbol{\Lambda}(\tau)\tilde{\mathbf{L}} = \widetilde{\mathbf{R}}\mathbf{A}\;\underbrace{\mathbf{A}^{-1}\boldsymbol{\Lambda}(\tau)\mathbf{A}}_{\mathbf{P}_\mathrm{C}}\;\mathbf{A}^{-1}\tilde{\mathbf{L}}$$

- We have $\mathbf{P}^{N} = \widetilde{\mathbf{R}}\mathbf{A}\mathbf{P}_\mathrm{C}^{N}\mathbf{A}^{-1}\tilde{\mathbf{L}}$
- Are we done now?
  - $\mathbf{P}_\mathrm{C}$ replaces $\mathbf{P}$ for large τ ✓ Same eigenvalue as $\mathbf{P}$ ✓
  - $\mathbf{P}_\mathrm{C}$ is a small matrix ✓
  - $\mathbf{P}_\mathrm{C}\mathbf{1} = \mathbf{1}$ (without proof) ✓
  - $\mathbf{P}_\mathrm{C}$ can be interpreted as the transition matrix between the metastable states. ✓
  - $\mathbf{P}_\mathrm{C}$ is a Koopman matrix. (without proof) ✓
  - $\mathbf{P}_\mathrm{C} \ngeq \mathbf{0}$

PCCA in PyEmma



pcca(*m*)

Runs PCCA++ [1]_ to compute a metastable decomposition of MSM states

After calling this method you can access `metastable_memberships()`, `metastable_distributions()`, `metastable_sets()` and `metastable_assignments()`.

Parameters: **m** (*int*) – Number of metastable sets

Returns: **pcca_obj** – An object containing all PCCA quantities. However, you can also ignore this return value and instead retrieve the quantities of your interest with the following MSM functions: `metastable_memberships()`, `metastable_distributions()`, `metastable_sets()` and `metastable_assignments()`.

- $\boldsymbol{\chi}$ … metastable memberships
- $\boldsymbol{\Pi\chi}$ … metastable distributions
- $\mathrm{argmax}_j\, \chi_{ij}$ … metastable assignments
- $M_i = \{s \mid \mathrm{argmax}_j\, \chi_{sj} = i\}$ … metastable sets $[M_i]_{i=1,\ldots,n}$

Further reading

- Susanna Röblitz, Marcus Weber, “Fuzzy spectral clustering by PCCA+: application to Markov state models and data classification”, *Advances in Data Analysis and Classification*, **7**, 147 (2013)
- Marcus Weber, Konstantin Fackeldey, "G-PCCA: Spectral Clustering for Non-reversible Markov Chains", *Konrad-Zuse-Zentrum für Informationstechnik Berlin, ZIB-Report* 15-35 (2015)

Appendix: Proof that $\mathbf{P}_{\mathrm{C}}\mathbf{1} = \mathbf{1}$

- Memberships must sum to one $\boldsymbol{\chi}\mathbf{1}_{n\times 1} = \mathbf{1}_{N\times 1}$
- The first right eigenvector is constant $\mathbf{R}\mathbf{e}_1 = \mathbf{1}_{N\times 1}$.
- $\Rightarrow \boldsymbol{\chi}\mathbf{1}_{n\times 1} = \mathbf{R}\mathbf{e}_1$
- Use definition of $\boldsymbol{\chi}$: $\boldsymbol{\chi}\mathbf{1}_{n\times 1} = \mathbf{R}\mathbf{A}\mathbf{1}_{n\times 1}$
- Therefore $\mathbf{R}\mathbf{e}_1 = \mathbf{R}\mathbf{A}\mathbf{1}_{n\times 1}$ which is satisfied by $\mathbf{A}\mathbf{1}_{n\times 1} = \mathbf{e}_1$.
- $\Rightarrow \mathbf{P}_{\mathrm{C}}\mathbf{1} = \mathbf{A}^{-1}\boldsymbol{\Lambda}(\tau)\mathbf{A}\mathbf{1} = \mathbf{A}^{-1}\boldsymbol{\Lambda}(\tau)\mathbf{e}_1 = \mathbf{A}^{-1}\mathbf{e}_1 = \mathbf{1}$

