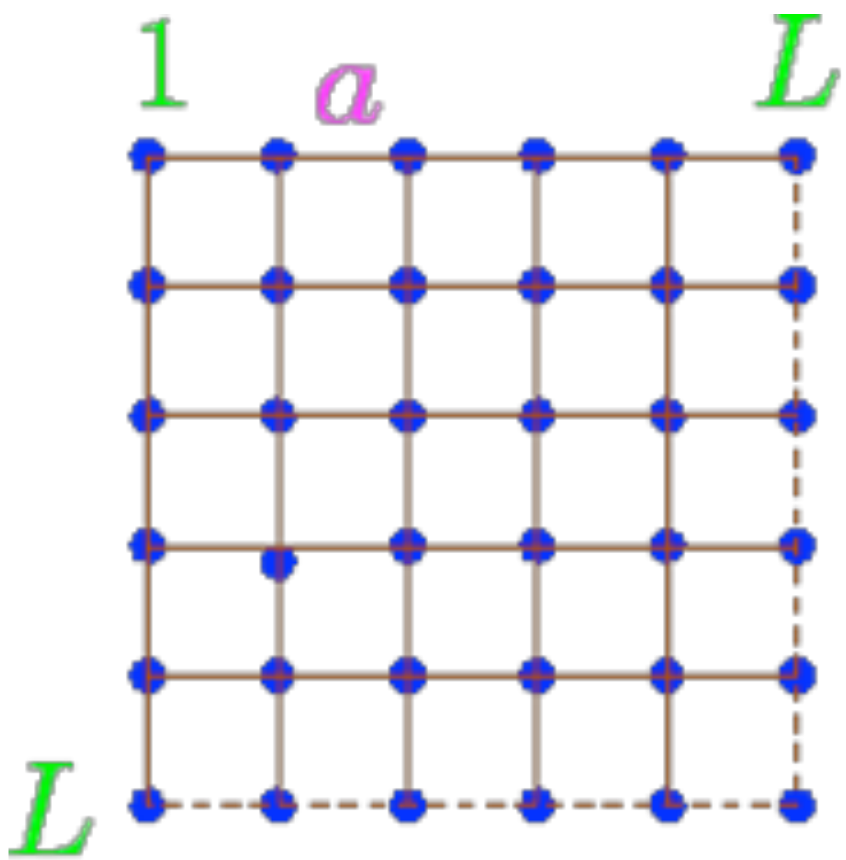
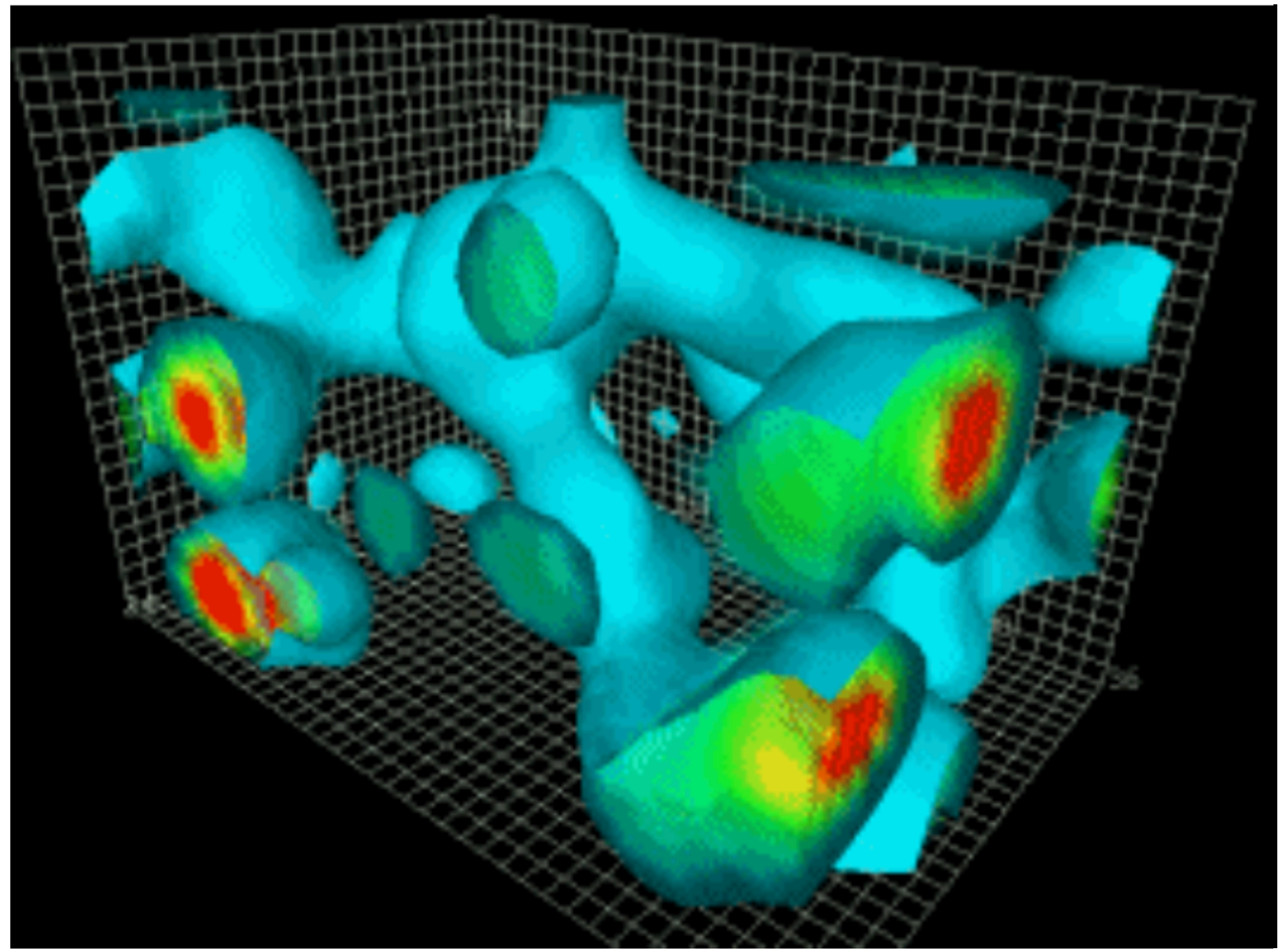
Extracting Excited States from Lattice Correlation Functions

Kimmy Cushman & George Fleming



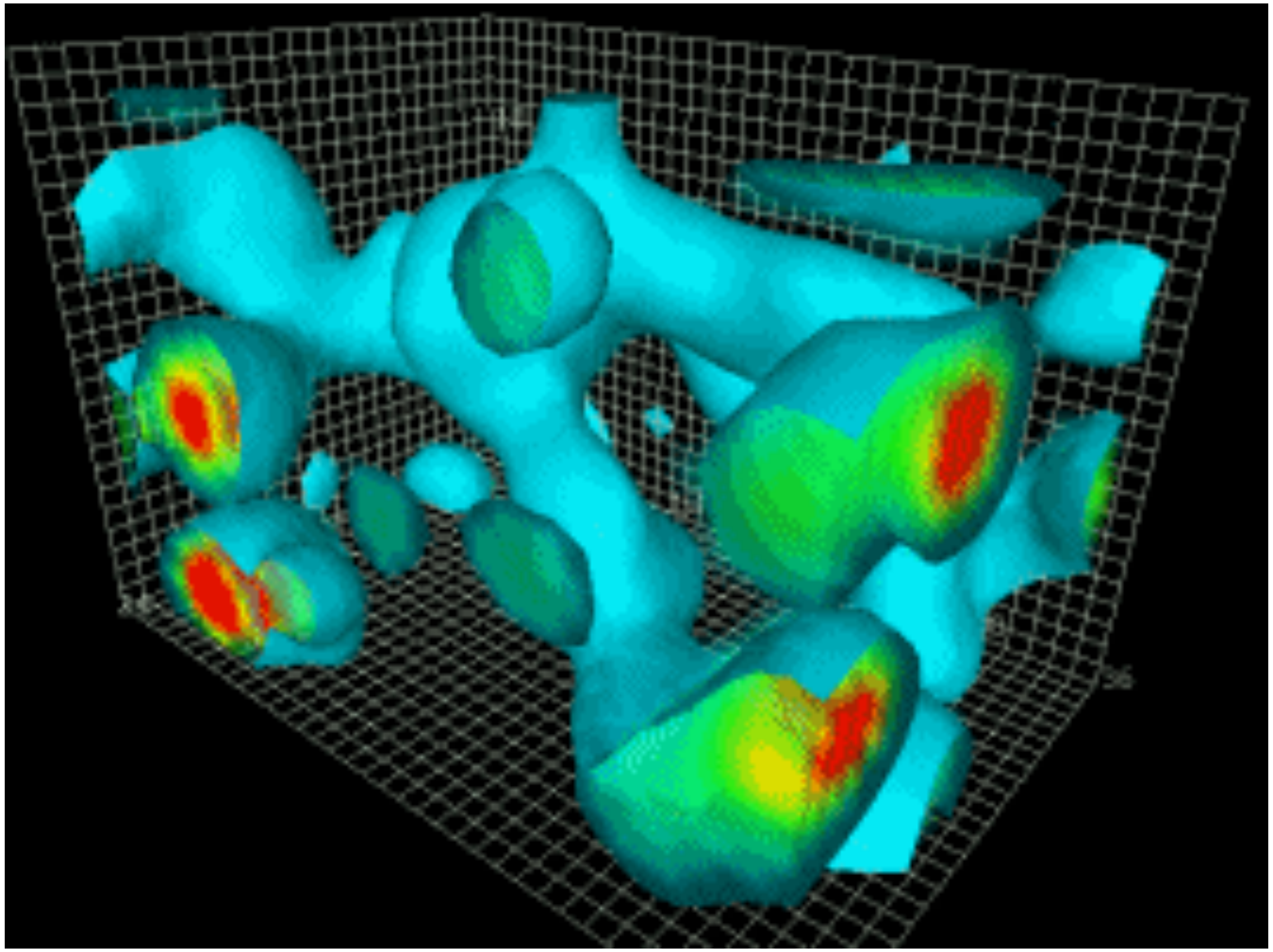


Outline

- Introduction and motivation
- Correlation functions
- Prony's method
- Bootstrapping and results
- Clustering
- Future work

Lattice QCD… what is it?

- Gauge field simulation from first principles
- Discretized space-time
- Finite volume with periodic conditions
- Volume ~ size of proton



http://www.physics.adelaide.edu.au/cssm/lattice/

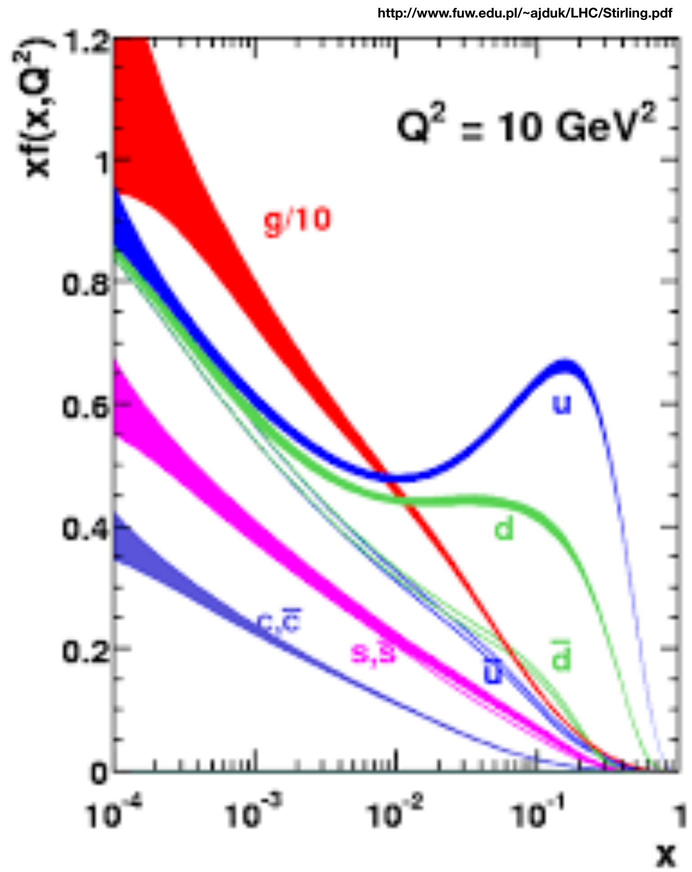
Motivation for QCD

- **QCD strongly coupled at <u>low energies</u>**

  -> not perturbative

  -jet fragmentation functions

  -parton distribution functions



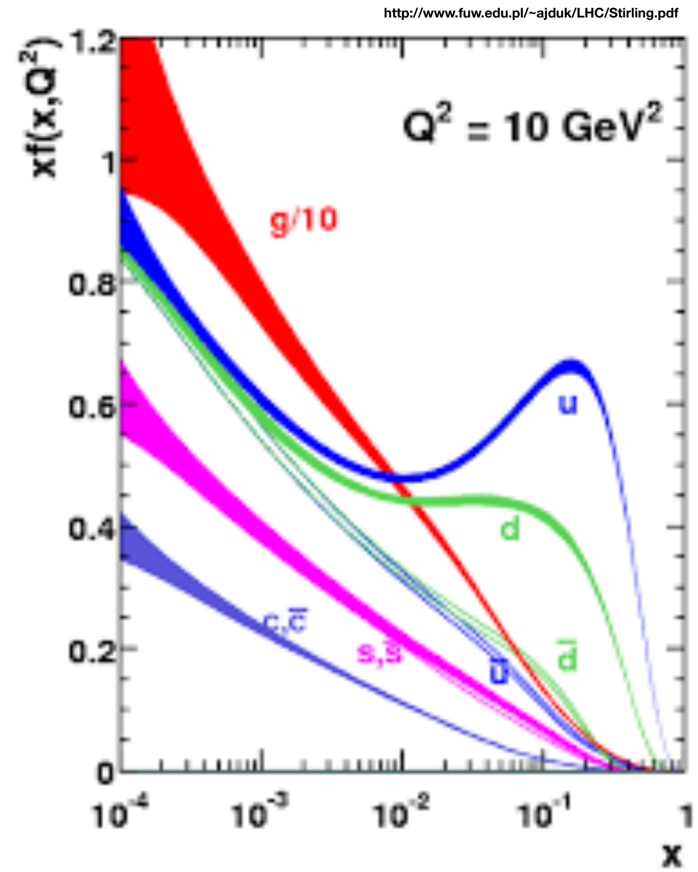
Motivation for QCD

- QCD strongly coupled at <u>low energies</u>

  -> not perturbative

  -jet fragmentation functions

  -parton distribution functions

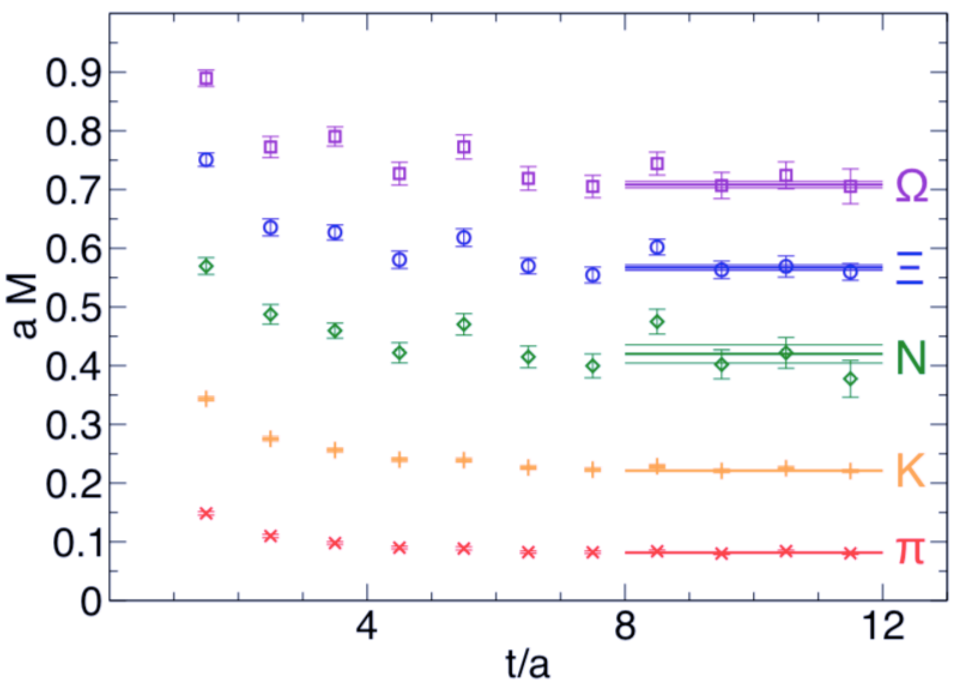


- Increased computational power = competitive and supplementary to experiment!

Image credit: Carlos Jones/ORNL

Motivation for QCD

1) > 95% of mass of hadrons come from QCD dynamics



Ab-initio Determination of Light Hadron Masses

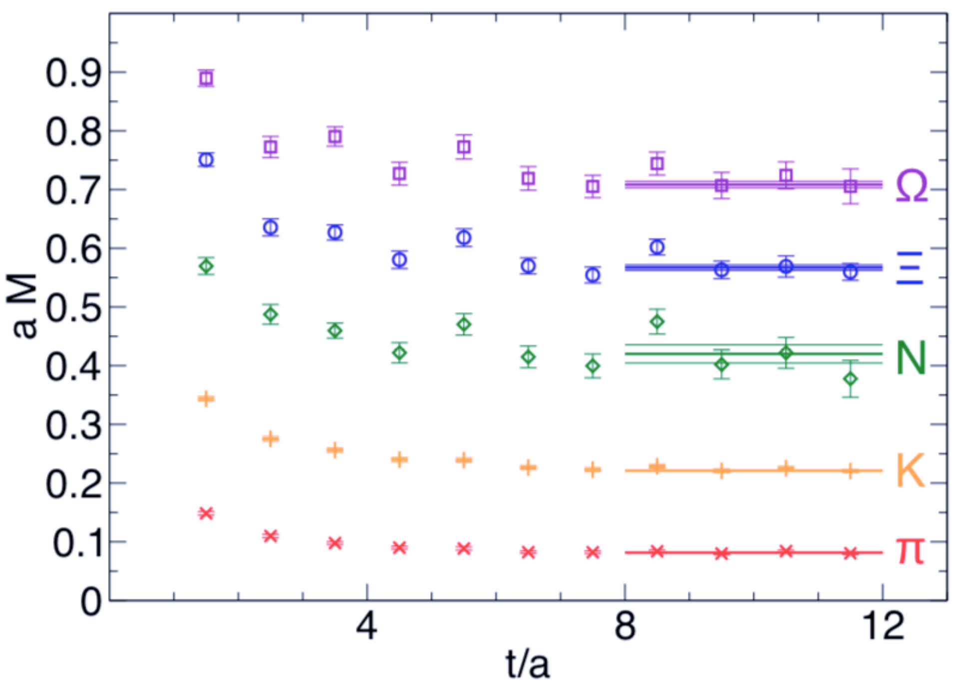
S. Dürr[1], Z. Fodor[1,2,3], J. Frison[4], C. Hoelbling[2,3,4],
R. Hoffmann[2], S. D. Katz[2,3], S. Krieg[2], T. Kurth[2],
L. Lellouch[4], T. Lippert[2,5], K.K. Szabo[2], G. Vulvert[4]

[1]NIC, DESY Zeuthen, D-15738 Zeuthen and FZ Jülich, D-52425 Jülich, Germany.
[2]Bergische Universität Wuppertal, Gaussstr. 20, D-42119 Wuppertal, Germany.
[3]Institute for Theoretical Physics, Eötvös University, H-1117 Budapest, Hungary.
[4]Centre de Physique Théorique*, Case 907, Campus de Luminy, F-13288 Marseille Cedex 9, France.
[5]Jülich Supercomputing Centre, FZ Jülich, D-52425 Jülich, Germany.

Budapest-Marseille-Wuppertal Collaboration

Motivation for QCD

1) > 95% of mass of hadrons come from QCD dynamics



Ab-initio Determination of Light Hadron Masses

S. Dürr[1], Z. Fodor[1,2,3], J. Frison[4], C. Hoelbling[2,3,4],
R. Hoffmann[2], S. D. Katz[2,3], S. Krieg[2], T. Kurth[2],
L. Lellouch[4], T. Lippert[2,5], K.K. Szabo[2], G. Vulvert[4]

[1]NIC, DESY Zeuthen, D-15738 Zeuthen and FZ Jülich, D-52425 Jülich, Germany.
[2]Bergische Universität Wuppertal, Gaussstr. 20, D-42119 Wuppertal, Germany.
[3]Institute for Theoretical Physics, Eötvös University, H-1117 Budapest, Hungary.
[4]Centre de Physique Théorique*, Case 907, Campus de Luminy, F-13288 Marseille Cedex 9, France.
[5]Jülich Supercomputing Centre, FZ Jülich, D-52425 Jülich, Germany.

Budapest-Marseille-Wuppertal Collaboration

2) Computation of matrix elements for weak flavor mixing

$$\Gamma(\pi \to l\nu) = \underbrace{\frac{G_F^2 |V_{ud}|^2 f_\pi^2}{8\pi}}_{} \underbrace{m_\pi m_l^2 \left(1 - \frac{m_l^2}{m_\pi^2}\right)^2}_{\text{Non-perturbative}}$$

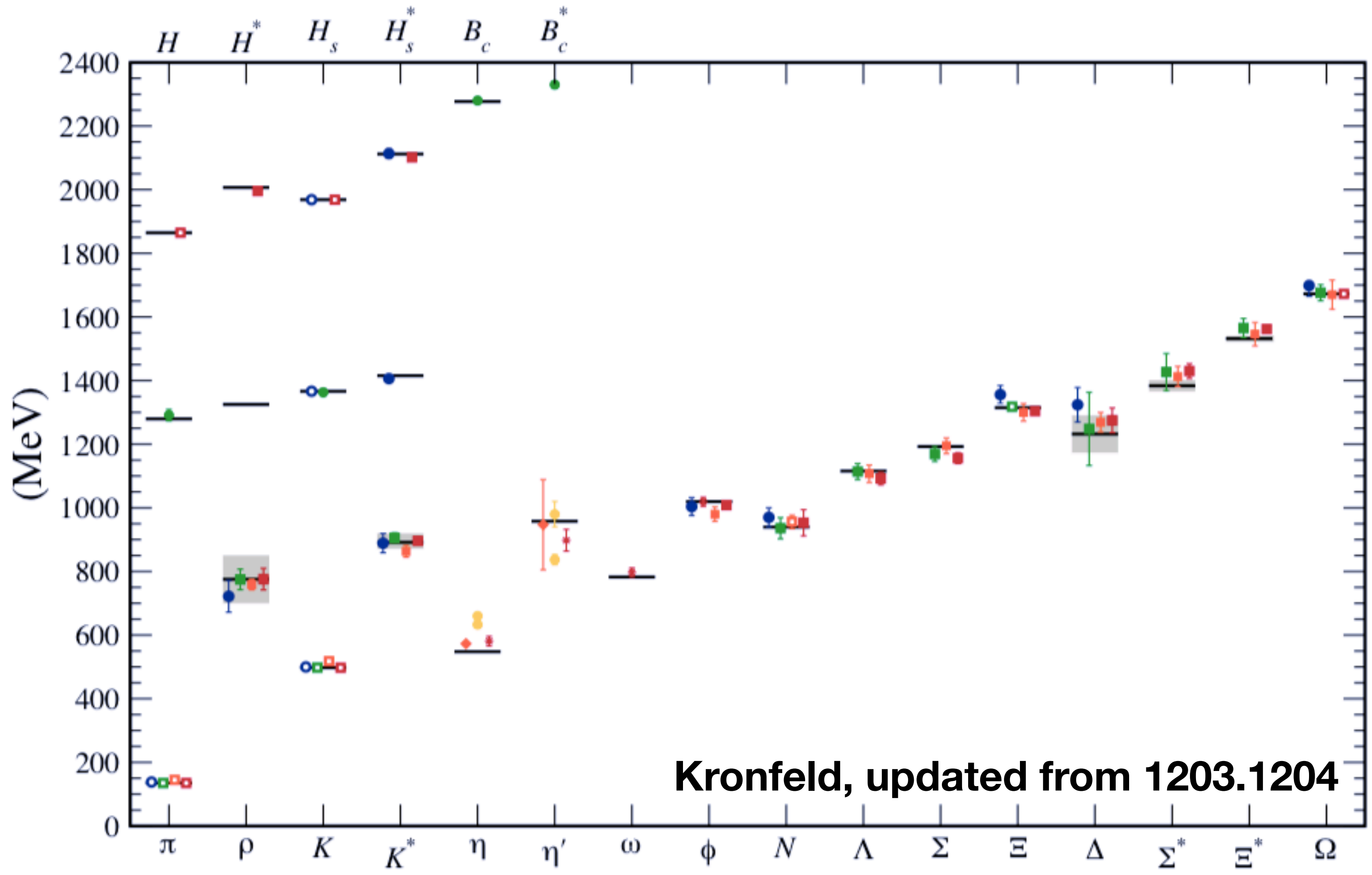
Perturbative

Non-perturbative

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

Motivation for Beyond the Standard Model

- **Predict spectrum of QCD baryons and mesons**



Motivation for Beyond the Standard Model

Motivation for Beyond the Standard Model

New gauge forces to explain…

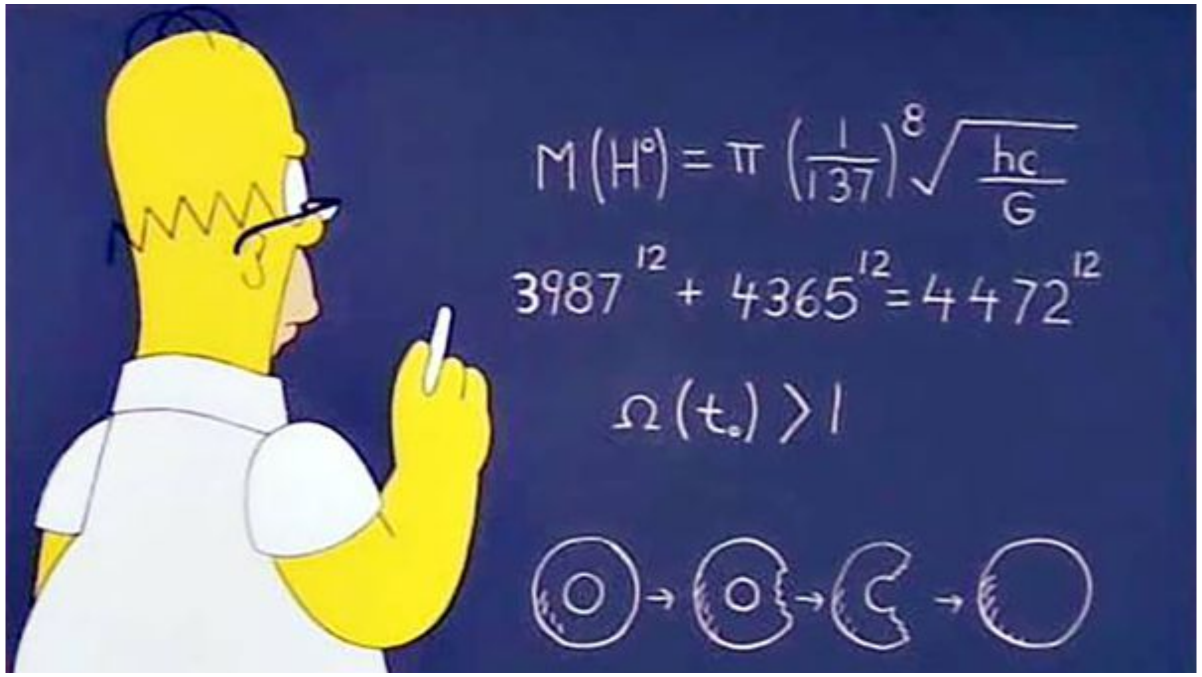
- **Higgs mechanism?**

https://www.symmetrymagazine.org/

- **Dark matter?**

https://www.nasa.gov/

Motivation for Beyond the Standard Model

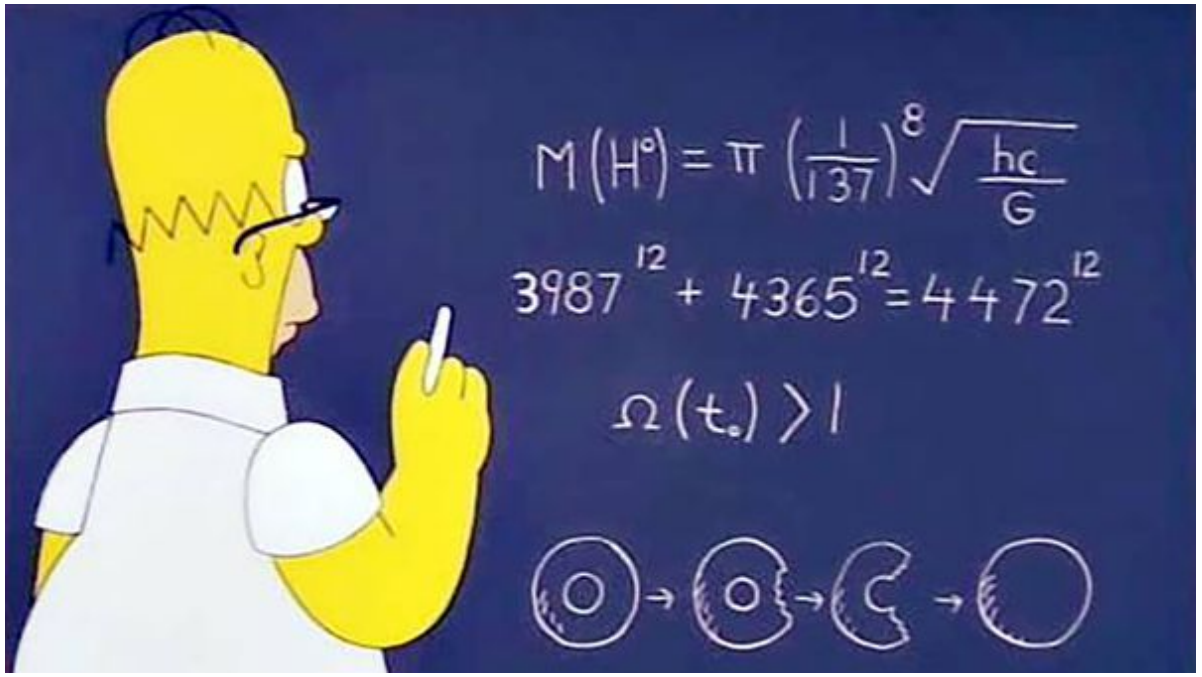


https://motls.blogspot.com/2015/03/did-homer-simpson-calculate-

Why composite Higgs?

1) Hierarchy problem - no longer a *fundamental* scalar

Motivation for Beyond the Standard Model

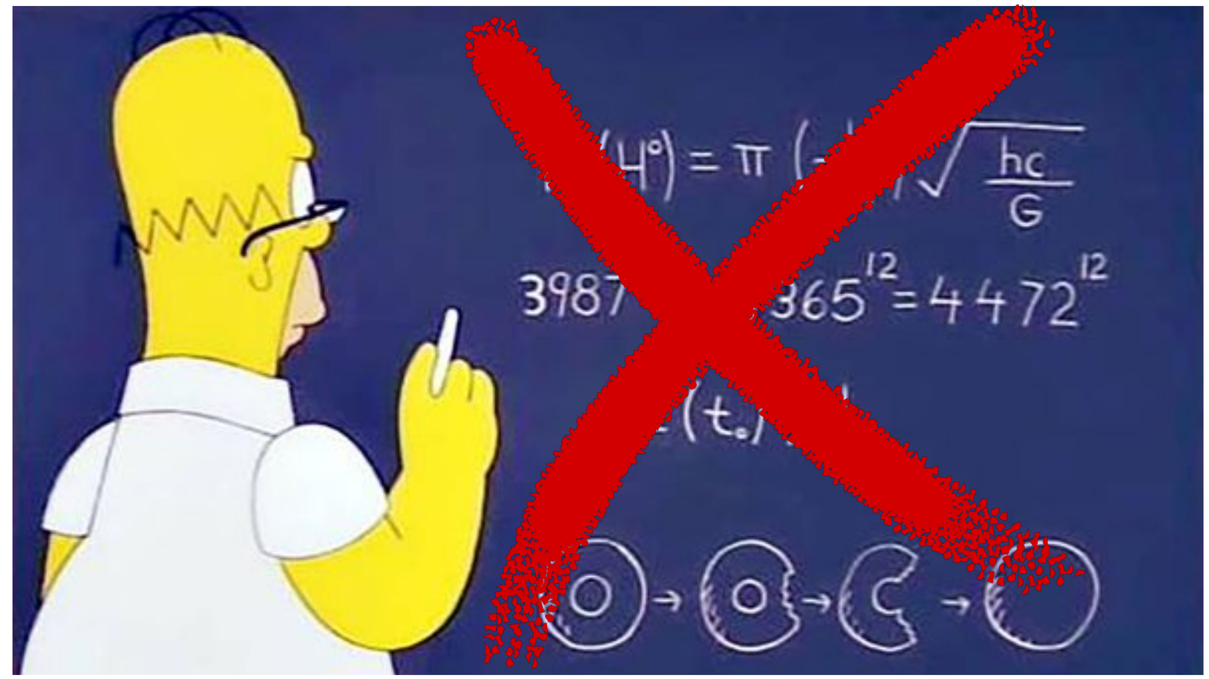


https://motls.blogspot.com/2015/03/did-homer-simpson-calculate-

Why composite Higgs?

1) **Hierarchy problem - no longer a *fundamental* scalar**
   - no fine tuning necessary if composite

Motivation for Beyond the Standard Model



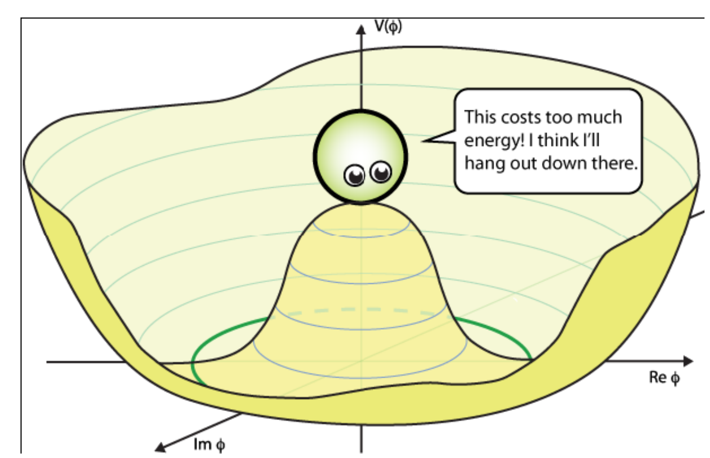
https://motls.blogspot.com/2015/03/did-homer-simpson-calculate-correct.html

Why composite Higgs?

1) **Hierarchy problem - no longer a *fundamental* scalar**
   - no fine tuning necessary if composite

2) **Dynamical symmetry breaking - Higgs model describes *effective* potential. Explains where potential comes from**

??
$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2$$
??

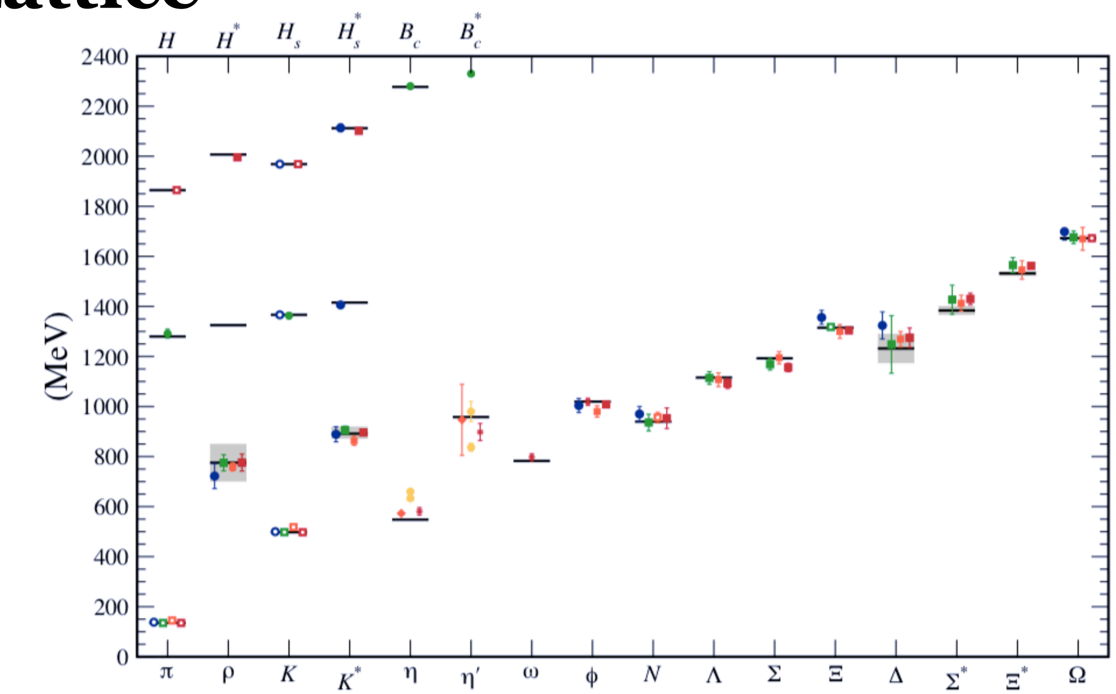


http://www.tifr.res.in/TSN/news_detail.php?id=69

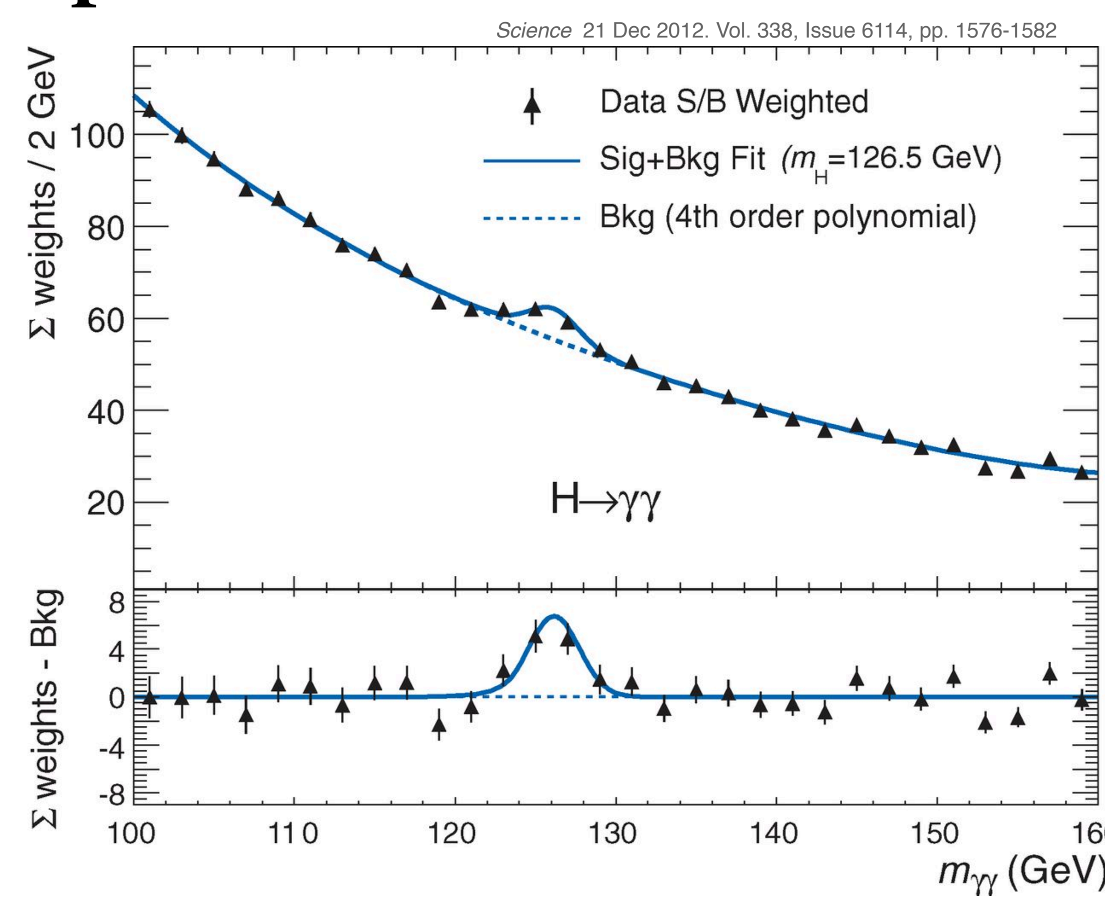
Motivation for Beyond the Standard Model

- **Every confining force has a spectrum of states**
- **New force of composite Higgs would have spectrum**

Lattice



Experiment



Feynman path integral trick

$$\langle 0|\bar{\pi}(t)\pi(0)|0\rangle = C_\pi(t) = \mathcal{O}$$

Feynman path integral trick

$$\langle 0|\bar{\pi}(t)\pi(0)|0\rangle = C_\pi(t) = \mathcal{O}$$

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathcal{D}} O[A]\, \mathrm{e}^{-iS[A]}}{\int_{\mathcal{D}} \mathrm{e}^{-iS[A]}}$$

Feynman path integral trick

$$\langle 0|\bar{\pi}(t)\pi(0)|0\rangle = C_\pi(t) = \mathcal{O}$$

$$\langle \mathcal{O} \rangle = \cancel{\frac{\int_\mathcal{D} \mathcal{O}[A]\, \mathrm{e}^{-iS[A]}}{\int_\mathcal{D} \mathrm{e}^{-iS[A]}}} \longrightarrow \langle \mathcal{O} \rangle = \frac{1}{Z}\sum_i \mathrm{e}^{-\beta S}\mathcal{O}_i$$

Feynman path integral trick

$$\langle 0|\bar{\pi}(t)\pi(0)|0\rangle = C_\pi(t) = \mathcal{O}$$

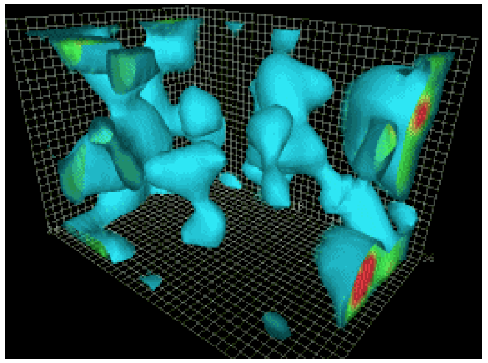
$$\cancel{\langle\mathcal{O}\rangle = \frac{\int_\mathcal{D} \mathcal{O}[A]\, \mathrm{e}^{-iS[A]}}{\int_\mathcal{D} \mathrm{e}^{-iS[A]}}} \longrightarrow \langle\mathcal{O}\rangle = \frac{1}{Z}\sum_i \mathrm{e}^{-\beta S}\mathcal{O}_i$$

Our only hope!

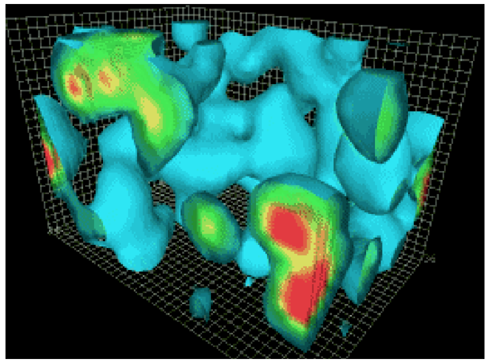
Importance Sampling

$$\langle 0|\bar{\pi}(t)\pi(0)|0\rangle = C_\pi(t) = \mathcal{O}$$

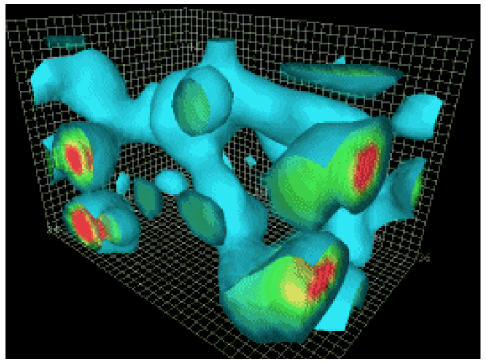
$$\langle \mathcal{O} \rangle = \cancel{\frac{\int_{\mathcal{D}} \mathcal{O}[A]\, \mathrm{e}^{-iS[A]}}{\int_{\mathcal{D}} \mathrm{e}^{-iS[A]}}} \longrightarrow \langle \mathcal{O} \rangle = \frac{1}{Z} \sum_i \mathrm{e}^{-\beta S} \mathcal{O}_i$$



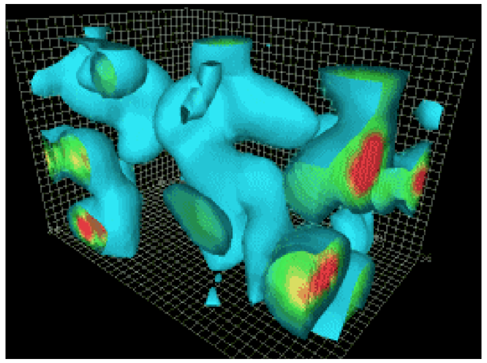
Configuration 1 $\mathcal{O}_1$



Configuration 2 $\mathcal{O}_2$



Configuration 3 $\mathcal{O}_3$



Configuration 4 $\mathcal{O}_4$

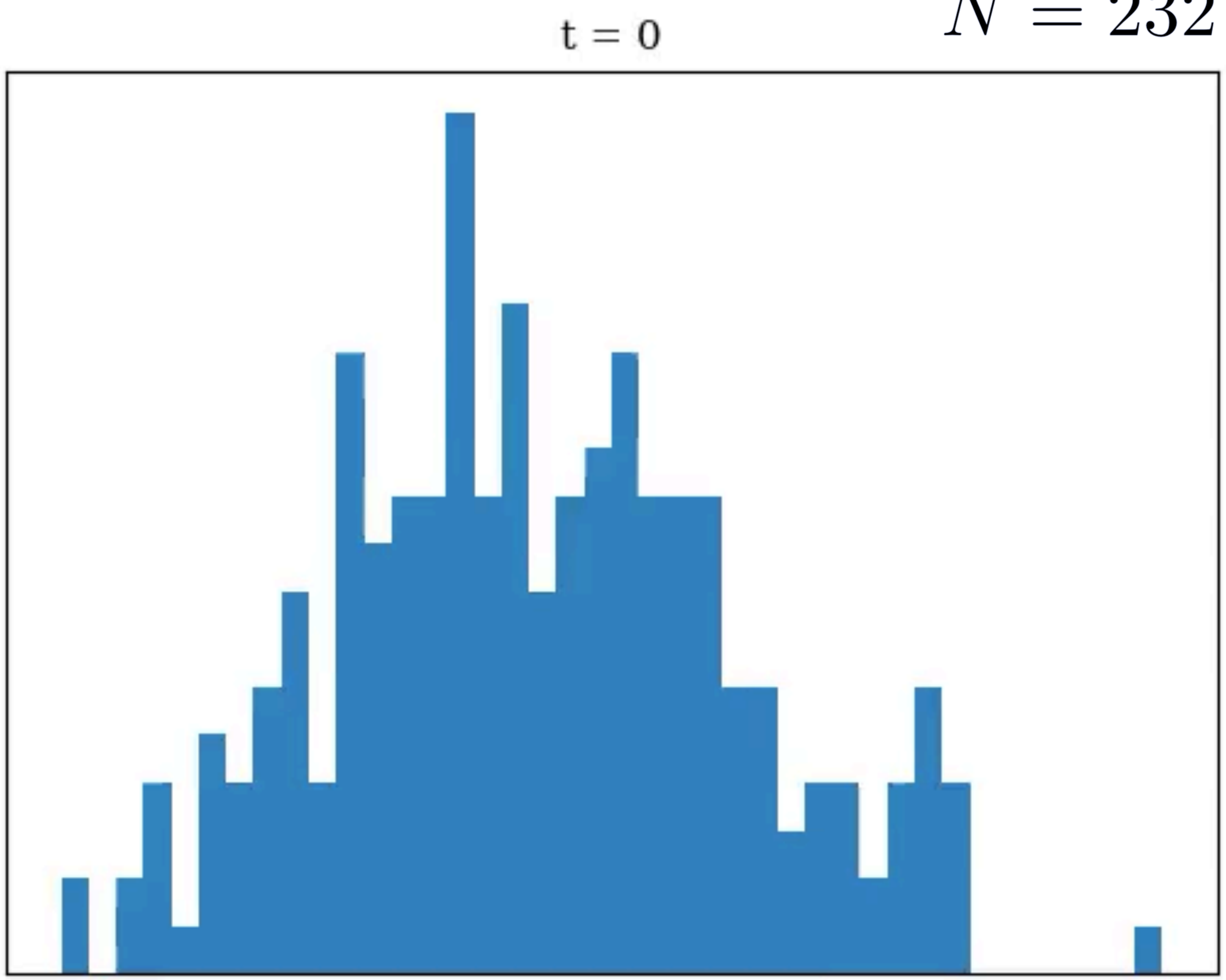
$$\langle \mathcal{O} \rangle = \frac{1}{N} \sum_i^N \mathcal{O}_i$$