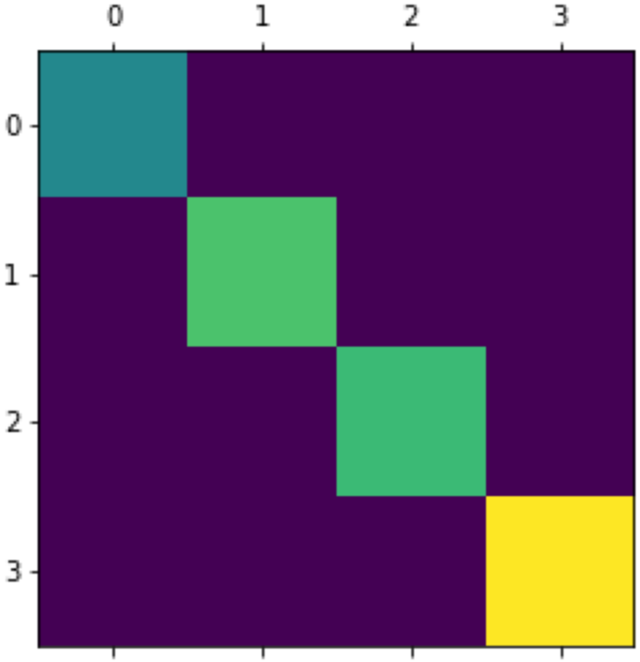
Appendix: Computing A

$$\mathbf{\Pi}_{\mathrm{C}} = \mathrm{diag}(\mathbf{A}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\boldsymbol{\pi})$$

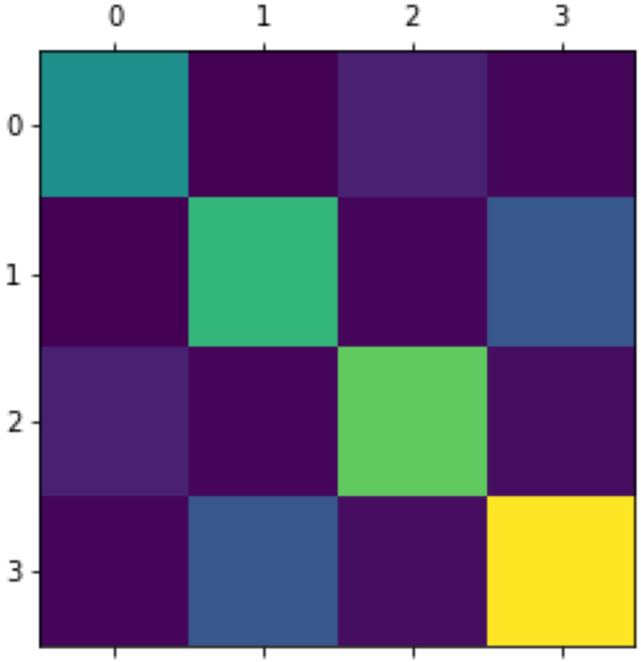
Stationary weight of the metastable states Inserted into the diagonal of a matrix.

$$\mathrm{Cov}(\boldsymbol{\chi}, \boldsymbol{\chi}) = \mathbf{A}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{\Pi}\mathbf{R}\mathbf{A}$$