Importance Sampling

$$\langle 0|\bar{\pi}(t)\pi(0)|0\rangle = C_\pi(t) = \mathcal{O} \qquad \langle \mathcal{O} \rangle = \frac{1}{N}\sum_i^N \mathcal{O}_i$$

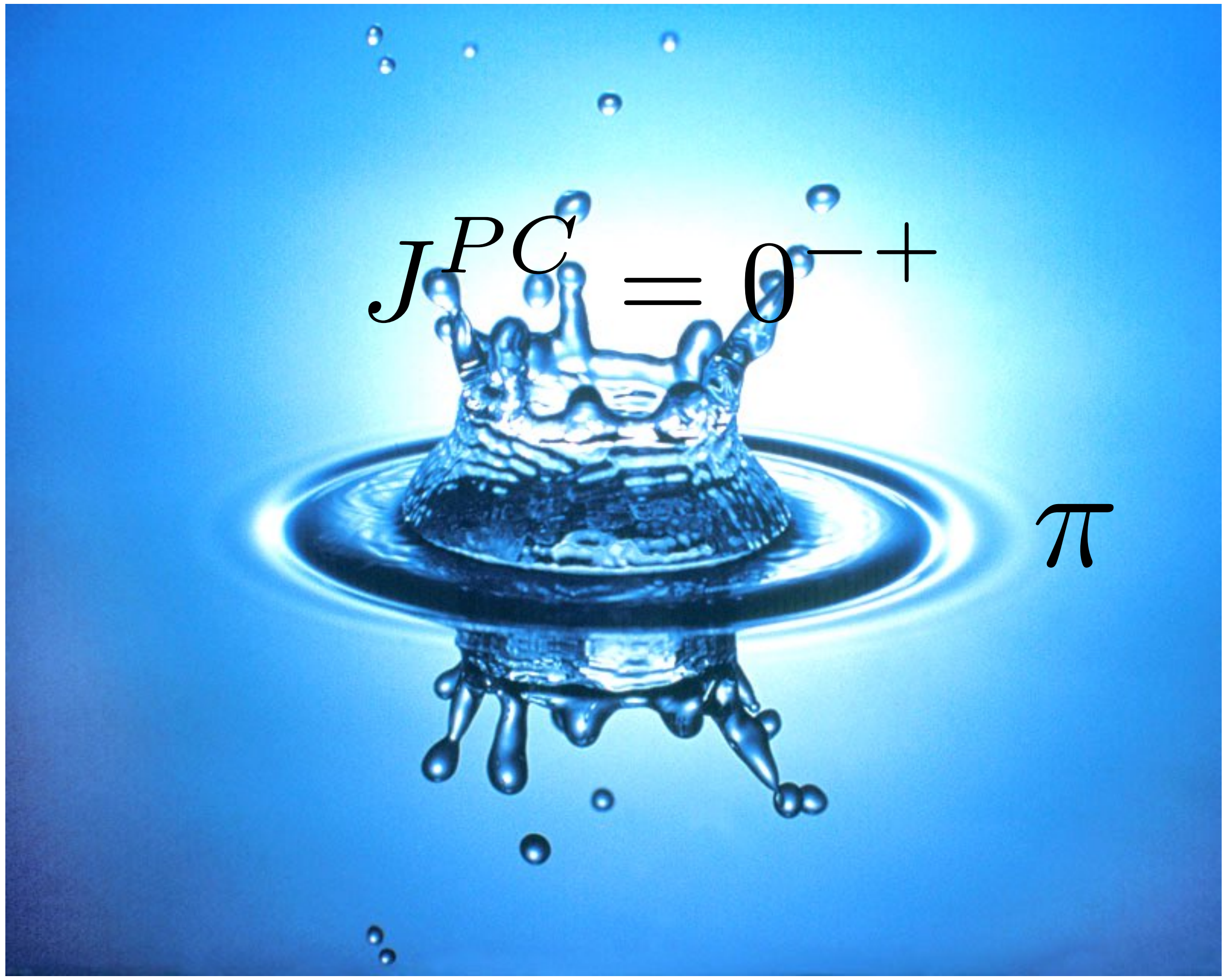
- SU(3) N_f = 8
- 64^3 x 128
- β = 4.8
- m_q = 0.00125

$N = 232$



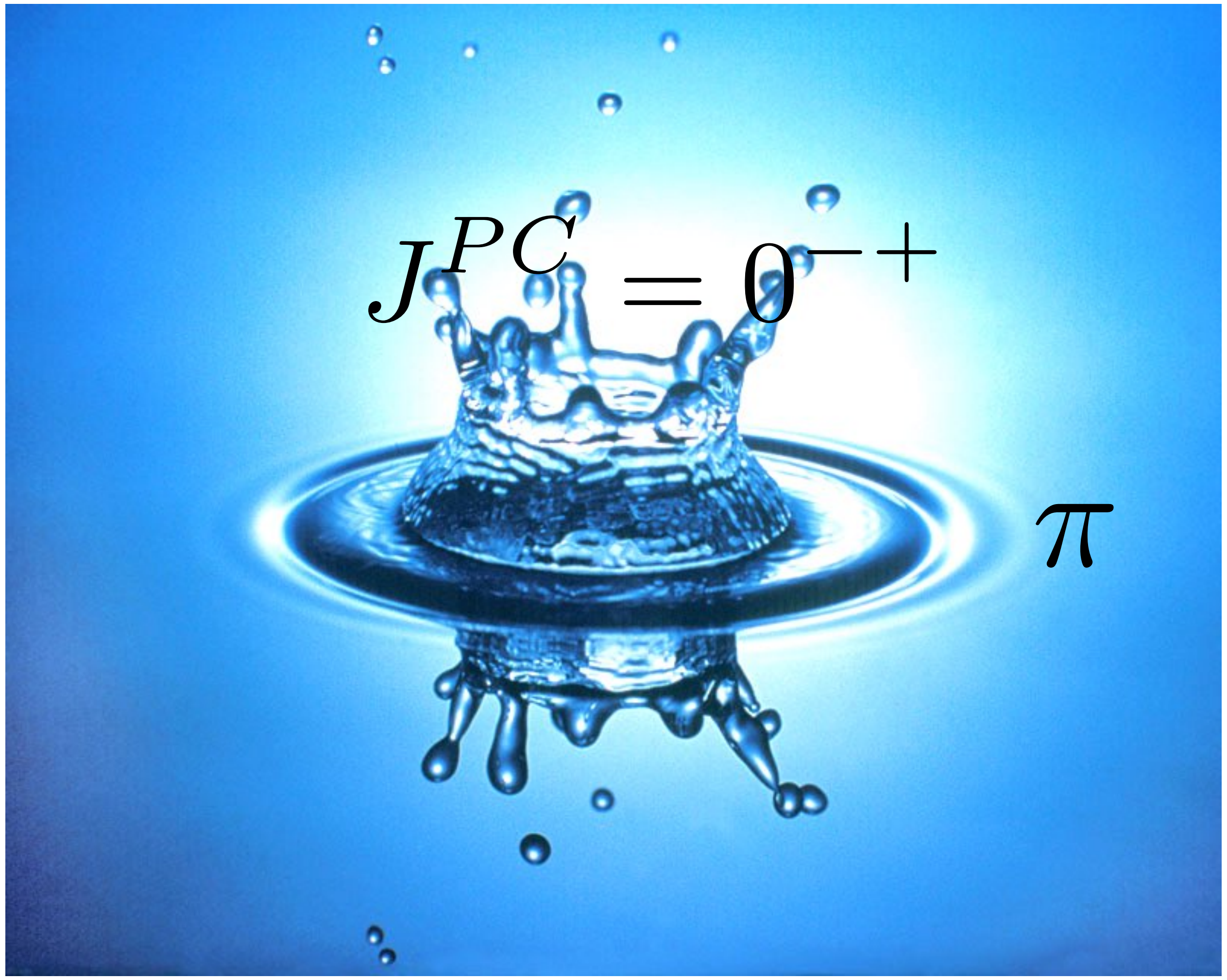
Energy Spectrum + Matrix Elements

$C(t) = \langle 0|\bar{\pi}(t)\pi(0)|0\rangle$



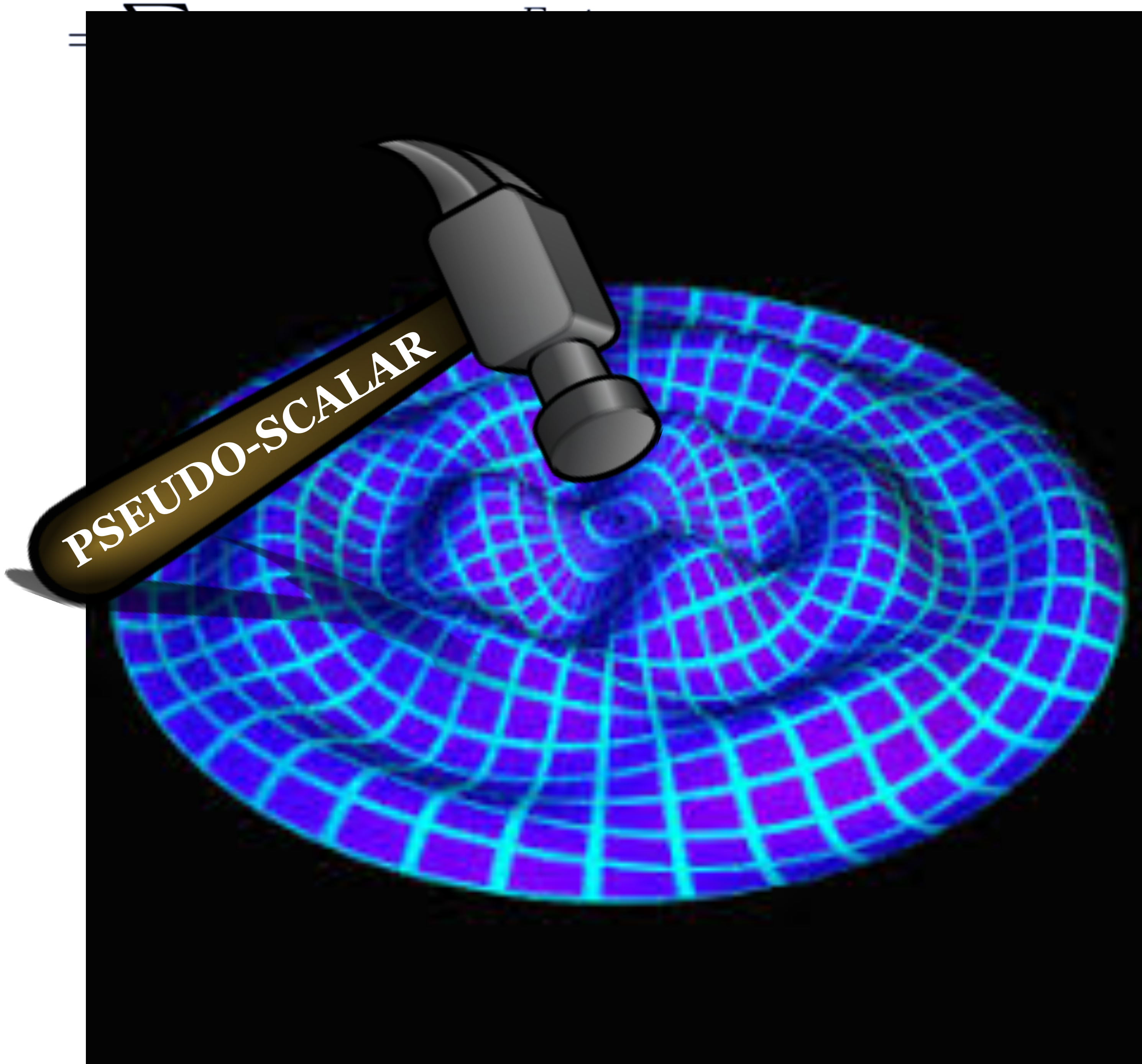
Energy Spectrum + Matrix Elements

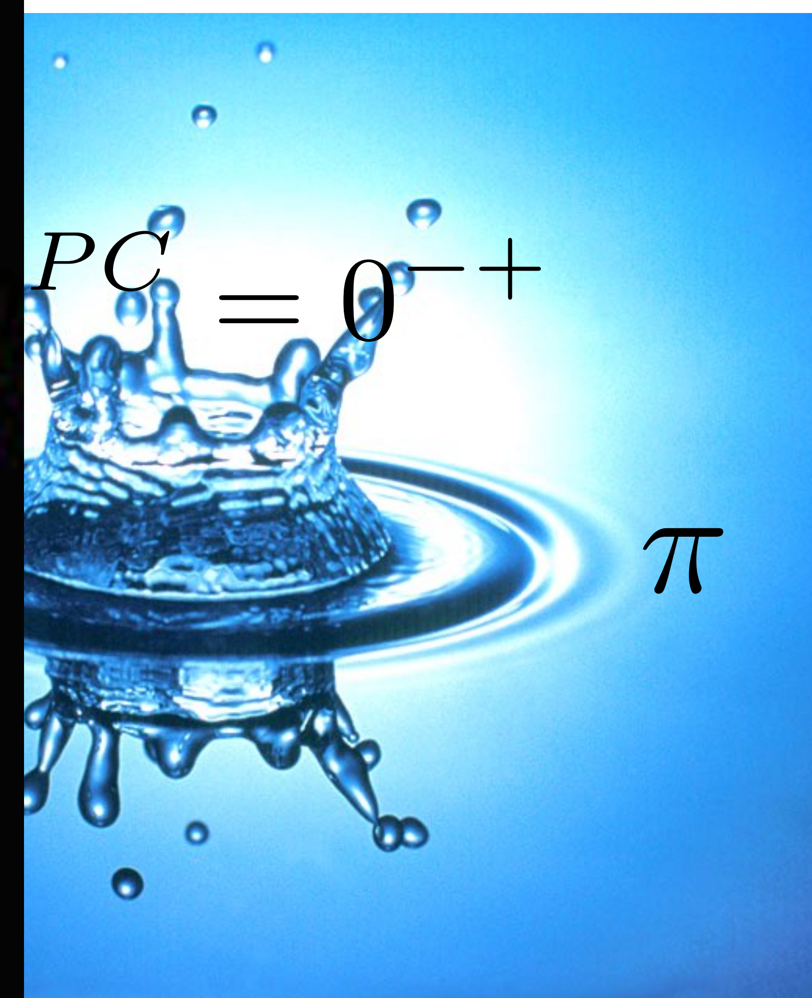
$$C(t) = \langle 0|\bar{\pi}(t)\pi(0)|0\rangle$$
$$= \sum_{m}^{\infty} \langle 0|\pi(0)|E_m\rangle e^{-E_m t} \langle E_m|\pi(0)|0\rangle$$



Energy Spectrum + Matrix Elements

$$C(t) = \langle 0|\bar{\pi}(t)\pi(0)|0\rangle$$



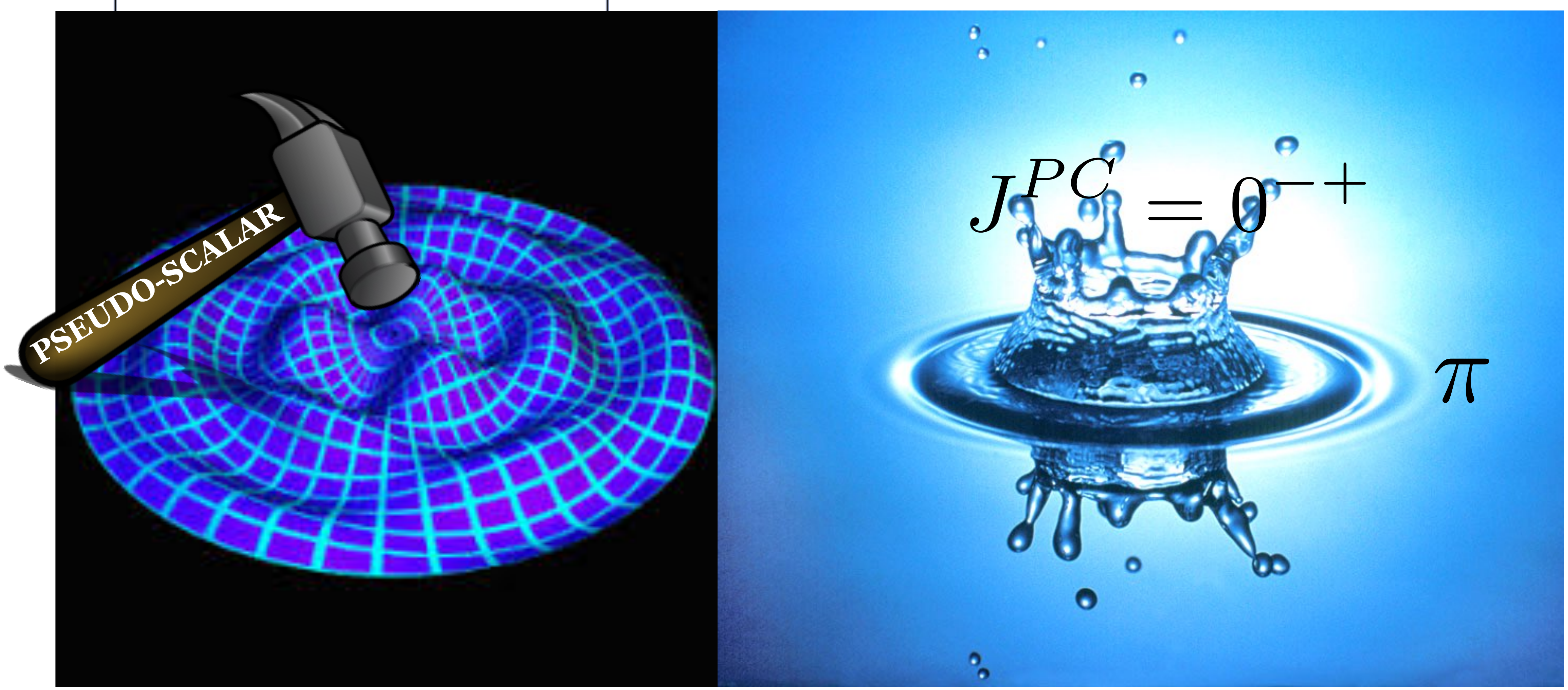


http://watersoundimage.yolasite.com/what-is-a-w-s-image.php

Energy Spectrum + Matrix Elements

$$C(t) = \langle 0|\bar{\pi}(t)\pi(0)|0\rangle$$
$$= \sum_{m}^{\infty} \langle 0|\pi(0)|E_m\rangle \mathrm{e}^{-E_m t} \langle E_m|\pi(0)|0\rangle$$

$$\boxed{\Rightarrow C(t) = \sum^{\infty} a_m\, \mathrm{e}^{-E_m t}}$$



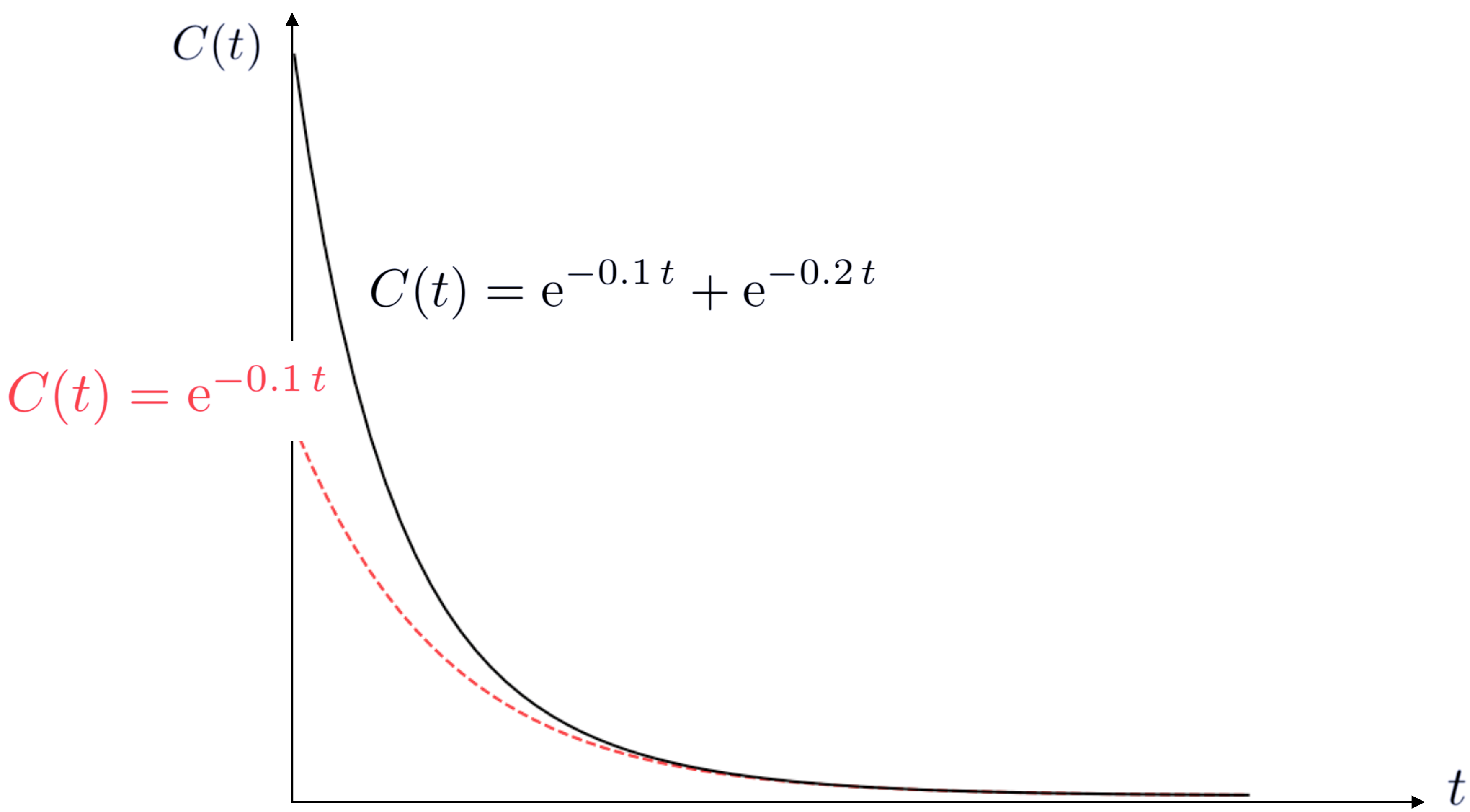
http://watersoundimage.yolasite.com/what-is-a-w-s-image.php

Energy Spectrum + Matrix Elements

$$
\begin{aligned}
C(t) &= \langle 0|\bar{\pi}(t)\pi(0)|0\rangle \\
&= \sum_{m}^{\infty} \langle 0|\pi(0)|E_m\rangle \mathrm{e}^{-E_m t} \langle E_m|\pi(0)|0\rangle
\end{aligned}
$$

$$
\begin{aligned}
C(t) &= a_0 \mathrm{e}^{-E_0 t} + a_1 \mathrm{e}^{-E_1 t} + \cdots \\
&\approx a_0 \mathrm{e}^{-E_0 t}
\end{aligned}
$$

$$
\boxed{\Rightarrow C(t) = \sum_{m}^{\infty} a_m \, \mathrm{e}^{-E_m t}}
$$



Energy Spectrum + Matrix Elements

$$C(t) = \langle 0|\bar{\pi}(t)\pi(0)|0\rangle$$
$$= \sum_m^{\infty} \langle 0|\pi(0)|E_m\rangle \mathrm{e}^{-E_m t}\langle E_m|\pi(0)|0\rangle$$

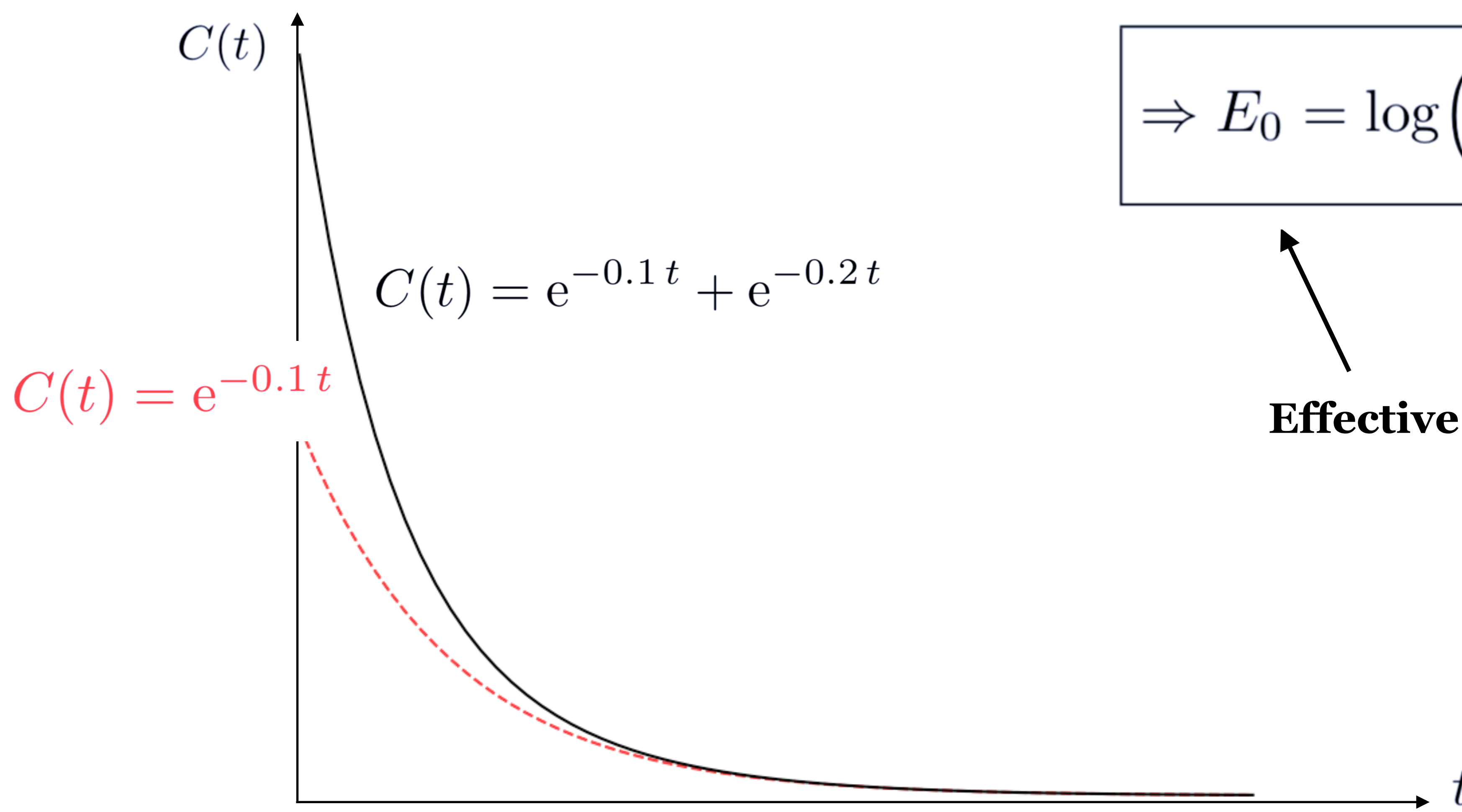
$$\boxed{\Rightarrow C(t) = \sum_m^{\infty} a_m\, \mathrm{e}^{-E_m t}}$$

$$C(t) = a_0\mathrm{e}^{-E_0 t} + a_1\mathrm{e}^{-E_1 t} + \cdots$$
$$\approx a_0\mathrm{e}^{-E_0 t}$$

$$\frac{C(t+1)}{C(t)} \approx \frac{a_1\mathrm{e}^{-E_0(t+1)}}{a_1\mathrm{e}^{-E_0 t}}$$
$$= \mathrm{e}^{-E_0}$$

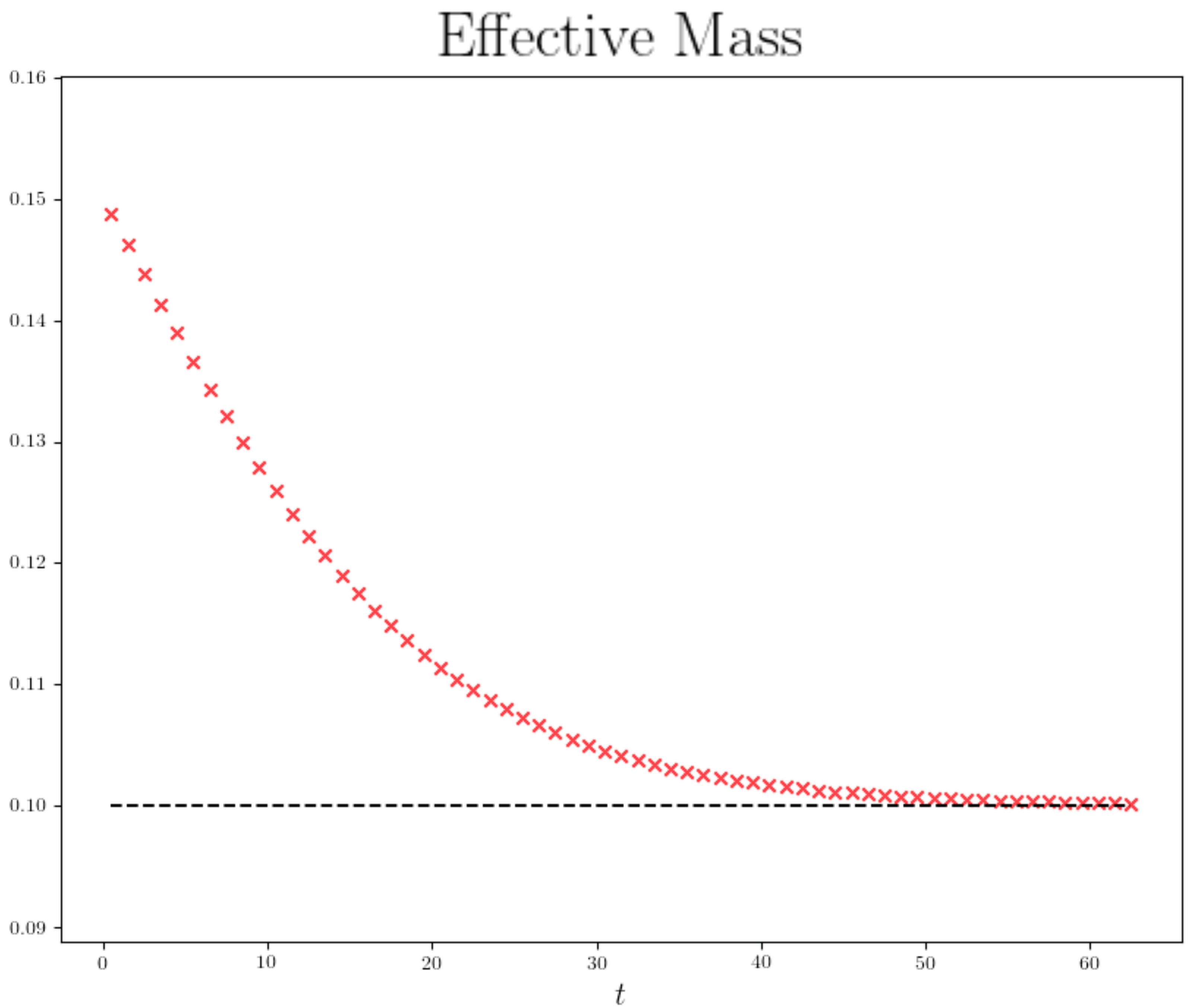
$$\boxed{\Rightarrow E_0 = \log\left(\frac{C(t)}{C(t+1)}\right)}$$

Effective mass



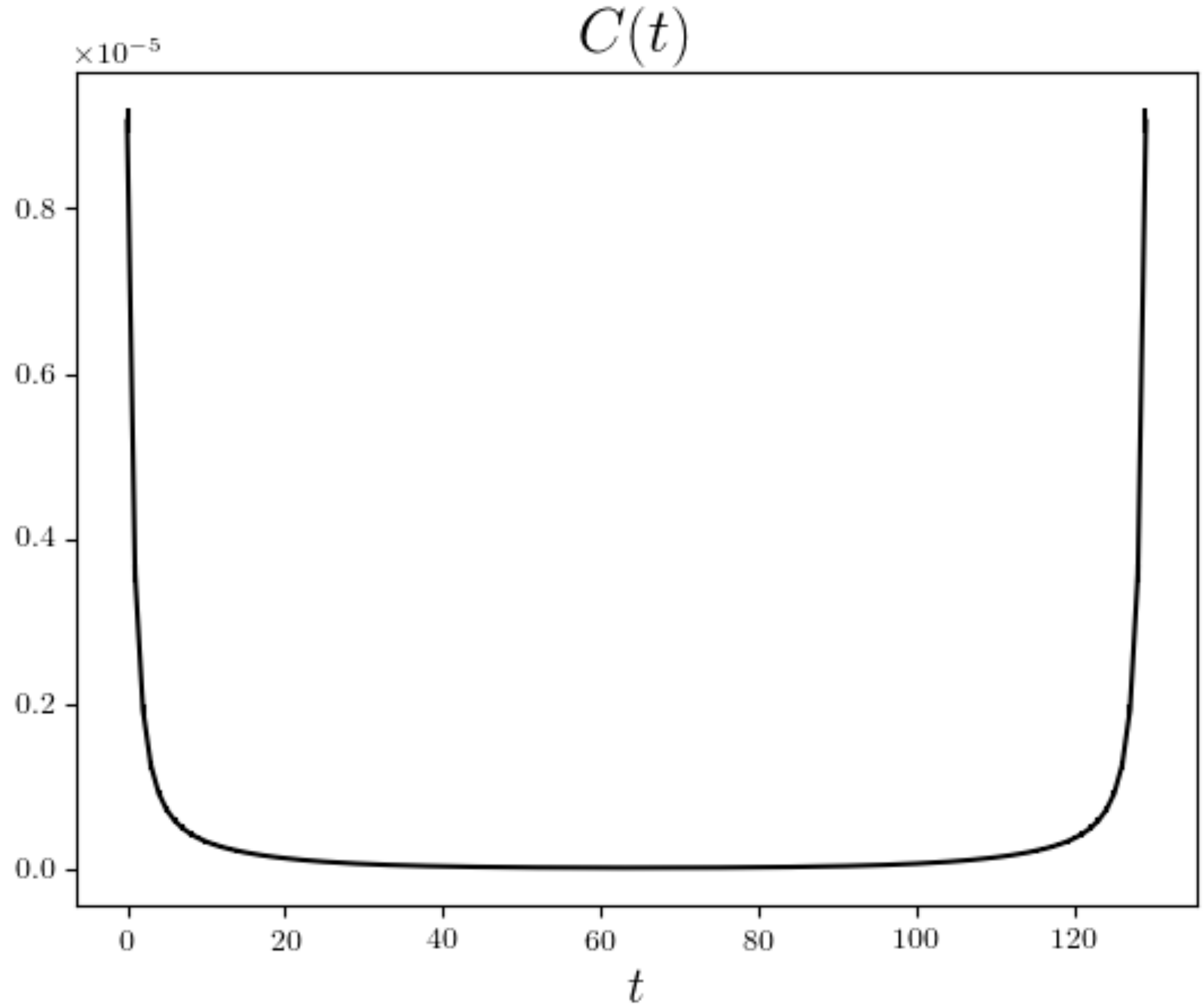
Pion Effective Mass

$$\boxed{\Rightarrow E_1 = \log\left(\frac{C(t)}{C(t+1)}\right)}$$



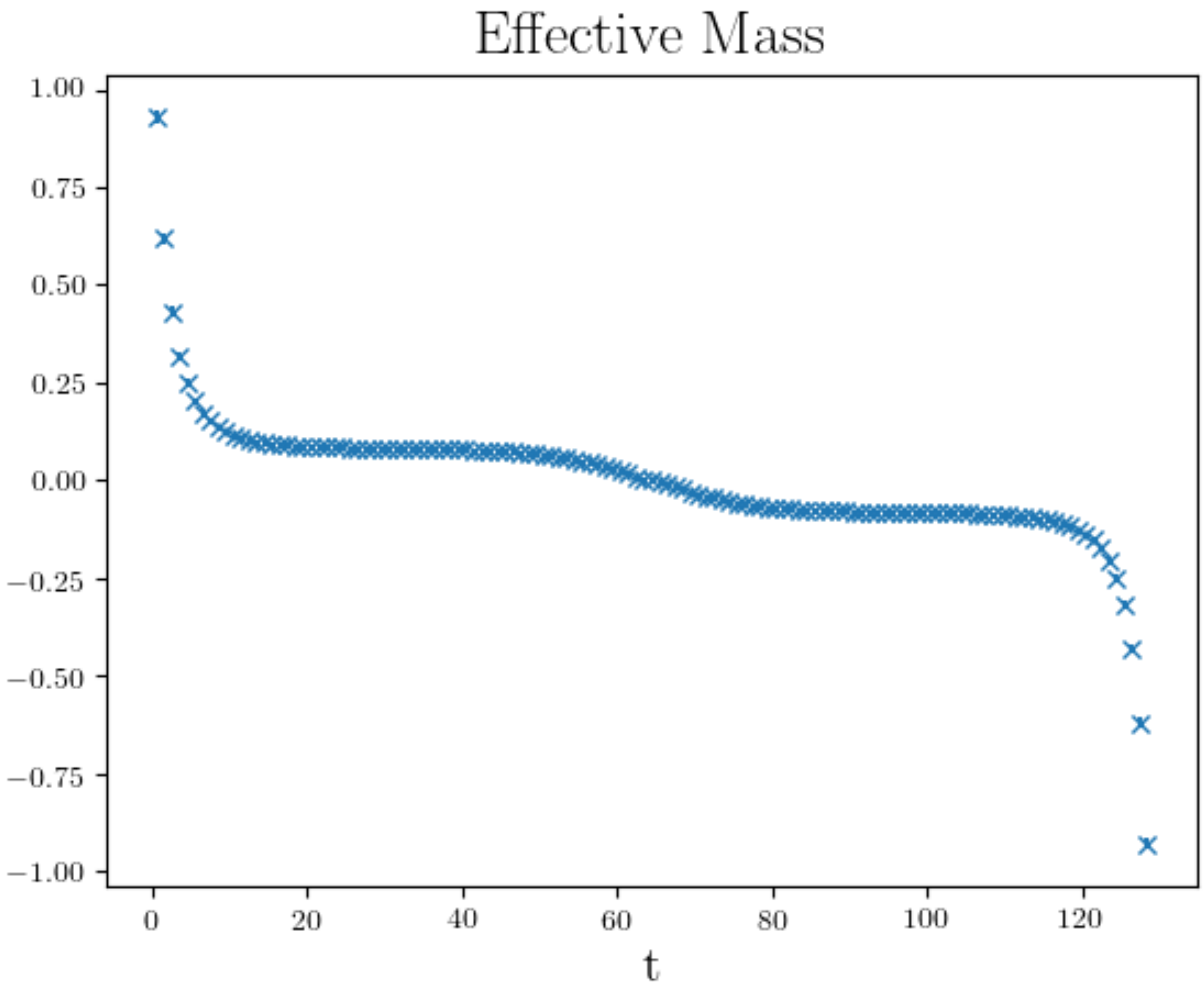
Pion Effective Mass

$$\boxed{\Rightarrow E_1 = \log\Big(\frac{C(t)}{C(t+1)}\Big)}$$



Pion Effective Mass

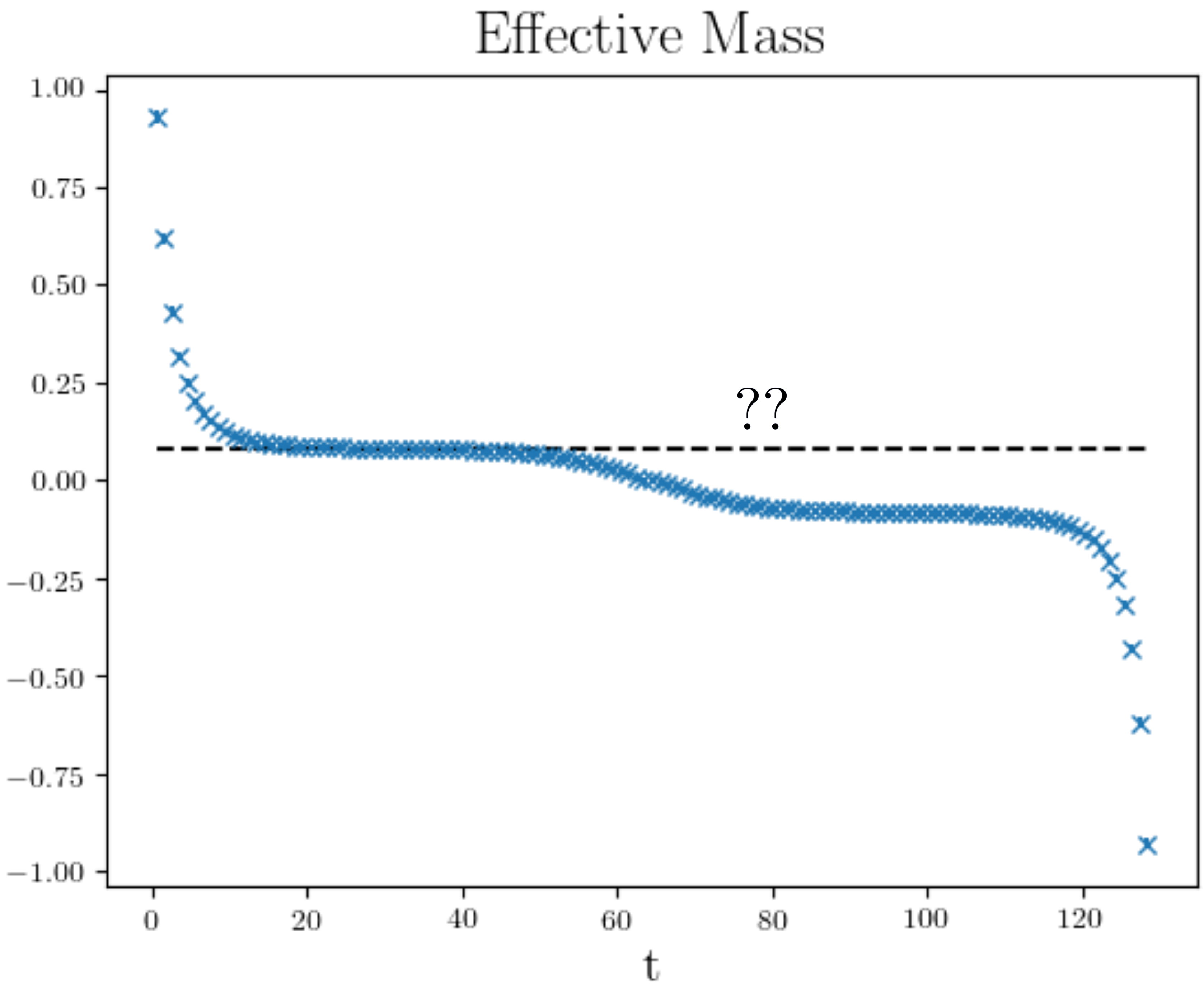
$$\Rightarrow E_1 = \log\left(\frac{C(t)}{C(t+1)}\right)$$



Pion Effective Mass

$$\Rightarrow E_1 = \log\left(\frac{C(t)}{C(t+1)}\right)$$

$$\Rightarrow C(t) = \sum_m^{\infty} a_m \, \mathrm{e}^{-E_m t}$$

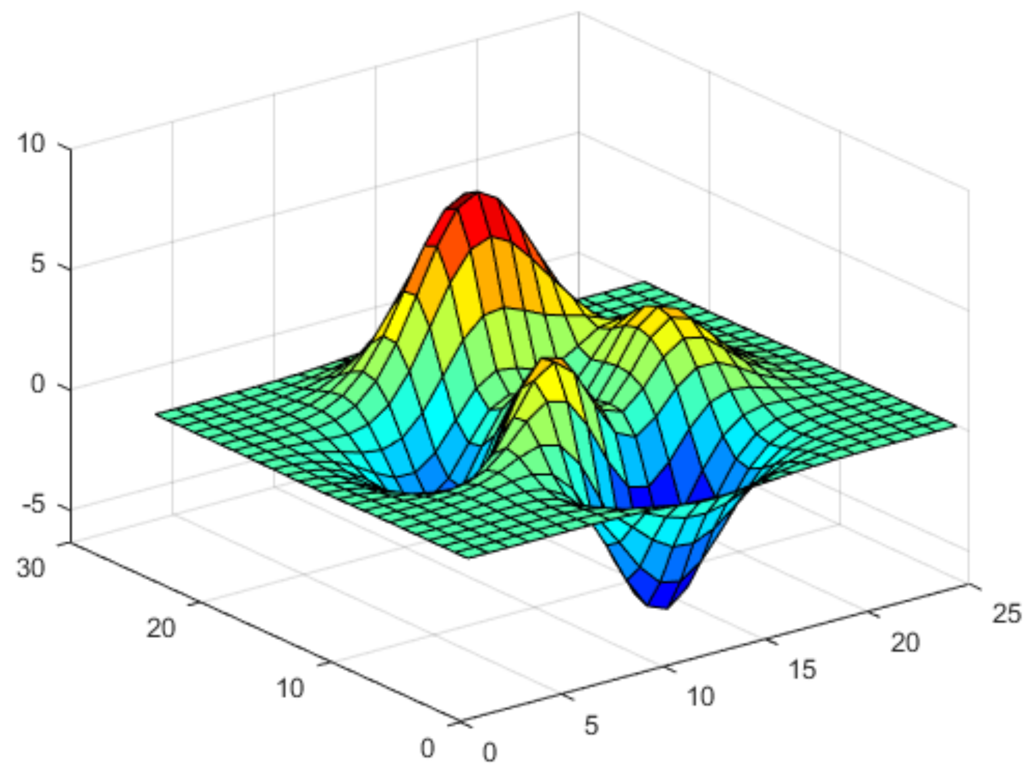


Standard Method for Excited States

Fit more exponentials! …. but

$$\boxed{\Rightarrow C(t) = \sum_{m}^{\infty} a_m \, \mathrm{e}^{-E_m t}}$$

- **Difficult/time consuming with decaying exponentials**
- **Finding global minimum - best** a_m **,** E_m **is hard**
- $2M$ **dimensional parameter space**
- **Prone to user bias - choosing initial values**



https://www.mathworks.com

Algebraic Approach - Prony 1700's

Prony, G. R. B. "J. de Lh Ecole Polytechnique." Paris 1 (1795): 24.