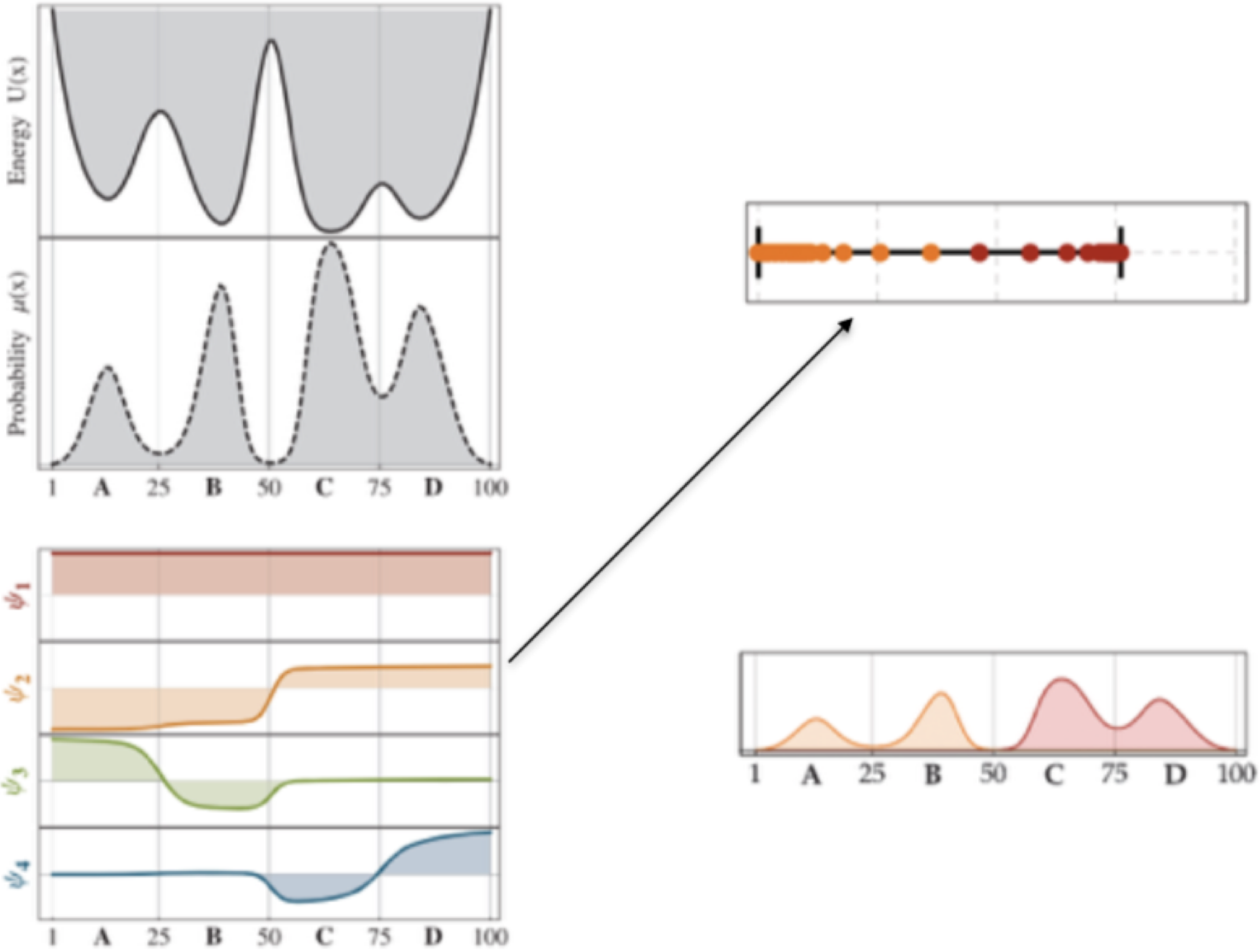
Overlap matrix of metastable states, weighted by stationary distribution