G. Fleming, S. Cohen, H. Lin, V. Pereyra (2009)

Algebraic Approach - Prony 1700's

Prony, G. R. B. "J. de Lh Ecole Polytechnique." Paris 1 (1795): 24.

$$
\begin{aligned}
y_n(t) &\equiv C(t+n) \\
&= \sum_m^M a_m \mathrm{e}^{-E_m(t+n)} \\
&= \sum_m^M a_m \mathrm{e}^{-E_m t} \mathrm{e}^{-E_m n}
\end{aligned}
$$

$$
\boxed{\Rightarrow y_n(t) = \sum_m^M A_m(t) z_m^n}
$$

$$
\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{2M-1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_M \\ z_1^2 & z_2^2 & \cdots & z_M^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{2M-1} & z_2^{2M-1} & \cdots & z_M^{2M-1} \end{pmatrix}}_{\text{Vandermode matrix}} \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_M \end{pmatrix}
$$

G. Fleming, S. Cohen, H. Lin, V. Pereyra (2009)

Algebraic Approach - Prony 1700's

Prony, G. R. B. "J. de Lh Ecole Polytechnique." Paris 1 (1795): 24.

$$
\begin{aligned}
y_n(t) &\equiv C(t+n) \\
&= \sum_m^M a_m \mathrm{e}^{-E_m(t+n)} \\
&= \sum_m^M a_m \mathrm{e}^{-E_m t} \mathrm{e}^{-E_m n}
\end{aligned}
$$

$$\boxed{\Rightarrow y_n(t) = \sum_m^M A_m(t) z_m^n}$$

$$
\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{2M-1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_M \\ z_1^2 & z_2^2 & \cdots & z_M^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{2M-1} & z_2^{2M-1} & \cdots & z_M^{2M-1} \end{pmatrix}}_{\text{Vandermode matrix}} \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_M \end{pmatrix}
$$

Non-linear

linear

G. Fleming, S. Cohen, H. Lin, V. Pereyra (2009)

Algebraic Approach - Prony 1700's

Prony, G. R. B. "J. de Lh Ecole Polytechnique." Paris 1 (1795): 24.

$$
\begin{aligned}
y_n(t) &\equiv C(t+n) \\
&= \sum_m^M a_m \mathrm{e}^{-E_m(t+n)} \\
&= \sum_m^M a_m \mathrm{e}^{-E_m t} \mathrm{e}^{-E_m n}
\end{aligned}
$$

$$
\boxed{\Rightarrow y_n(t) = \sum_m^M A_m(t) z_m^n}
$$

$$
\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{2M-1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_M \\ z_1^2 & z_2^2 & \cdots & z_M^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{2M-1} & z_2^{2M-1} & \cdots & z_M^{2M-1} \end{pmatrix}}_{\text{Vandermode matrix}} \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_M \end{pmatrix}
$$

Non-linear **linear**

$$
0 = \begin{vmatrix} \underbrace{\begin{matrix} y_0 & y_1 & \cdots & y_{M-2} & y_{M-1} \\ y_1 & y_2 & \cdots & y_{M-1} & y_M \\ y_2 & y_3 & \cdots & y_M & y_{M+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_M & y_{M+1} & \cdots & y_{2M-2} & y_{2M-1} \end{matrix}}_{\text{Hankel matrix}} & \begin{matrix} 1 \\ z \\ z^2 \\ \vdots \\ z^M \end{matrix} \end{vmatrix}
$$

We can do this!!

G. Fleming, S. Cohen, H. Lin, V. Pereyra (2009)

Prony's Method for M= 3

M = 3 means 6 y's

$$0 = \begin{vmatrix} y_0 & y_1 & y_2 & 1 \\ y_1 & y_2 & y_3 & z \\ y_2 & y_3 & y_4 & z^2 \\ y_3 & y_4 & y_5 & z^3 \end{vmatrix}$$

Prony's Method for M= 3

M = 3 means 6 y's

$$0 = \begin{vmatrix} y_0 & y_1 & y_2 & 1 \\ y_1 & y_2 & y_3 & z \\ y_2 & y_3 & y_4 & z^2 \\ y_3 & y_4 & y_5 & z^3 \end{vmatrix}$$

Need to solve M^{th} order polynomial

$$0 = p_0 + p_1 z + p_2 z^2 + p_3 z^3 \qquad z_m = \mathrm{e}^{-E_m}$$

Prony's Method, Linear Prediction, Matrix Prony

E. Berkowitz, A. Nicholson, C. Chang et al. (2017)

S. Beane, W. Detmold, T. Luu et al. (2009)

Prony's Method, Linear Prediction, Matrix Prony

E. Berkowitz, A. Nicholson, C. Chang et al. (2017)

S. Beane, W. Detmold, T. Luu et al. (2009)

$$C(t) \to \begin{pmatrix} C_{11}(t) & C_{12}(t) \\ C_{21}(t) & C_{22}(t) \end{pmatrix}$$

$$C_{ij}(t) = \langle 0|\mathcal{O}_i(t)\mathcal{O}_j(0)|0\rangle$$

Prony's Method, Linear Prediction, Matrix Prony

E. Berkowitz, A. Nicholson, C. Chang et al. (2017)

S. Beane, W. Detmold, T. Luu et al. (2009)

$$C(t) \to \begin{pmatrix} C_{11}(t) & C_{12}(t) \\ C_{21}(t) & C_{22}(t) \end{pmatrix}$$

$$C_{ij}(t) = \langle 0|\mathcal{O}_i(t)\mathcal{O}_j(0)|0\rangle$$

Effective mass becomes generalized eigenvalue problem

$$0 = C(t+1) - zC(t) \qquad\qquad \mathrm{e}^{-E_0} = \frac{C(t+1)}{C(t)}$$

Prony's Method, Linear Prediction, Matrix Prony

E. Berkowitz, A. Nicholson, C. Chang et al. (2017)

S. Beane, W. Detmold, T. Luu et al. (2009)

$$C(t) \to \begin{pmatrix} C_{11}(t) & C_{12}(t) \\ C_{21}(t) & C_{22}(t) \end{pmatrix}$$

$$C_{ij}(t) = \langle 0|\mathcal{O}_i(t)\mathcal{O}_j(0)|0\rangle$$

Effective mass becomes generalized eigenvalue problem

$$0 = C(t+1) - zC(t) \qquad \mathrm{e}^{-E_0} = \frac{C(t+1)}{C(t)}$$

$$C(t+1)v = \lambda C(t)v \qquad \lambda = \mathrm{e}^{-E}$$

Prony's Method, Linear Prediction, Matrix Prony

Effective mass becomes generalized eigenvalue problem

$$\boxed{\lambda = \mathrm{e}^{-E} \qquad\qquad C(t+1)v = \lambda C(t)v}$$

$$C(t+1)v_0 = \mathrm{e}^{-E_0} C(t) v_0$$

$$C(t+1)v_1 = \mathrm{e}^{-E_1} C(t) v_1$$

Prony's Method, Linear Prediction, Matrix Prony

Effective mass becomes generalized eigenvalue problem

$$\boxed{\lambda = \mathrm{e}^{-E} \qquad\qquad C(t+1)v = \lambda C(t)v}$$

maximal overlap

$C(t+1)v_0 = \mathrm{e}^{-E_0}C(t)v_0 \longleftarrow$ **with ground state**

$C(t+1)v_1 = \mathrm{e}^{-E_1}C(t)v_1 \longleftarrow$ **with first excited state**

Prony's Method, Linear Prediction, Matrix Prony

Effective mass becomes generalized eigenvalue problem

$$\lambda = \mathrm{e}^{-E} \qquad\qquad C(t+1)v = \lambda C(t)v$$

maximal overlap

$$C(t+1)v_0 = \mathrm{e}^{-E_0} C(t) v_0 \longleftarrow$$ **with ground state**

$$C(t+1)v_1 = \mathrm{e}^{-E_1} C(t) v_1 \longleftarrow$$ **with first excited state**

Ground state matrix equation

$$\begin{pmatrix} C_{11}(t+1) & C_{12}(t+1) \\ C_{21}(t+1) & C_{22}(t+1) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \mathrm{e}^{-E_0} \begin{pmatrix} C_{11}(t) & C_{12}(t) \\ C_{21}(t) & C_{22}(t) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Prony's Method, Linear Prediction, Matrix Prony

Effective mass becomes generalized eigenvalue problem

$$\lambda = \mathrm{e}^{-E} \qquad\qquad C(t+1)v = \lambda C(t)v$$

maximal overlap

$$C(t+1)v_0 = \mathrm{e}^{-E_0}C(t)v_0 \longleftarrow$$ **with ground state**

$$C(t+1)v_1 = \mathrm{e}^{-E_1}C(t)v_1 \longleftarrow$$ **with first excited state**

Ground state matrix equation

$$\begin{pmatrix} C_{11}(t+1) & C_{12}(t+1) \\ C_{21}(t+1) & C_{22}(t+1) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \mathrm{e}^{-E_0} \begin{pmatrix} C_{11}(t) & C_{12}(t) \\ C_{21}(t) & C_{22}(t) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\mathrm{e}^{-E_0(t+1)} \begin{pmatrix} \langle 0|\mathcal{O}_1(b_1\mathcal{O}_1 + b_2\mathcal{O}_2)|0\rangle \\ \langle 0|\mathcal{O}_2(b_1\mathcal{O}_1 + b_2\mathcal{O}_2)|0\rangle \end{pmatrix} = \mathrm{e}^{-E_0}\mathrm{e}^{-E_0 t} \begin{pmatrix} \langle 0|\mathcal{O}_1(b_1\mathcal{O}_1 + b_2\mathcal{O}_2)|0\rangle \\ \langle 0|\mathcal{O}_2(b_1\mathcal{O}_1 + b_2\mathcal{O}_2)|0\rangle \end{pmatrix}$$

Prony’s Method, Linear Prediction, Matrix Prony

Hankel Matrix determinant specifies eigenvalue problem

$$0 = \begin{vmatrix} y_0 & y_1 & y_2 & 1 \\ y_1 & y_2 & y_3 & z \\ y_2 & y_3 & y_4 & z^2 \\ y_3 & y_4 & y_5 & z^3 \end{vmatrix}$$

M **solutions**

$$\boxed{0 = p_0 + p_1 z + p_2 z^2 + p_3 z^3 \qquad z_m = \mathrm{e}^{-E_m}}$$

Scalar equation

Prony's Method, Linear Prediction, Matrix Prony

Hankel Matrix determinant specifies eigenvalue problem

$$0 = \begin{vmatrix} y_0 & y_1 & y_2 & 1 \\ y_1 & y_2 & y_3 & z \\ y_2 & y_3 & y_4 & z^2 \\ y_3 & y_4 & y_5 & z^3 \end{vmatrix}$$

M **solutions**

$$\boxed{0 = p_0 + p_1 z + p_2 z^2 + p_3 z^3 \qquad z_m = \mathrm{e}^{-E_m}}$$

Scalar equation

$M \times N$ **solutions**

$$\boxed{0 = P_0 + \lambda P_1 + \lambda^2 P_2 + \lambda^3 P_3 \qquad \lambda_m = \mathrm{e}^{-E_m}}$$

Matrix equation

Prony's Method, Linear Prediction, Matrix Prony

Hankel Matrix determinant specifies eigenvalue problem

$$0 = \begin{vmatrix} y_0 & y_1 & y_2 & 1 \\ y_1 & y_2 & y_3 & z \\ y_2 & y_3 & y_4 & z^2 \\ y_3 & y_4 & y_5 & z^3 \end{vmatrix}$$

M **solutions**

$$\boxed{0 = p_0 + p_1 z + p_2 z^2 + p_3 z^3 \qquad z_m = \mathrm{e}^{-E_m}}$$

Scalar equation

Focus on labeling problem

$M \times N$ **solutions**

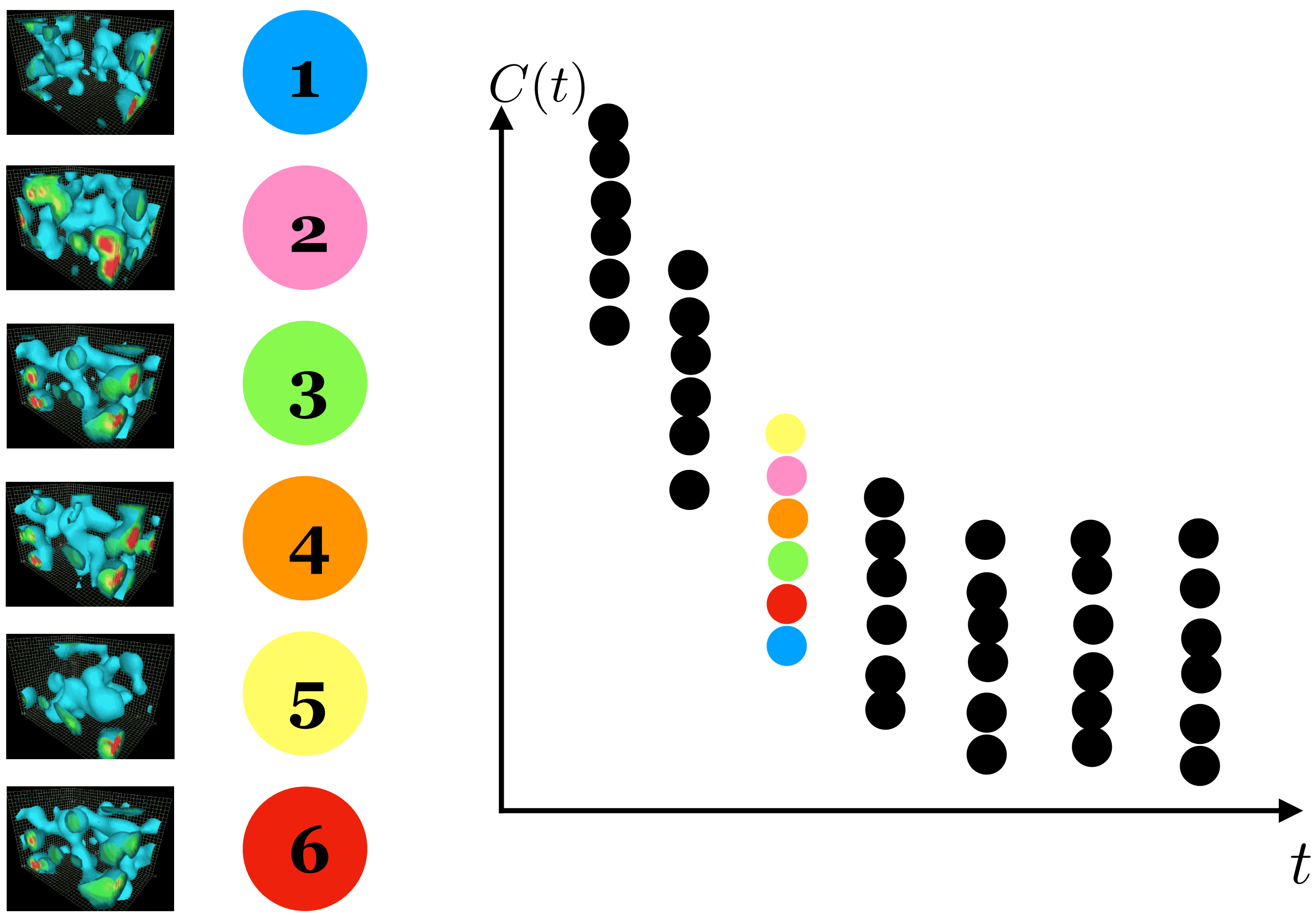
$$\boxed{0 = P_0 + \lambda P_1 + \lambda^2 P_2 + \lambda^3 P_3 \qquad \lambda_m = \mathrm{e}^{-E_m}}$$

Matrix equation

From theory to practice: the bootstrap

From theory to practice: the bootstrap

6 data points $C(t=t_o)$ for



From theory to practice: the bootstrap

6 data points $C(t=t_o)$ for

1
2
3
4
5
6

Bootstrap

$$\frac{1}{6}(1+1+2+3+5+5) = C_1(t=t_o)$$

$$\frac{1}{6}(1+2+2+3+4+6) = C_2(t=t_o)$$

$$\frac{1}{6}(2+3+3+3+4+5) = C_3(t=t_o)$$

$$\frac{1}{6}(1+1+2+5+6+6) = C_4(t=t_o)$$

$$\frac{1}{6}(1+2+3+4+4+5) = C_5(t=t_o)$$

$$\frac{1}{6}(1+2+3+5+5+6) = C_6(t=t_o)$$

$$\frac{1}{6}(2+3+3+4+6+6) = C_7(t=t_o)$$

$$\frac{1}{6}(1+1+2+3+5+5) = C_8(t=t_o)$$

From theory to practice: the bootstrap

$$\frac{1}{6}(1 + 1 + 2 + 3 + 5 + 5) = C_1(t=t_0)$$

Repeat for all times to obtain $C_1(0), C_1(1), C_1(2), C_1(3), \ldots C_1(T)$

From theory to practice: the bootstrap

$$\frac{1}{6}(1+1+2+3+5+5) = C_1(t=t_0)$$

Repeat for all times to obtain $C_1(0), C_1(1), C_1(2), C_1(3), \dots C_1(T)$

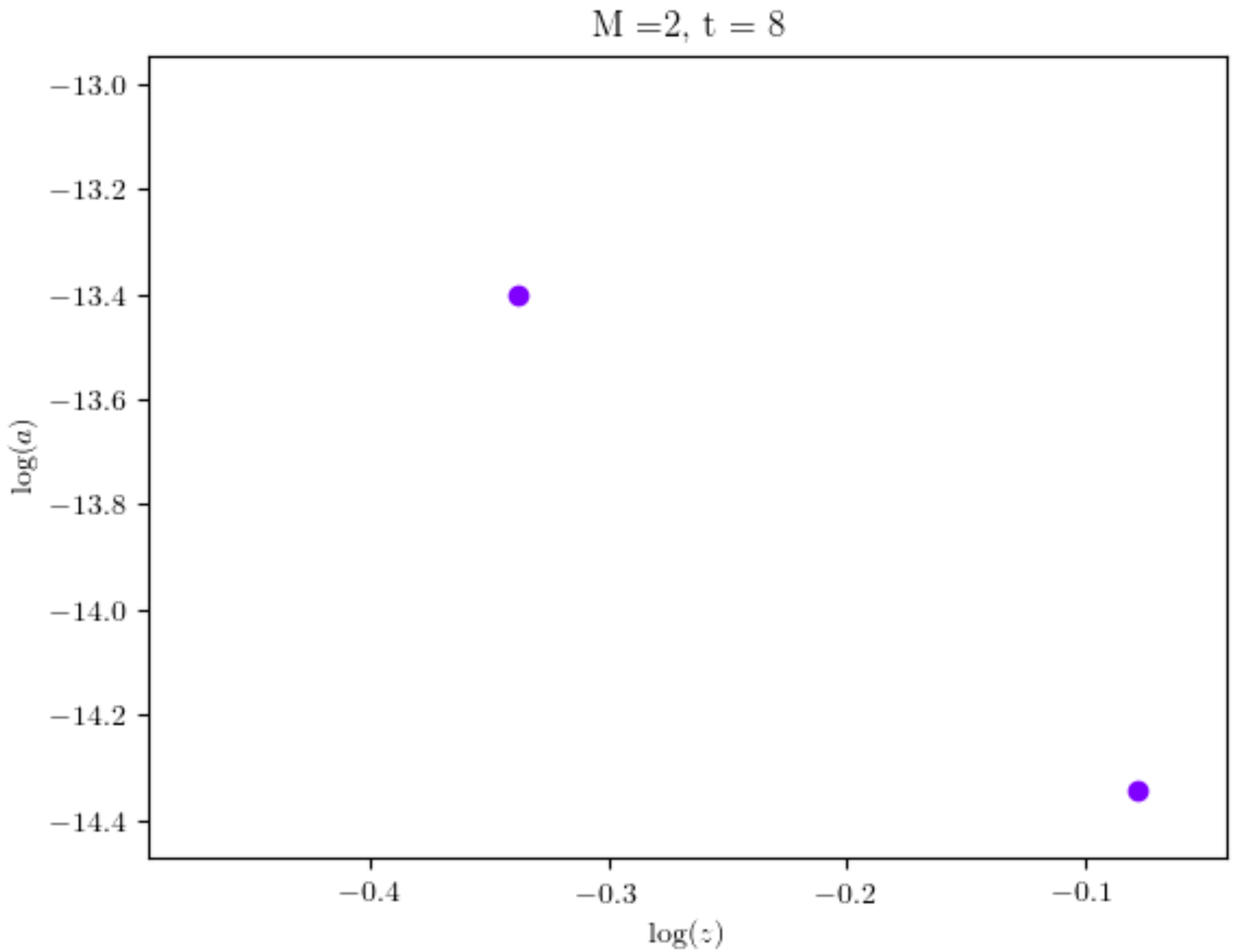
$$y_n(t) = C(t+n)$$

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ z_1 & z_2 \\ z_1^2 & z_2^2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \qquad 0 = \begin{vmatrix} y_0 & y_1 & 1 \\ y_1 & y_2 & z \\ y_2 & y_3 & z^2 \end{vmatrix}$$

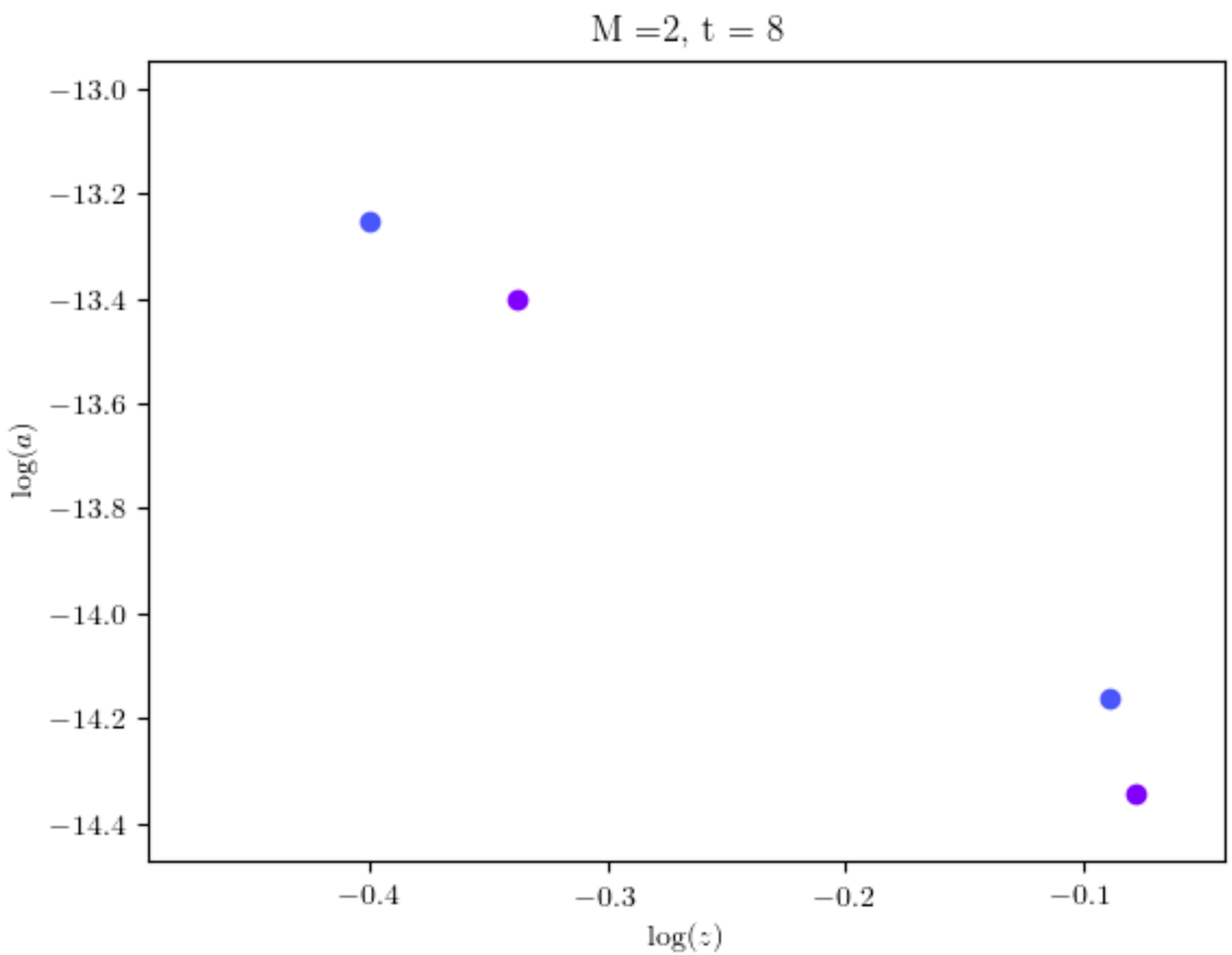
$$C_1(t) \Rightarrow \{\ (z_1, a_1),\ (z_2, a_2)\ \}$$

From theory to practice: the bootstrap

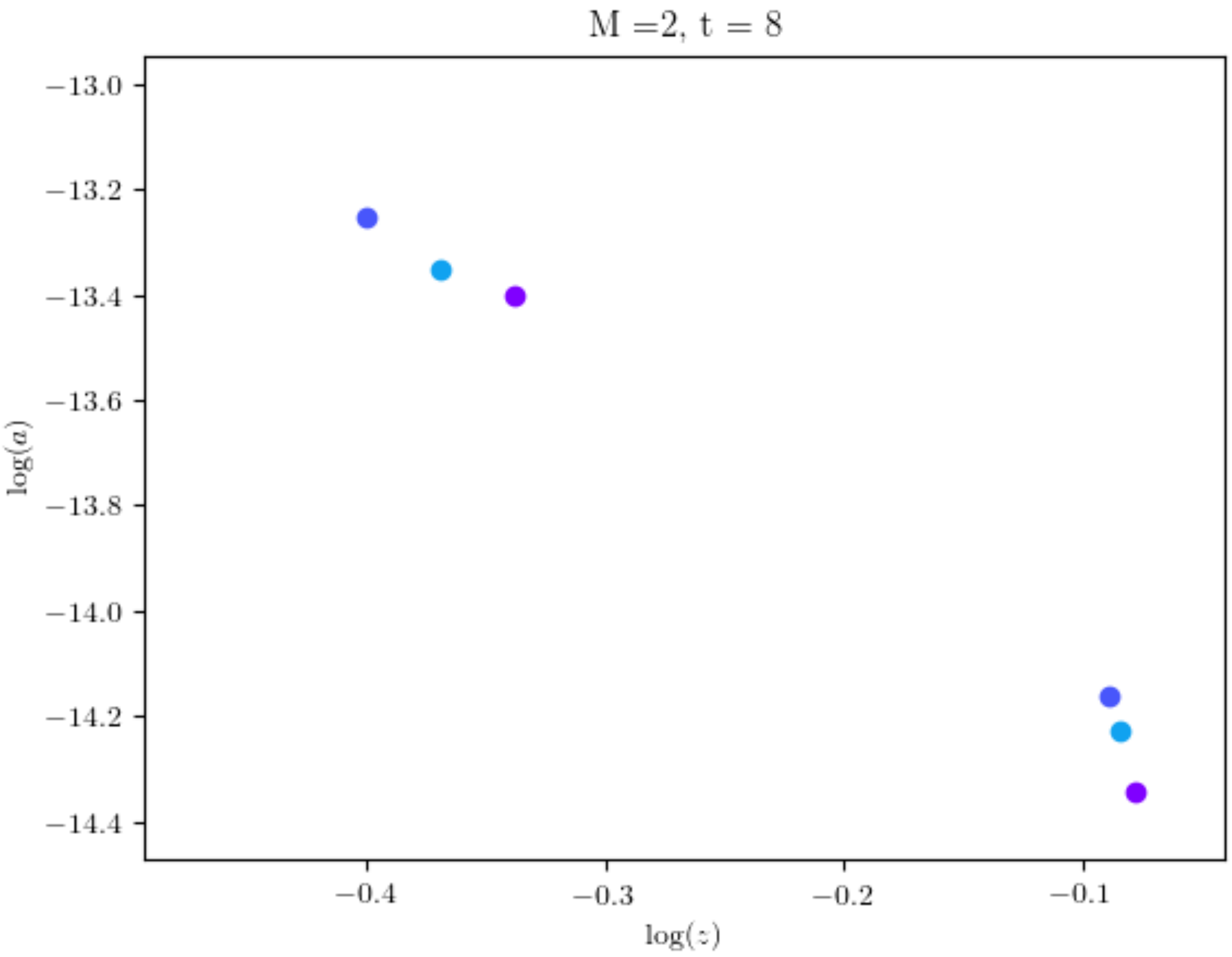
$a_m = |\langle 0|\pi|E_m\rangle|^2 \qquad z_m = \mathrm{e}^{-E_m}$