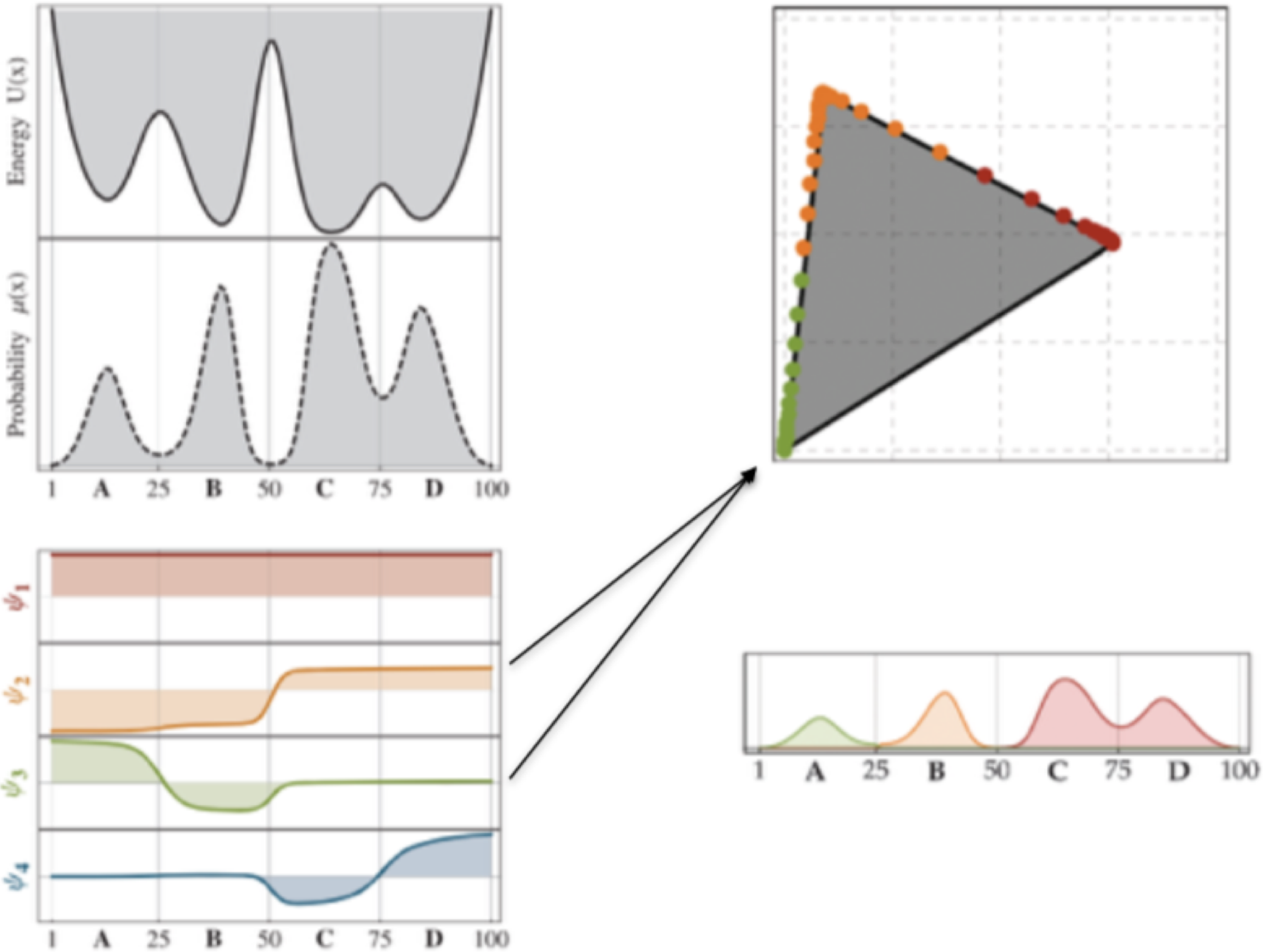


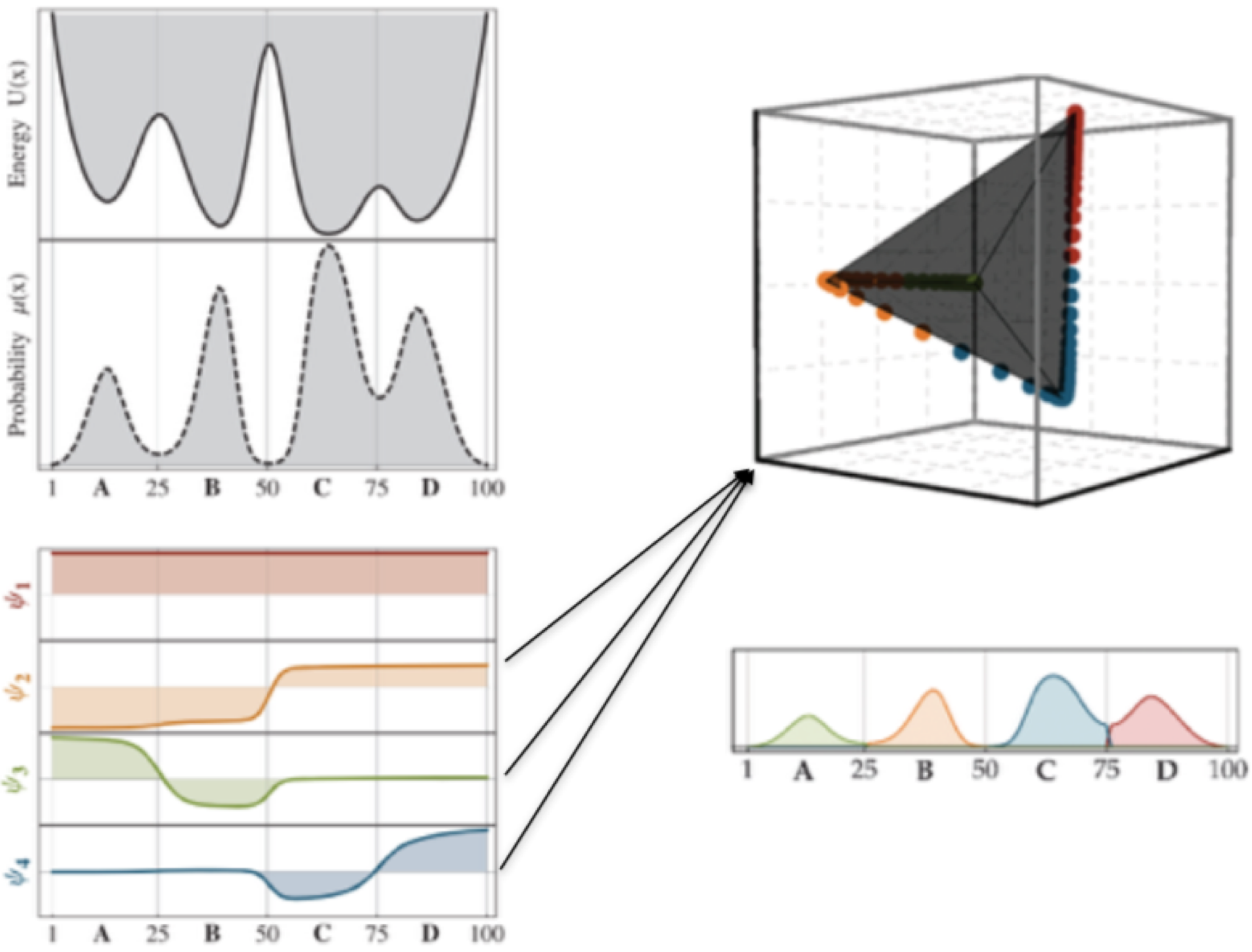
$$\approx$$



$$\mathrm{tr}(\mathbf{\Pi}_{\mathrm{C}}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{\Pi}\mathbf{R}\mathbf{A}) \rightarrow \text{min}$$


Energy U(x)
Probability μ(x)
1
A
25
B
50
C
75
D
100
ψ1
ψ2
ψ3
ψ4
1
A
25
B
50
C
75
D
100
1
A
25
B
50
C
75
D
100


Energy U(x)
Probability μ(x)
1
A
25
B
50
C
75
D
100
ψ_1
ψ_2
ψ_3
ψ_4
1
A
25
B
50
C
75
D
100
1
A
25
B
50
C
75
D
100


Energy U(x)
Probability μ(x)
1
A
25
B
50
C
75
D
100
ψ1
ψ2
ψ3
ψ4
1
A
25
B
50
C
75
D
100
1
A
25
B
50
C
75
D
100

$\mathbf{R} \in \mathbb{R}^{N \times n}$	matrix of dominant eigenvectors
$\boldsymbol{\chi} \in \mathbb{R}^{N \times n}$	matrix of memberships
$\boldsymbol{\chi} \geq 0$	non-negativity
$\sum_{i=1}^{n} \boldsymbol{\chi} = 1$	partition of 1
$\boldsymbol{\chi} \approx \mathbf{RA}$	spectral clustering