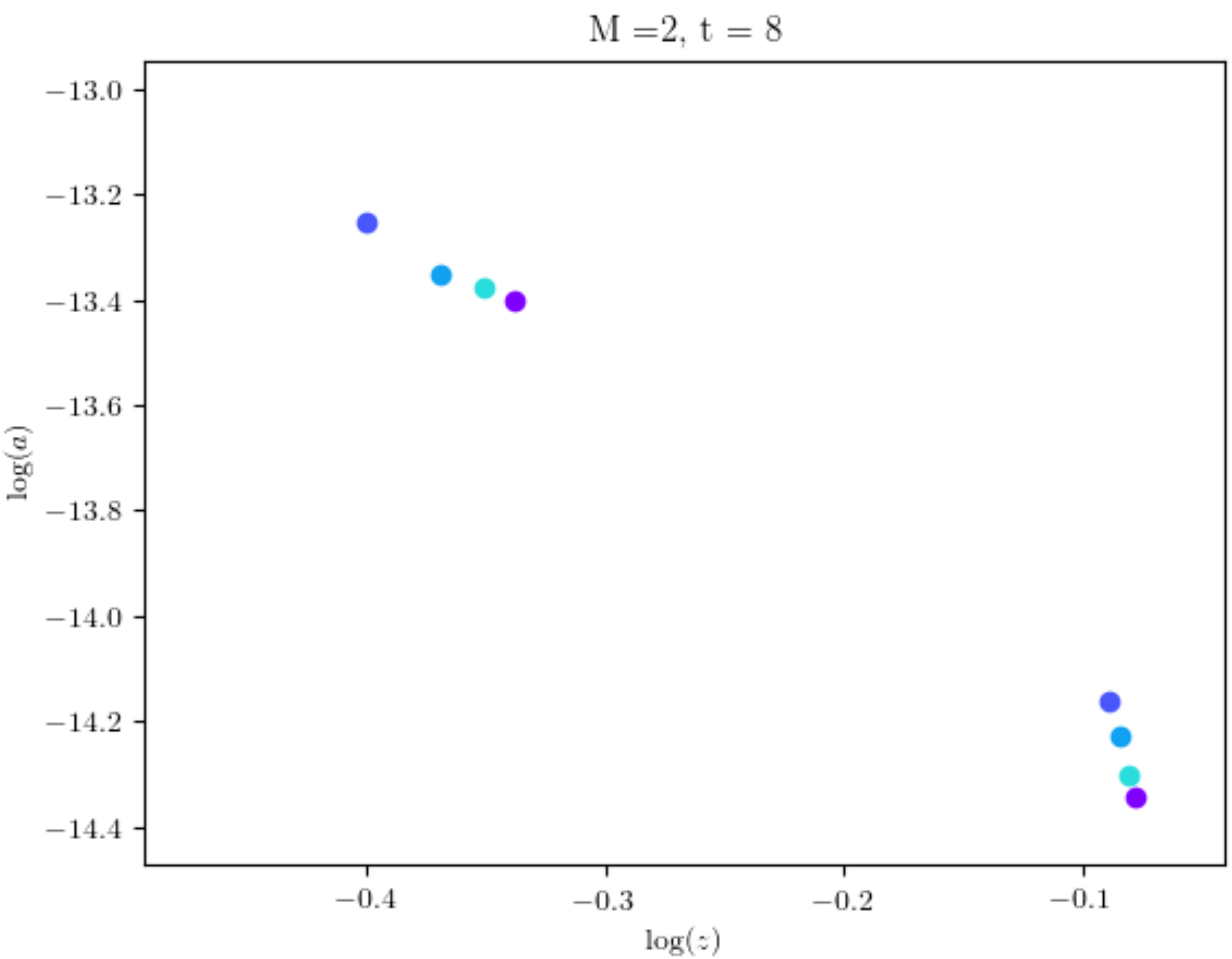
From theory to practice: the bootstrap



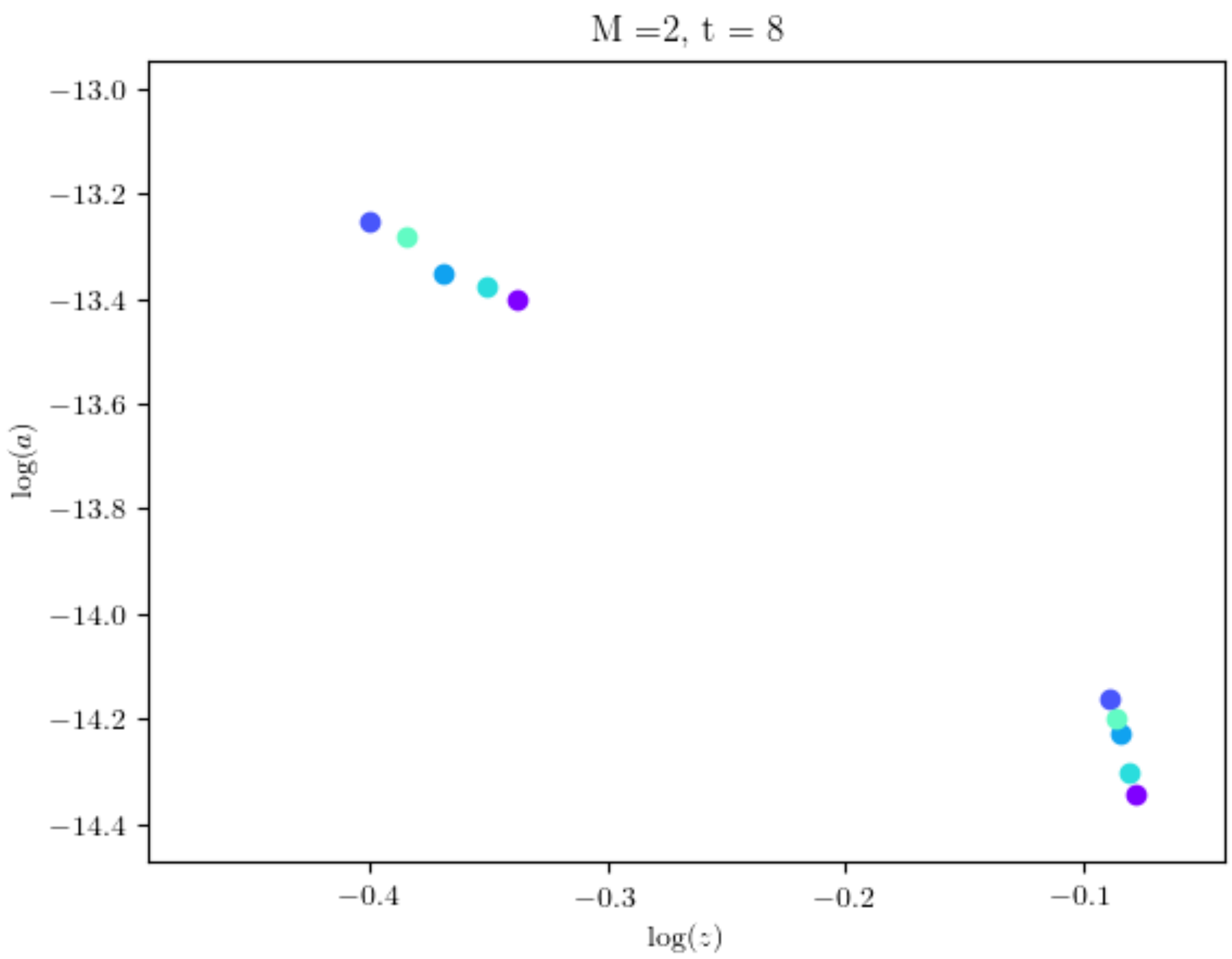
From theory to practice: the bootstrap



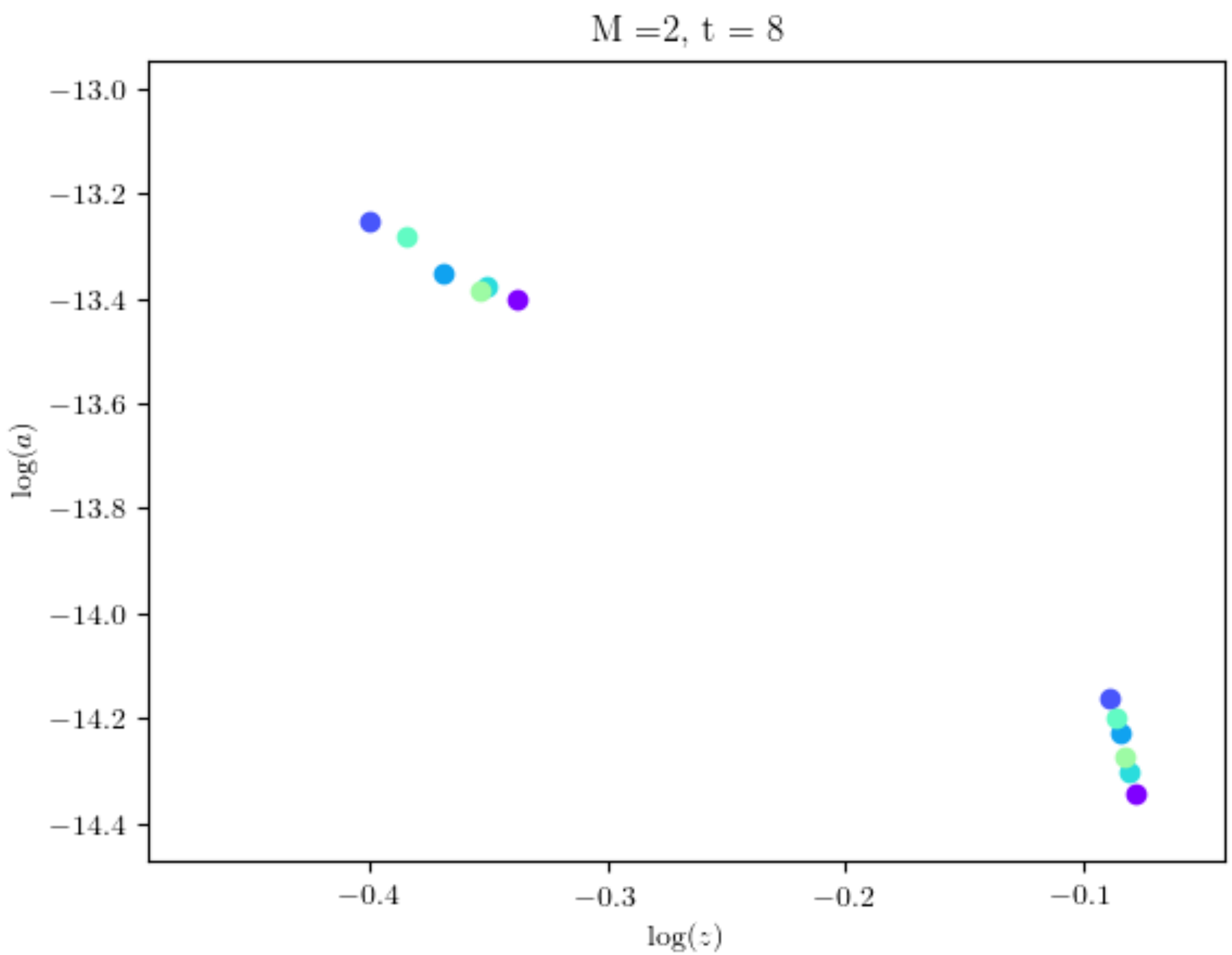
From theory to practice: the bootstrap



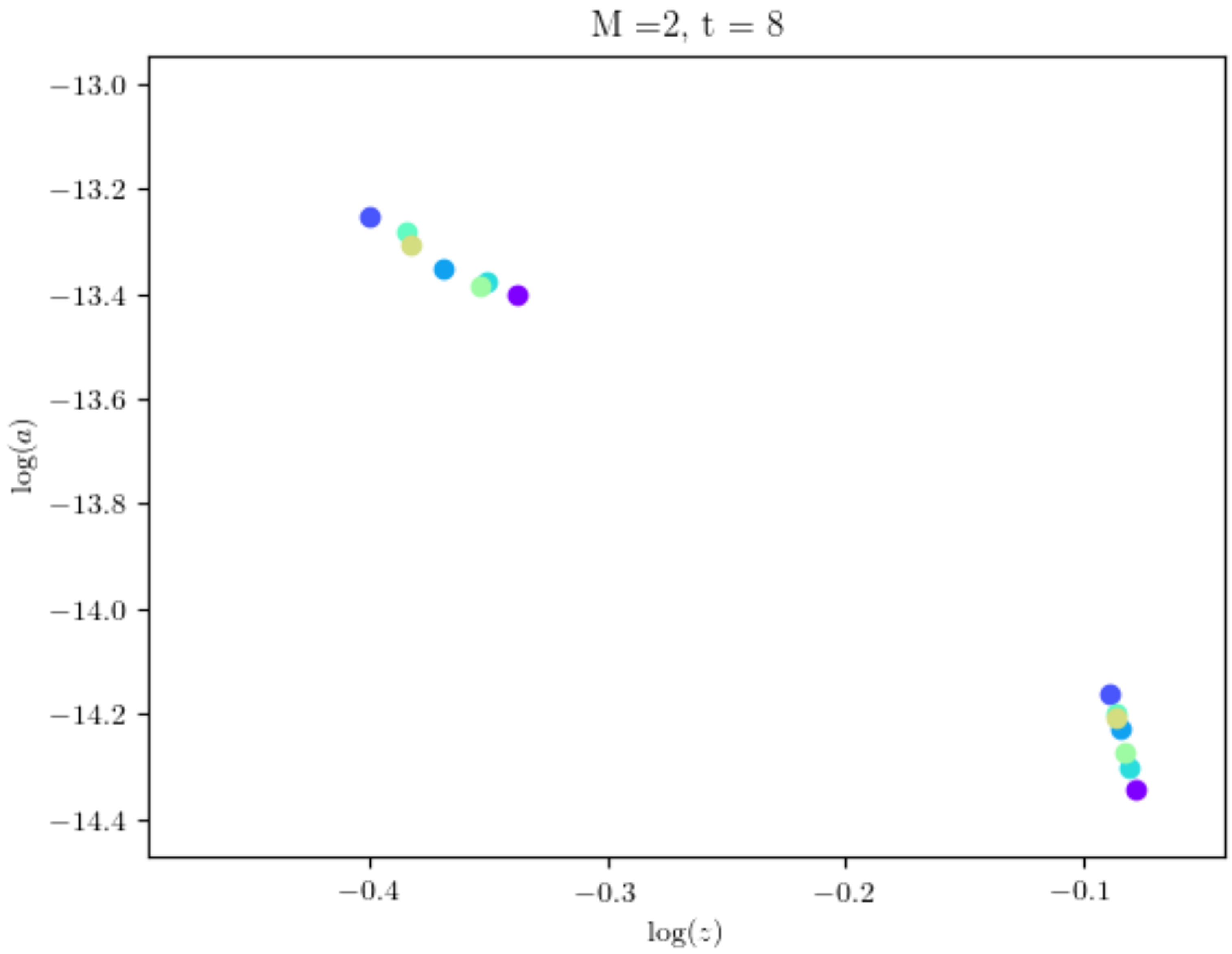
From theory to practice: the bootstrap



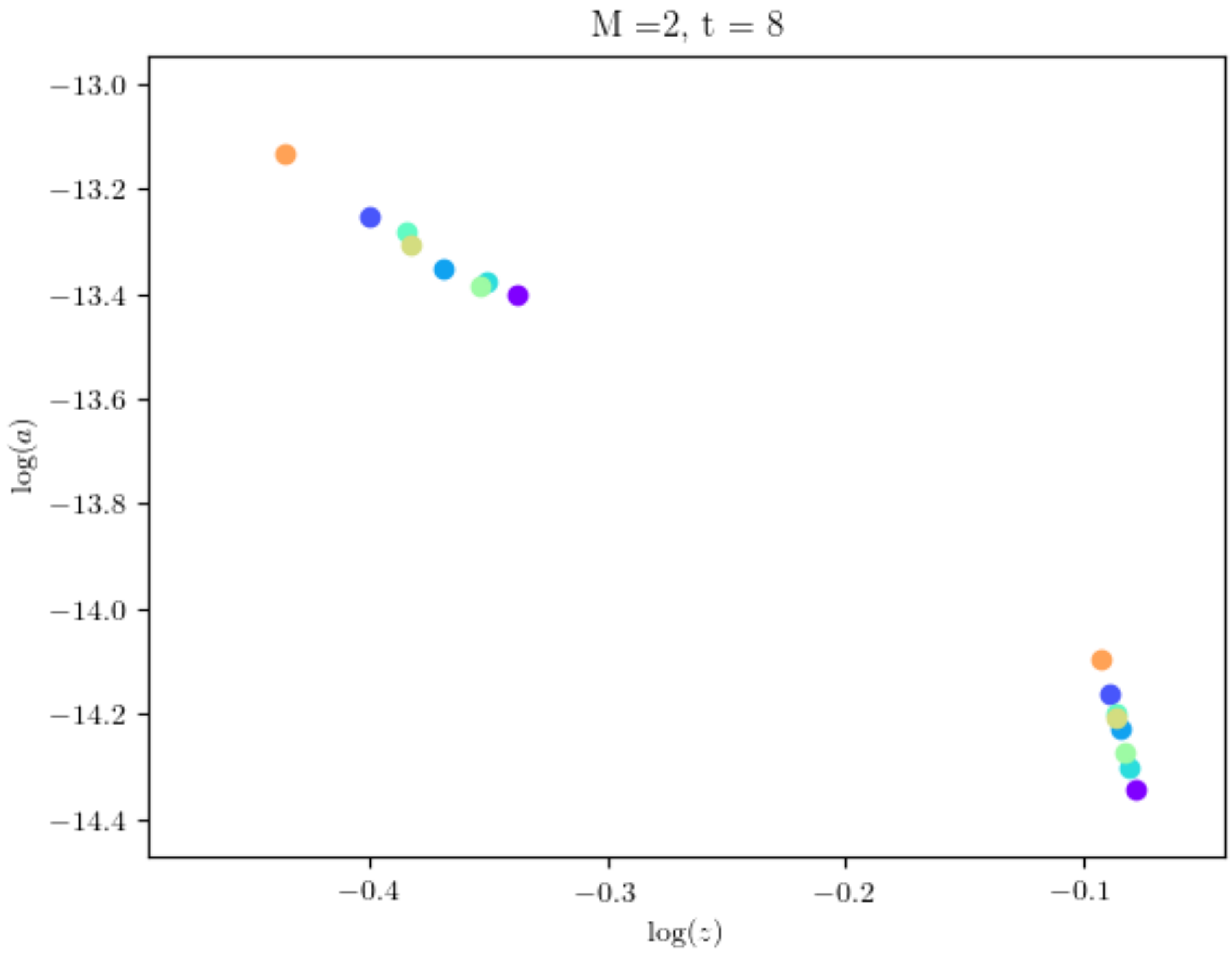
From theory to practice: the bootstrap



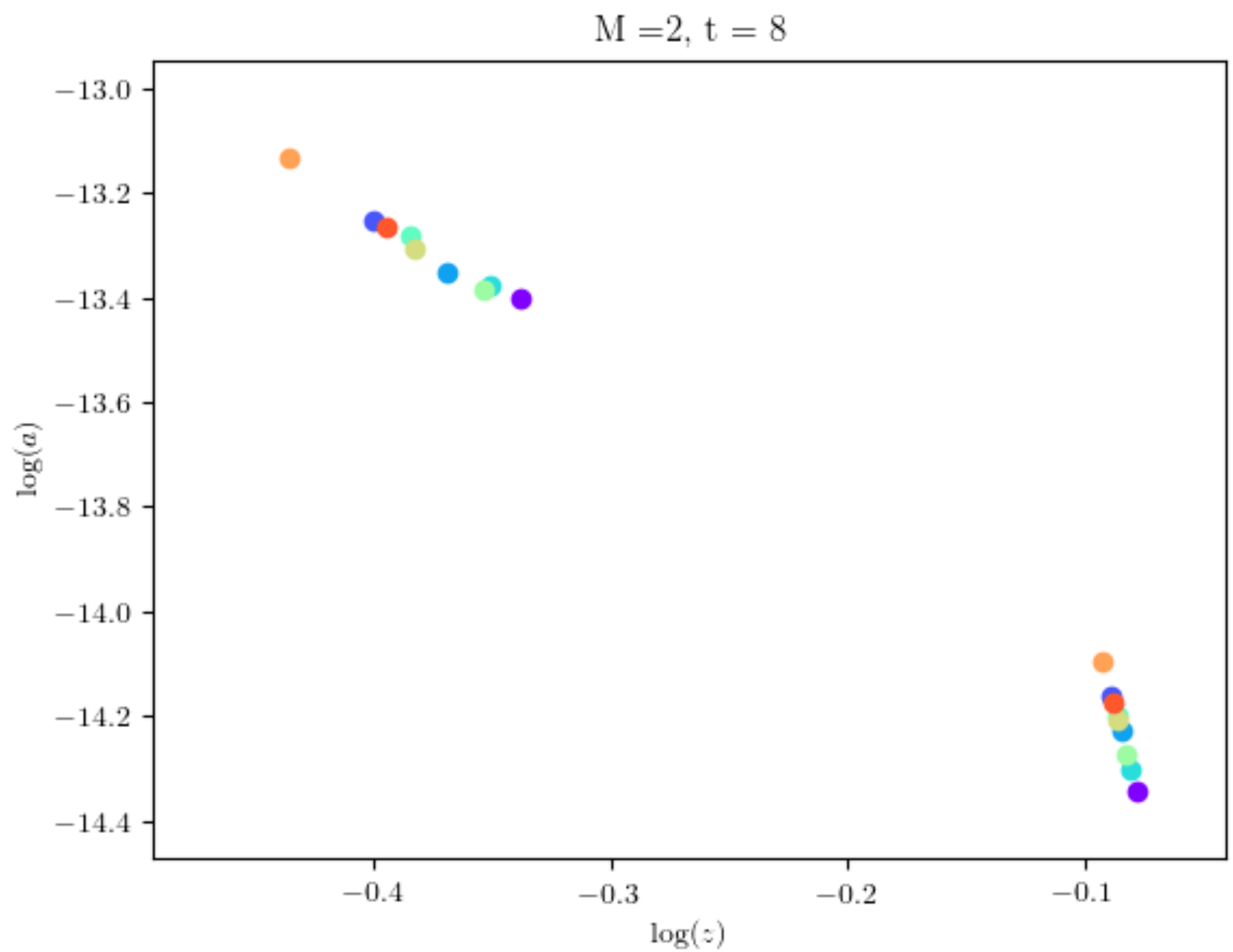
From theory to practice: the bootstrap



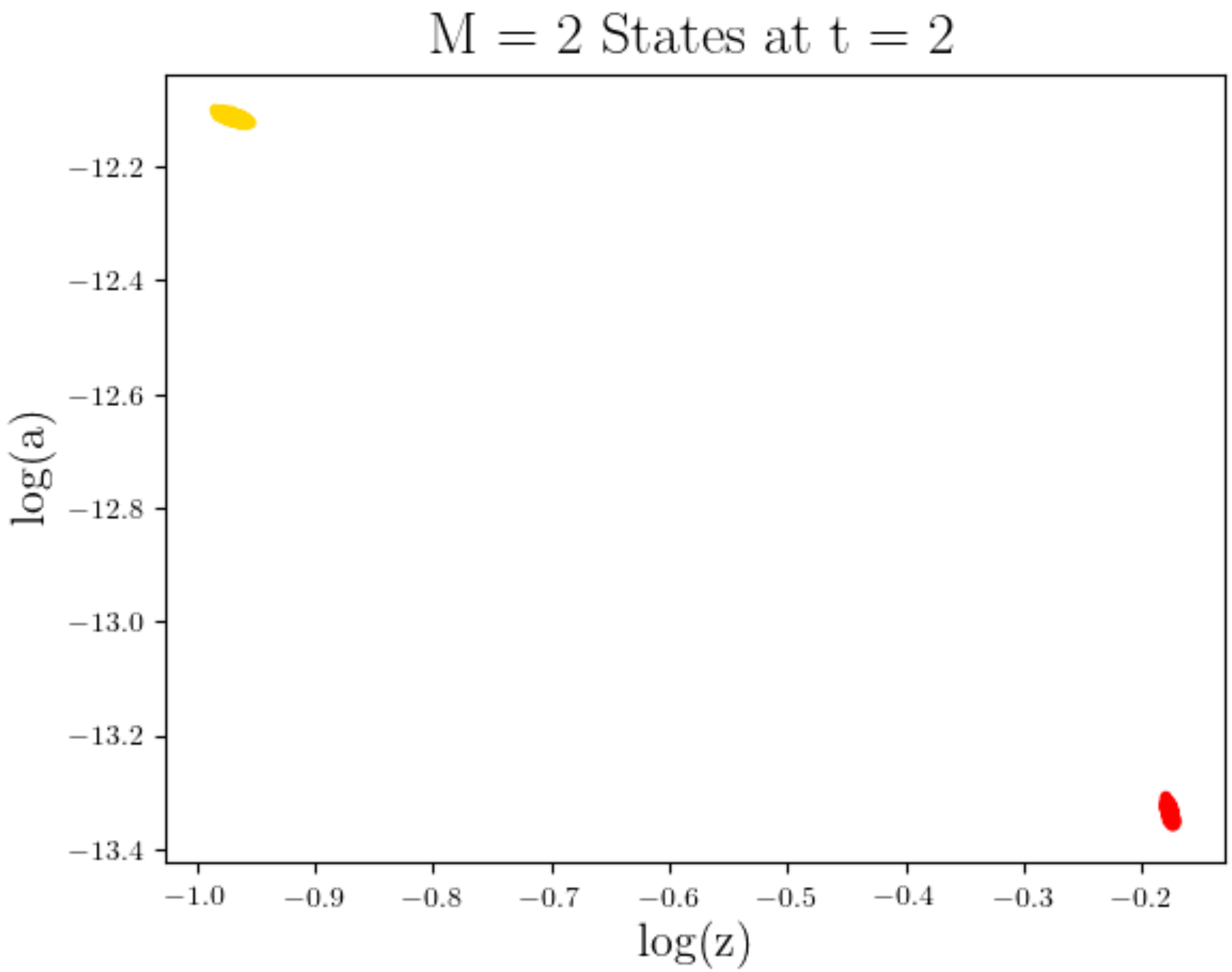
From theory to practice: the bootstrap



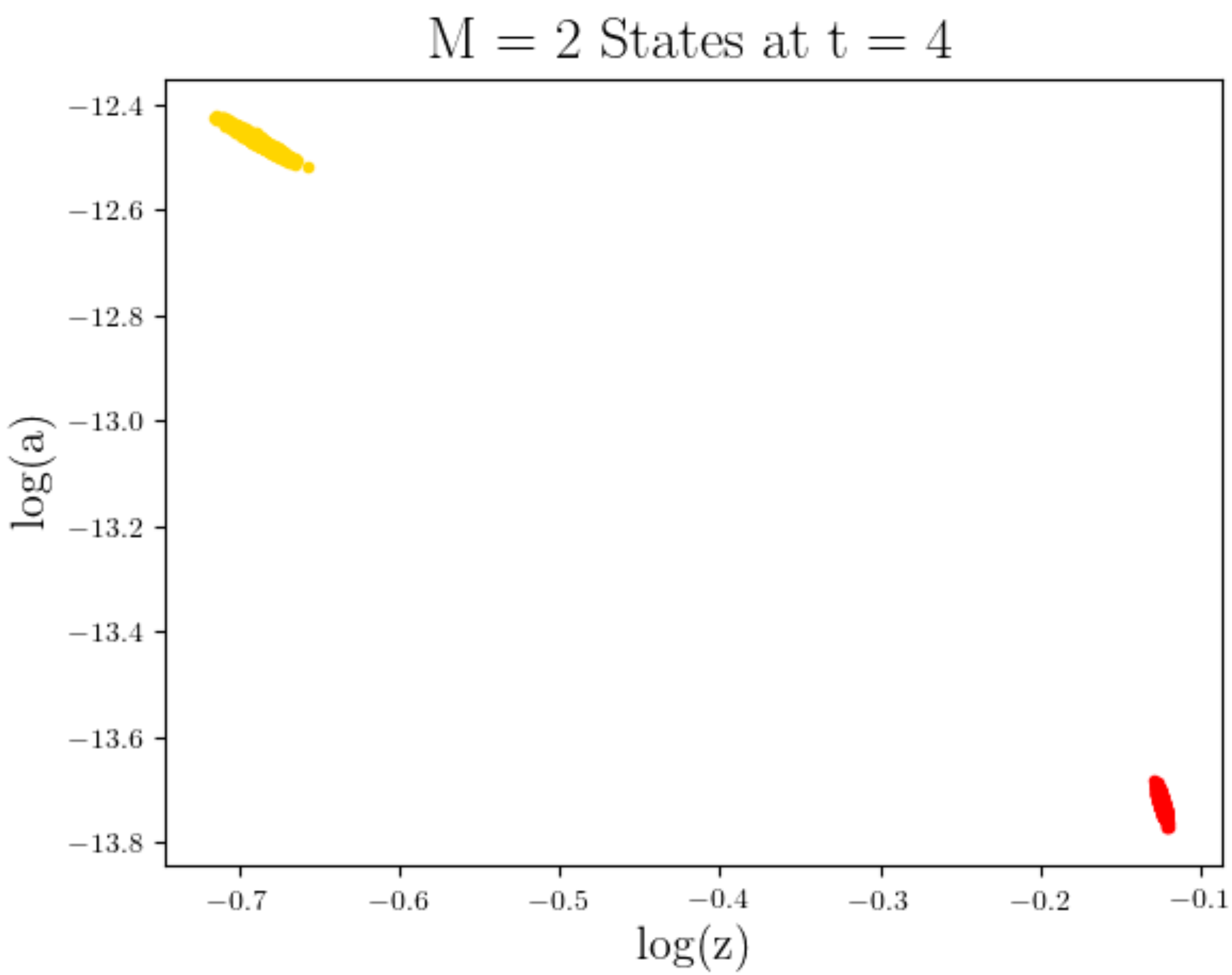
From theory to practice: the bootstrap



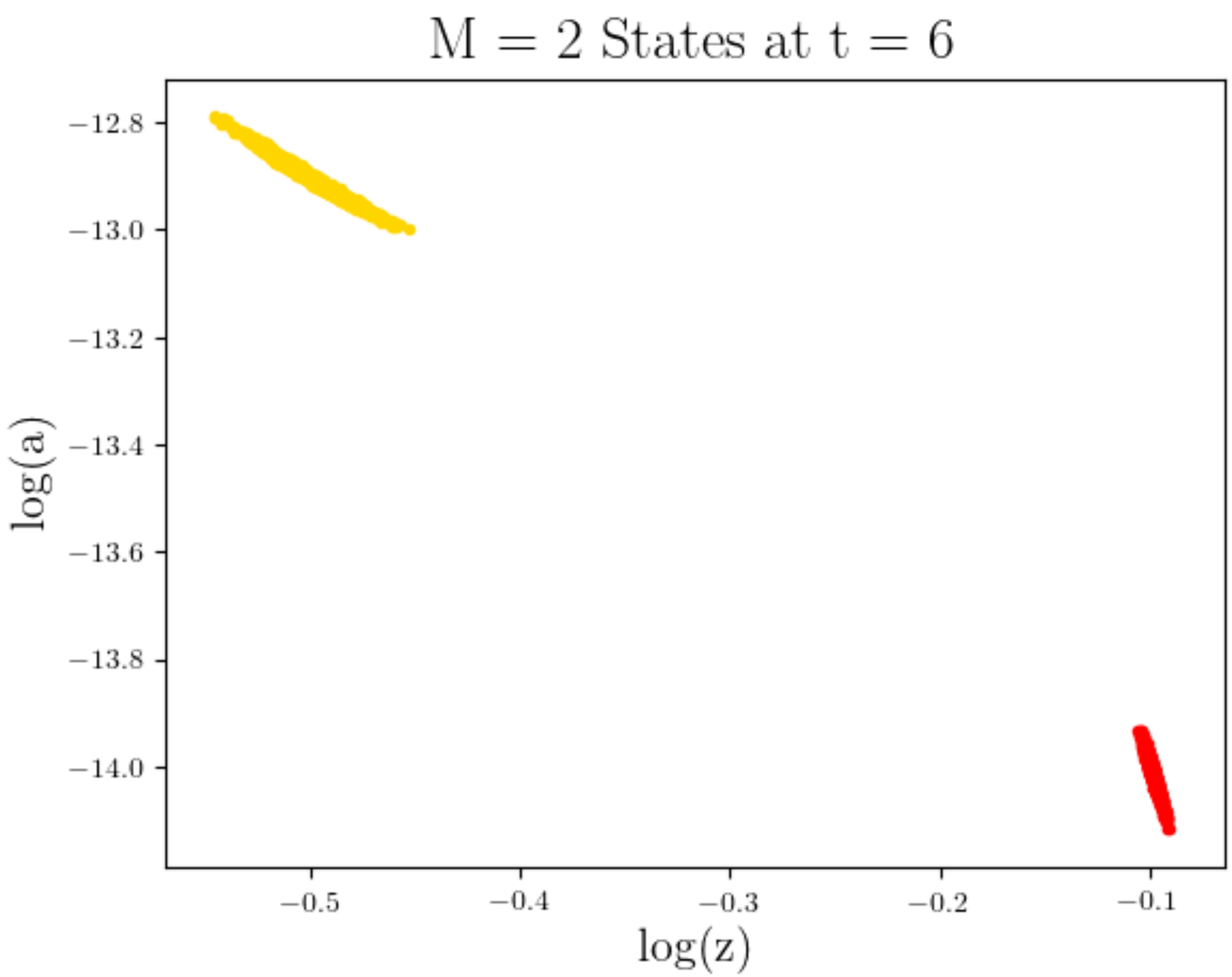
M = 2 State Extraction



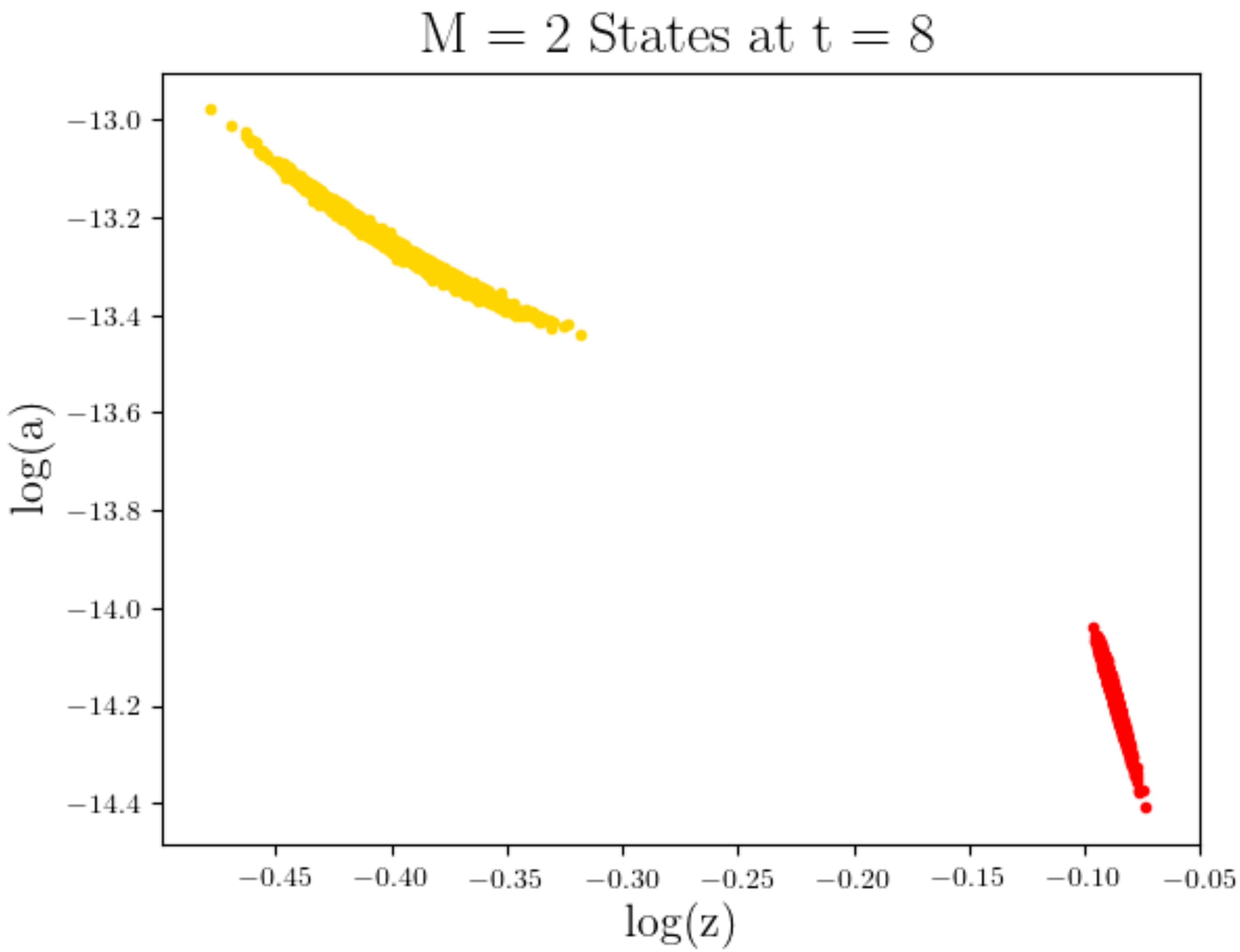
M = 2 State Extraction



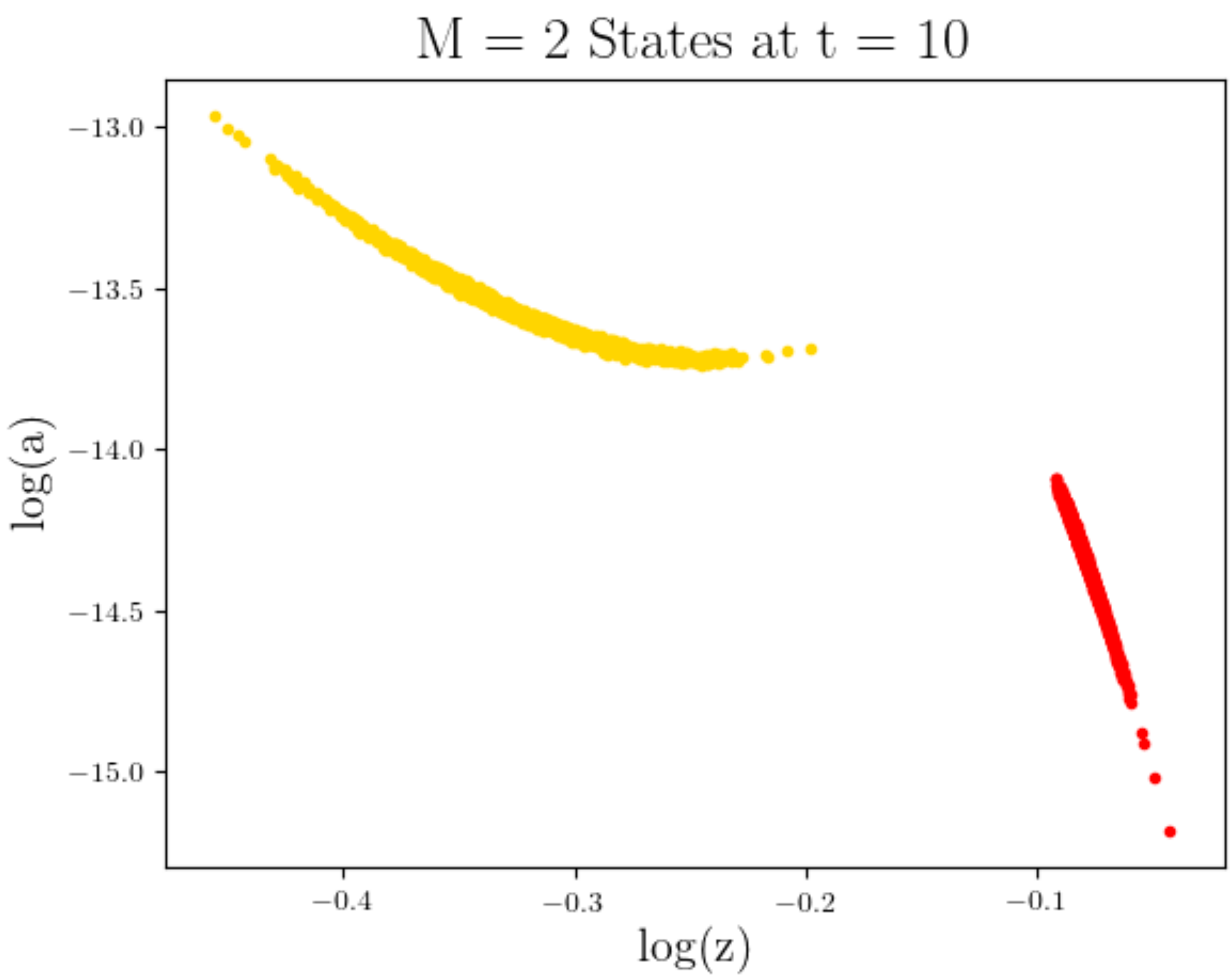
M = 2 State Extraction



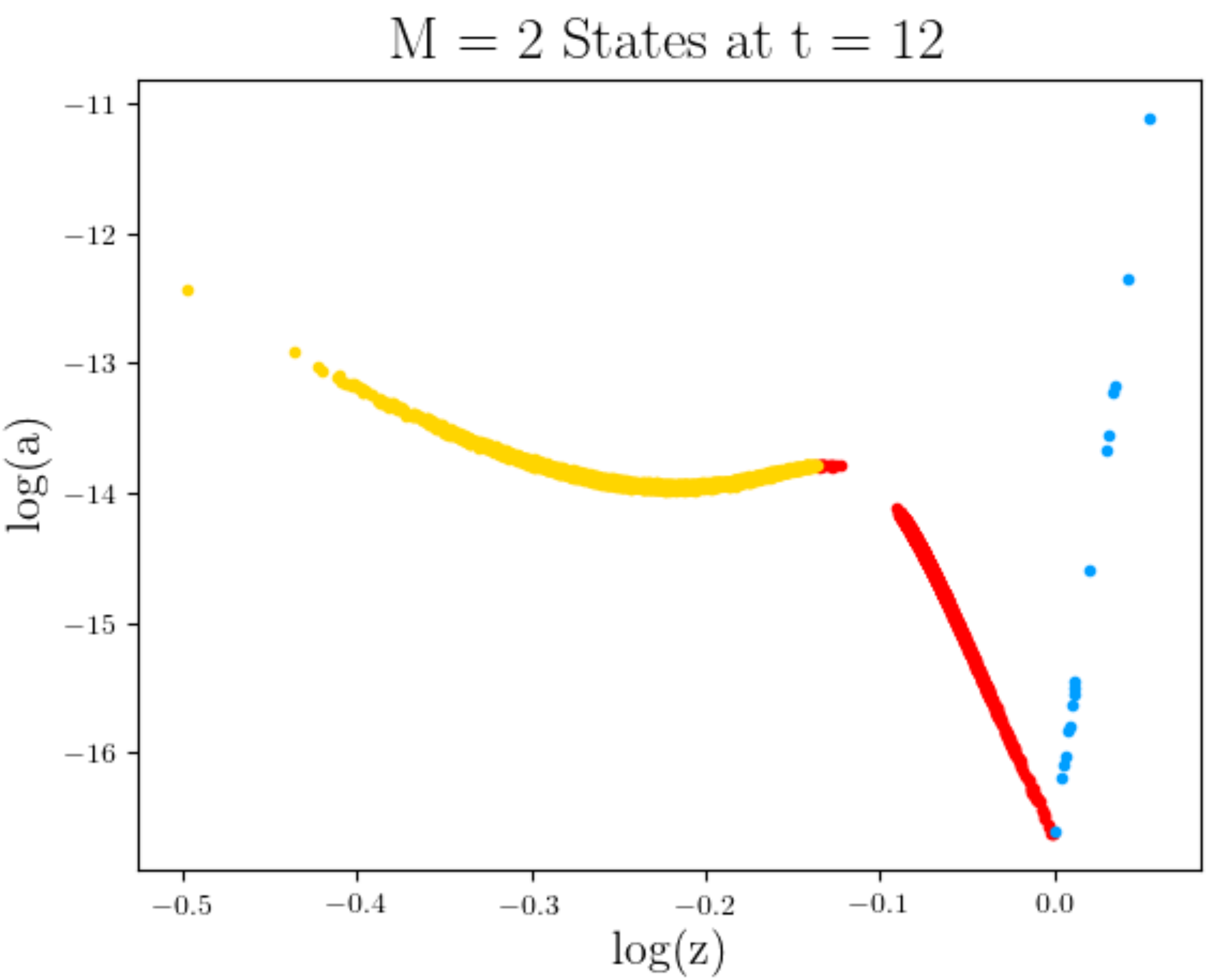
M = 2 State Extraction



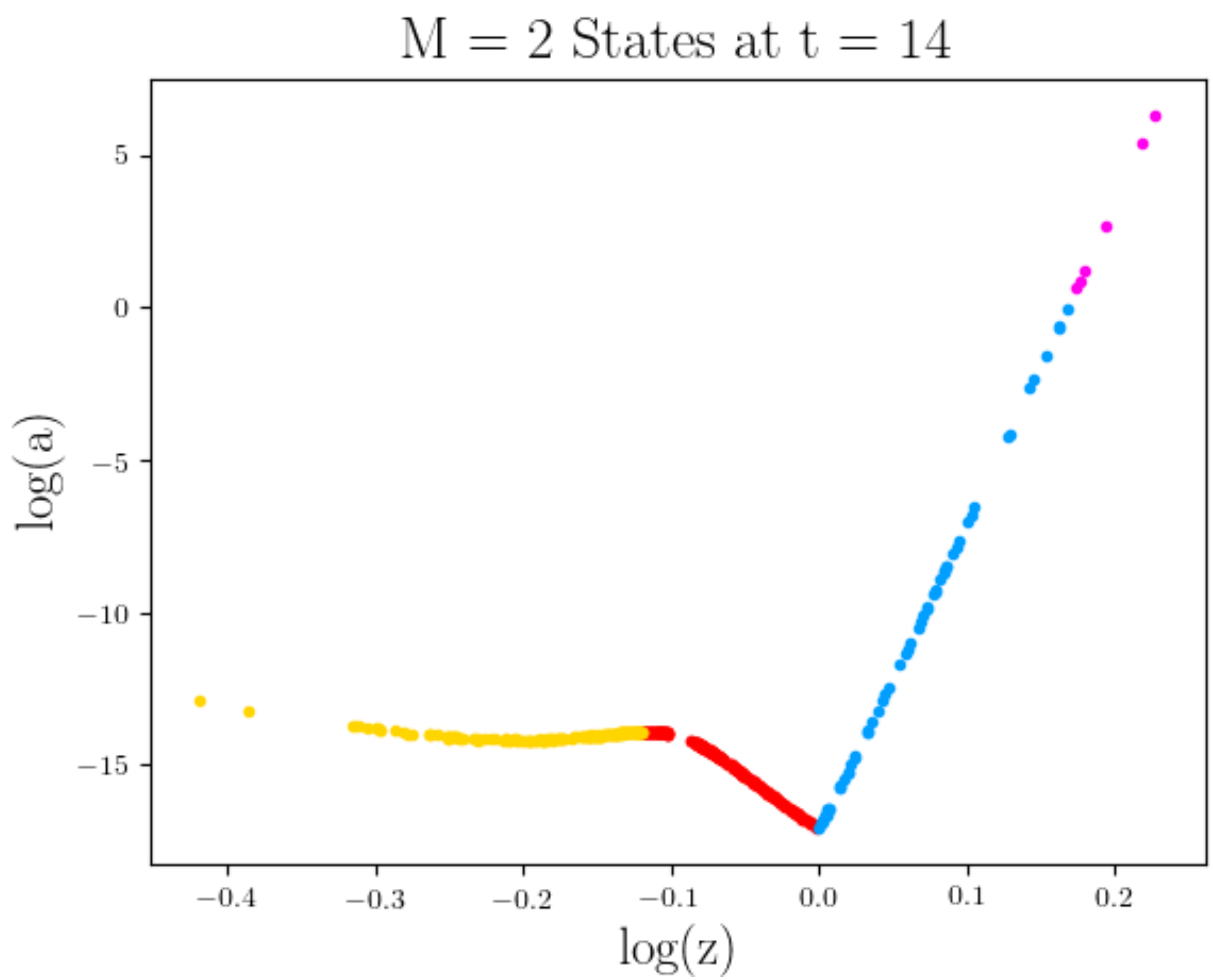
M = 2 State Extraction



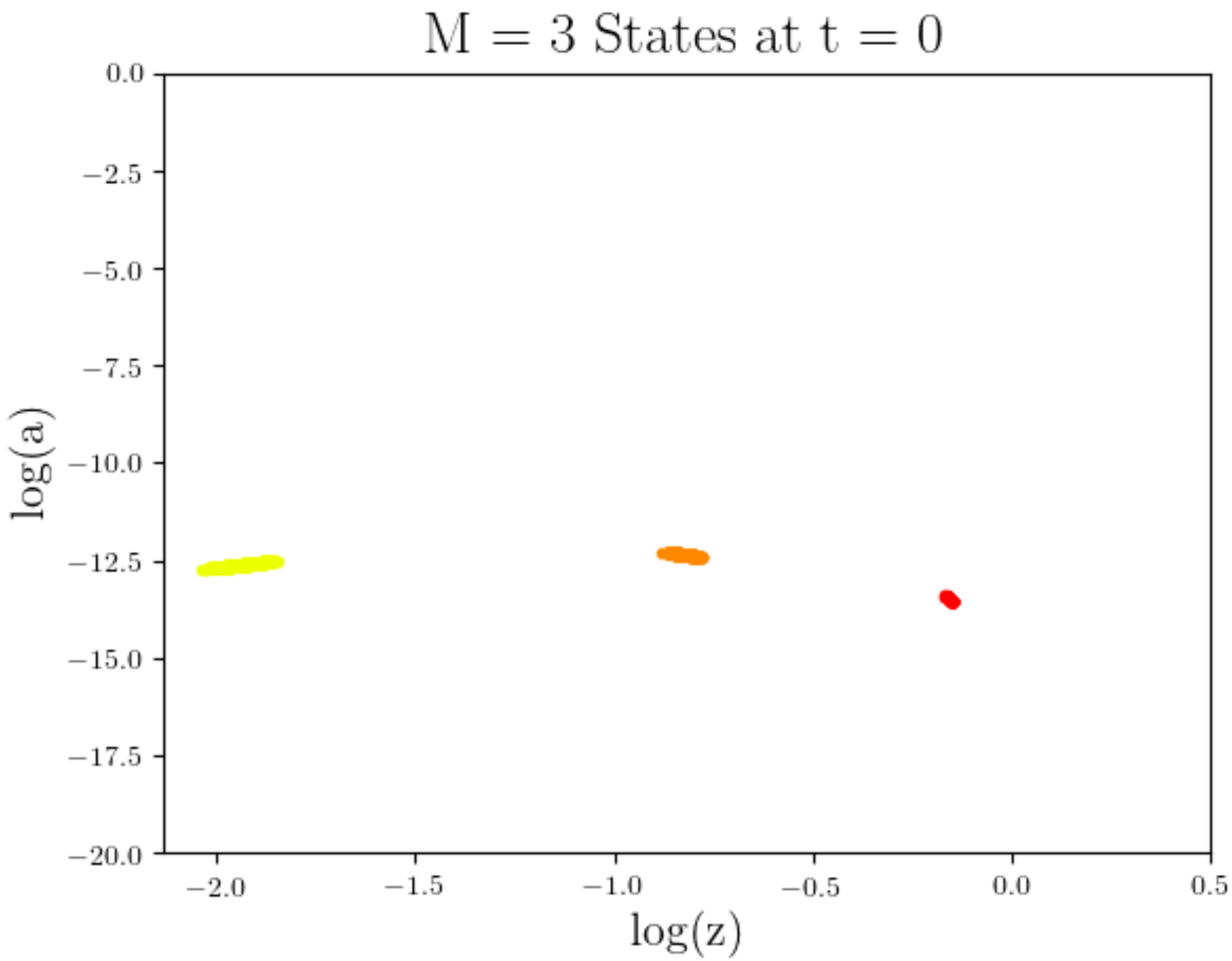
M = 2 State Extraction



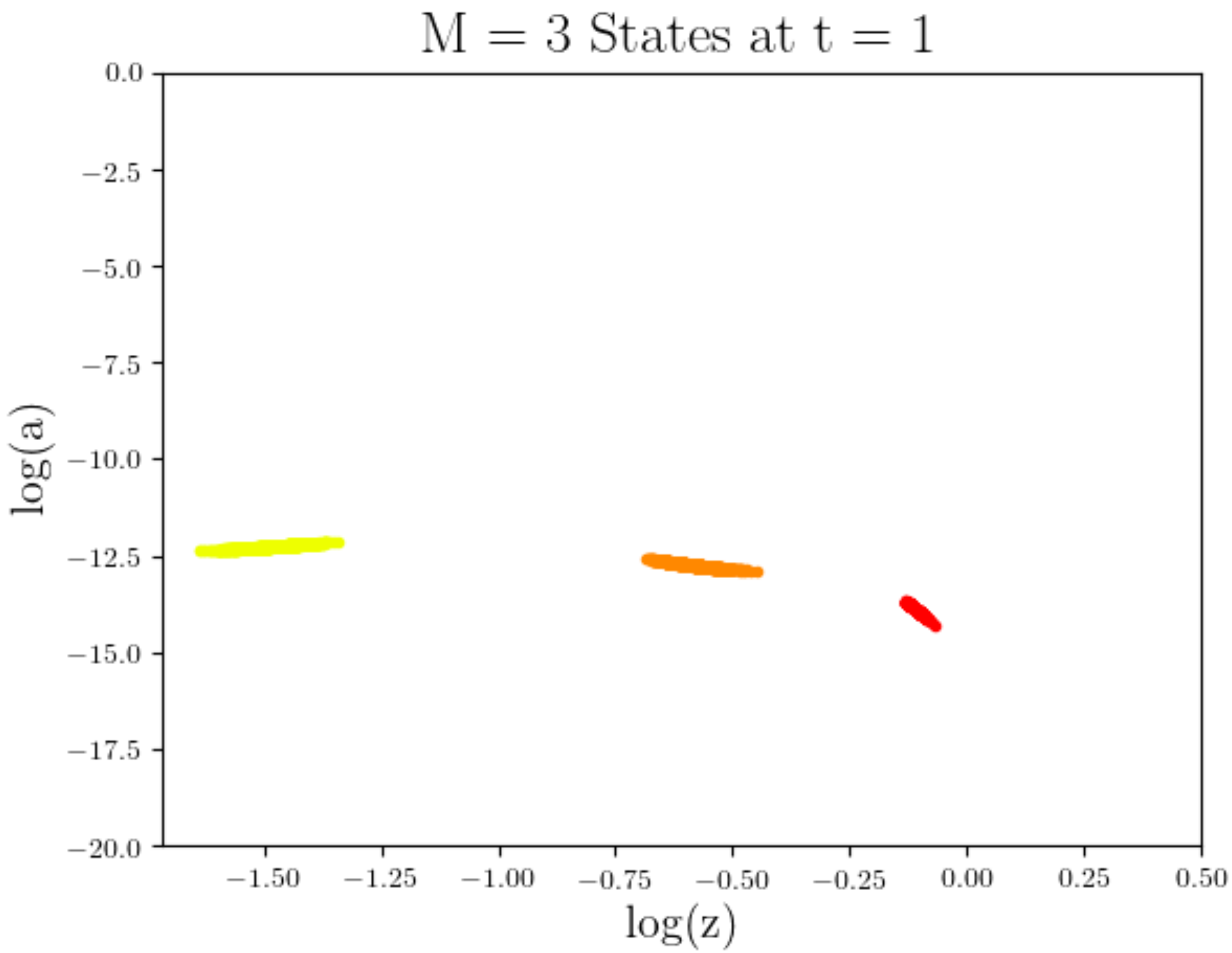
M = 2 State Extraction



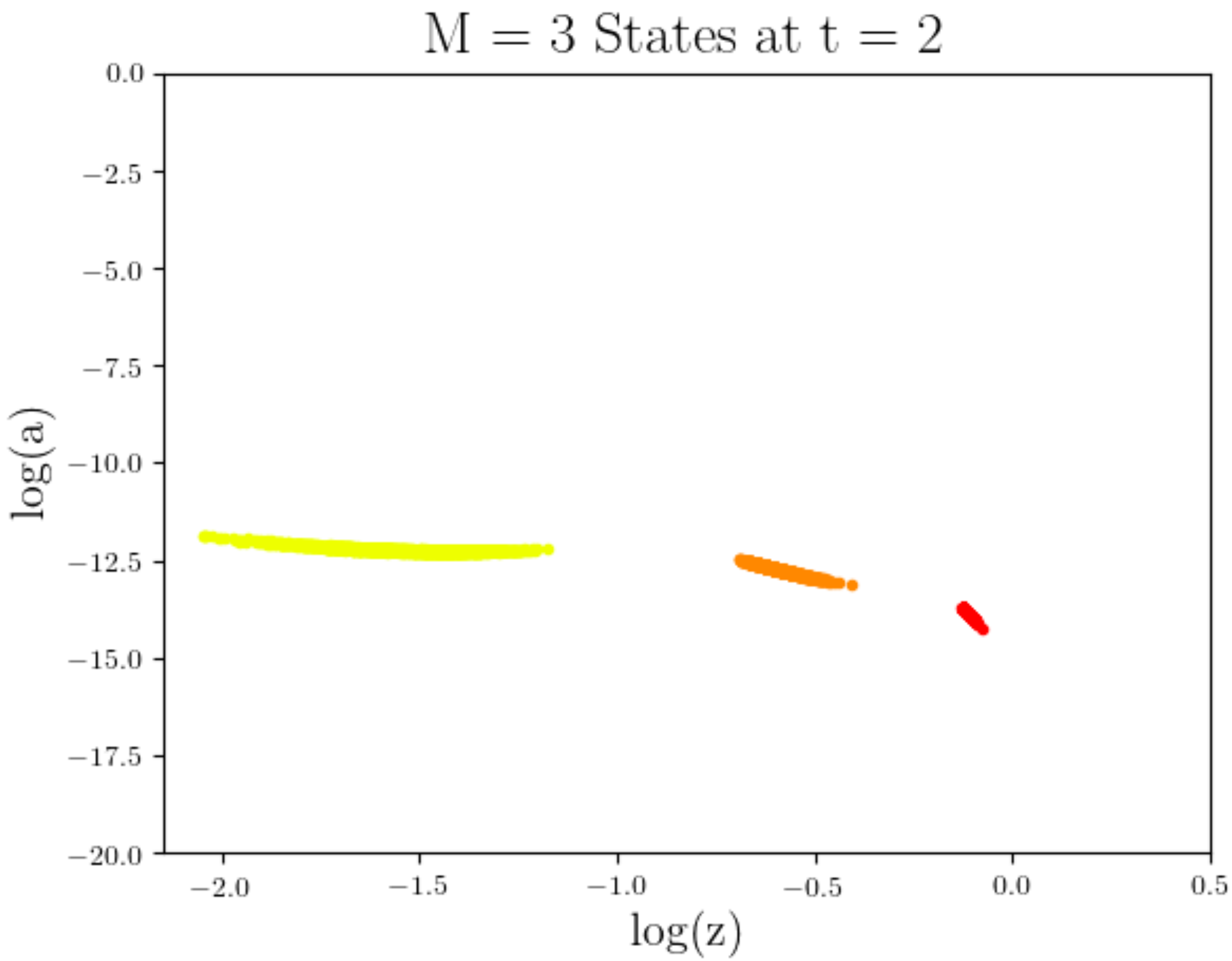
M = 3 State Extraction



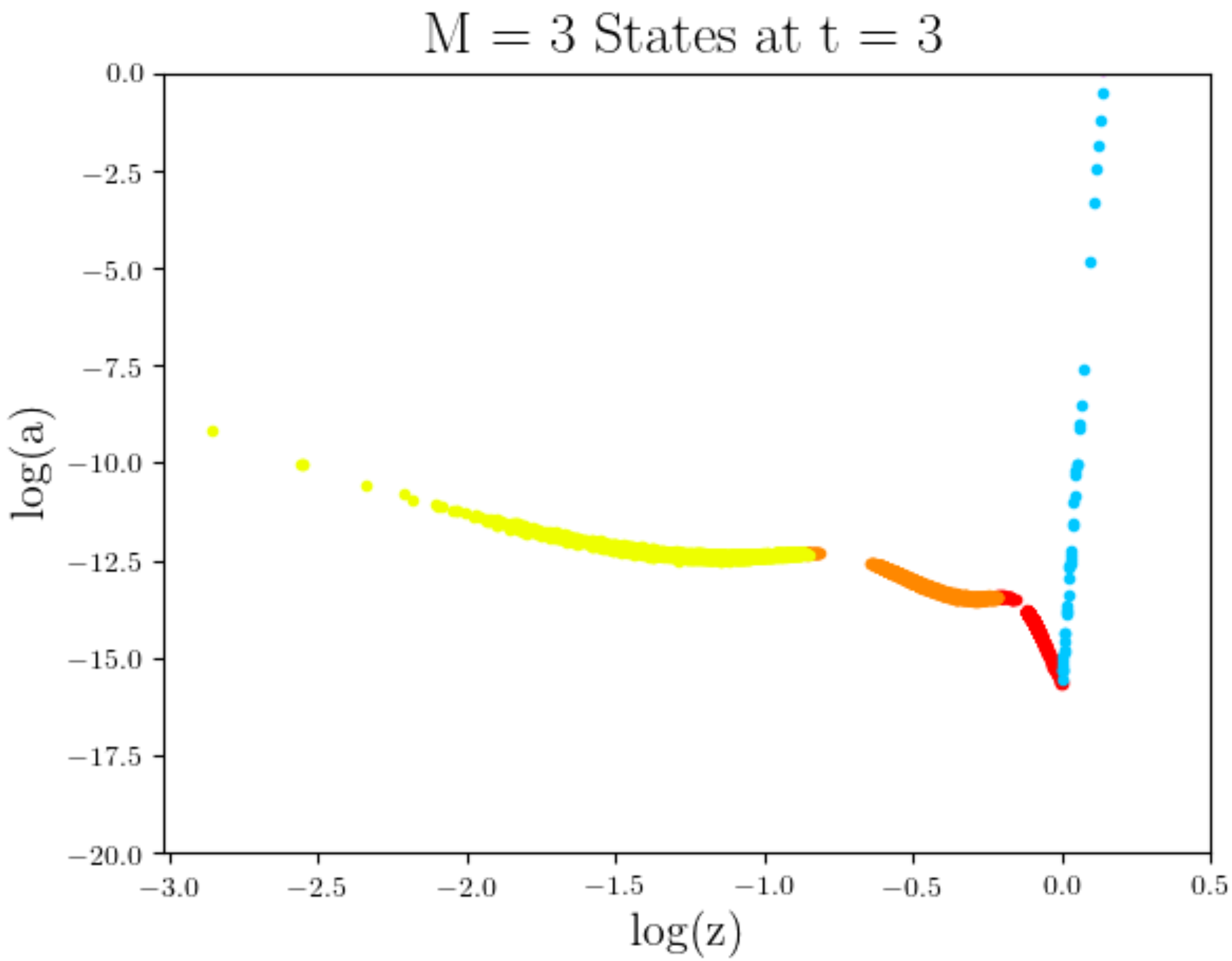
M = 3 State Extraction



M = 3 State Extraction



M = 3 State Extraction



Expectation Maximization Clustering

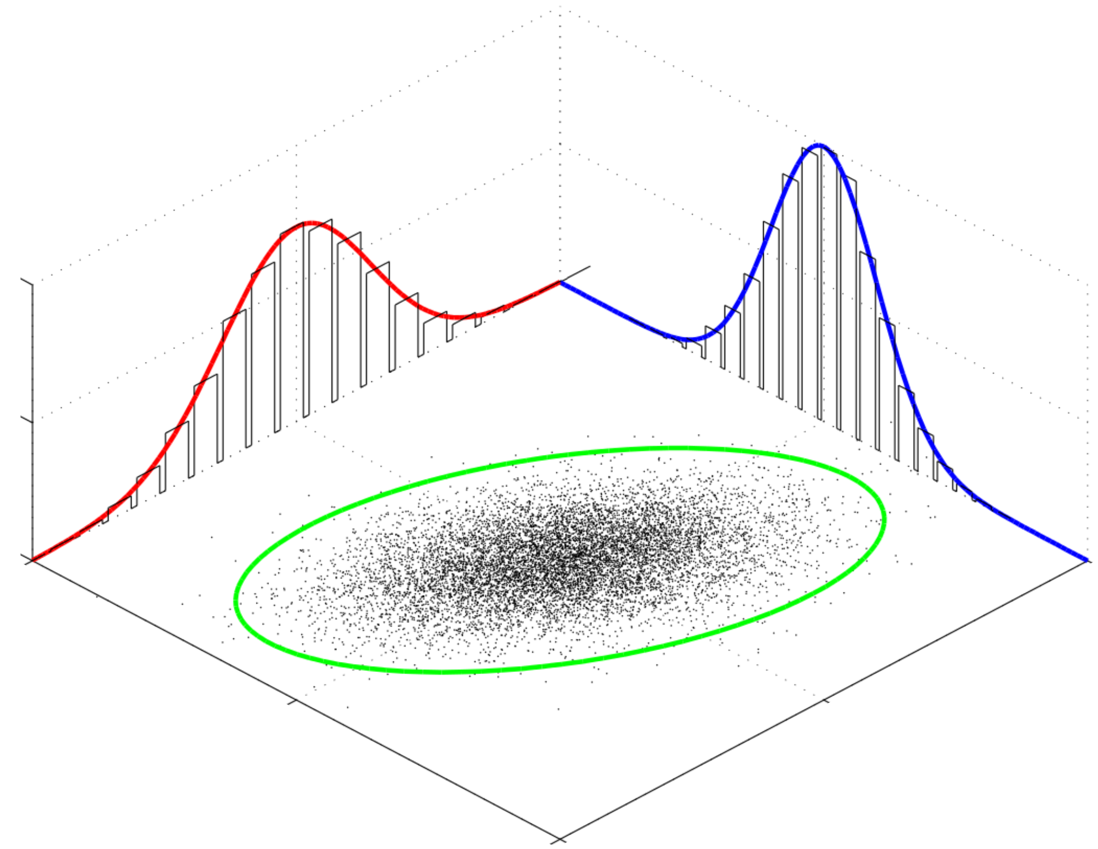
Dempster, A.P.; Laird, N.M.; Rubin, D.B. (1977). "Maximum Likelihood from Incomplete Data via the EM Algorithm". Journal of the Royal Statistical Society, Series B. 39 (1): 1–38.

$$\vec{\mu} = \left\langle \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle$$

$$\Sigma_{xx} = \sigma_x^2 = \langle (x - \mu_x)^2 \rangle$$

$$\Sigma_{yy} = \sigma_y^2 = \langle (y - \mu_y)^2 \rangle$$

$$\Sigma_{xy} = \Sigma_{yx}^{\dagger} = \langle (x - \mu_x)(y - \mu_y) \rangle$$



$$\boxed{p(\vec{x}) = \frac{1}{2\pi\sqrt{|\Sigma|}} \mathrm{e}^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})}}$$

K-Means Algorithm

Lloyd., S. P. (1982). "Least squares quantization in PCM"(PDF).
IEEE Transactions on Information Theory. 28 (2): 129–137.

1. Assign points to initial clusters
2. Compute the mean and covariance matrix for each cluster
3. For each bootstrap sample, find most probably permutation of points among clusters
4. Repeat 2. and 3. until the process converges

K-Means Algorithm

Lloyd., S. P. (1982). "Least squares quantization in PCM"(PDF).
IEEE Transactions on Information Theory. 28 (2): 129–137.

1. Assign points to initial clusters
2. Compute the mean and covariance matrix for each cluster
3. For each bootstrap sample, find most probably permutation of points among clusters
4. Repeat 2. and 3. until the process converges

K-Means Algorithm

Lloyd., S. P. (1982). "Least squares quantization in PCM"(PDF).
IEEE Transactions on Information Theory. 28 (2): 129–137.

1. Assign points to initial clusters
2. Compute the mean and covariance matrix for each cluster
3. For each bootstrap sample, find most probably permutation of points among clusters
4. Repeat 2. and 3. until the process converges

K-Means Algorithm

Lloyd., S. P. (1982). "Least squares quantization in PCM"(PDF).
IEEE Transactions on Information Theory. 28 (2): 129–137.

1. Assign points to initial clusters
2. Compute the mean and covariance matrix for each cluster
3. For each bootstrap sample, find most probably permutation of points among clusters
4. Repeat 2. and 3. until the process converges

K-Means Algorithm

Lloyd., S. P. (1982). "Least squares quantization in PCM"(PDF).
IEEE Transactions on Information Theory. 28 (2): 129–137.

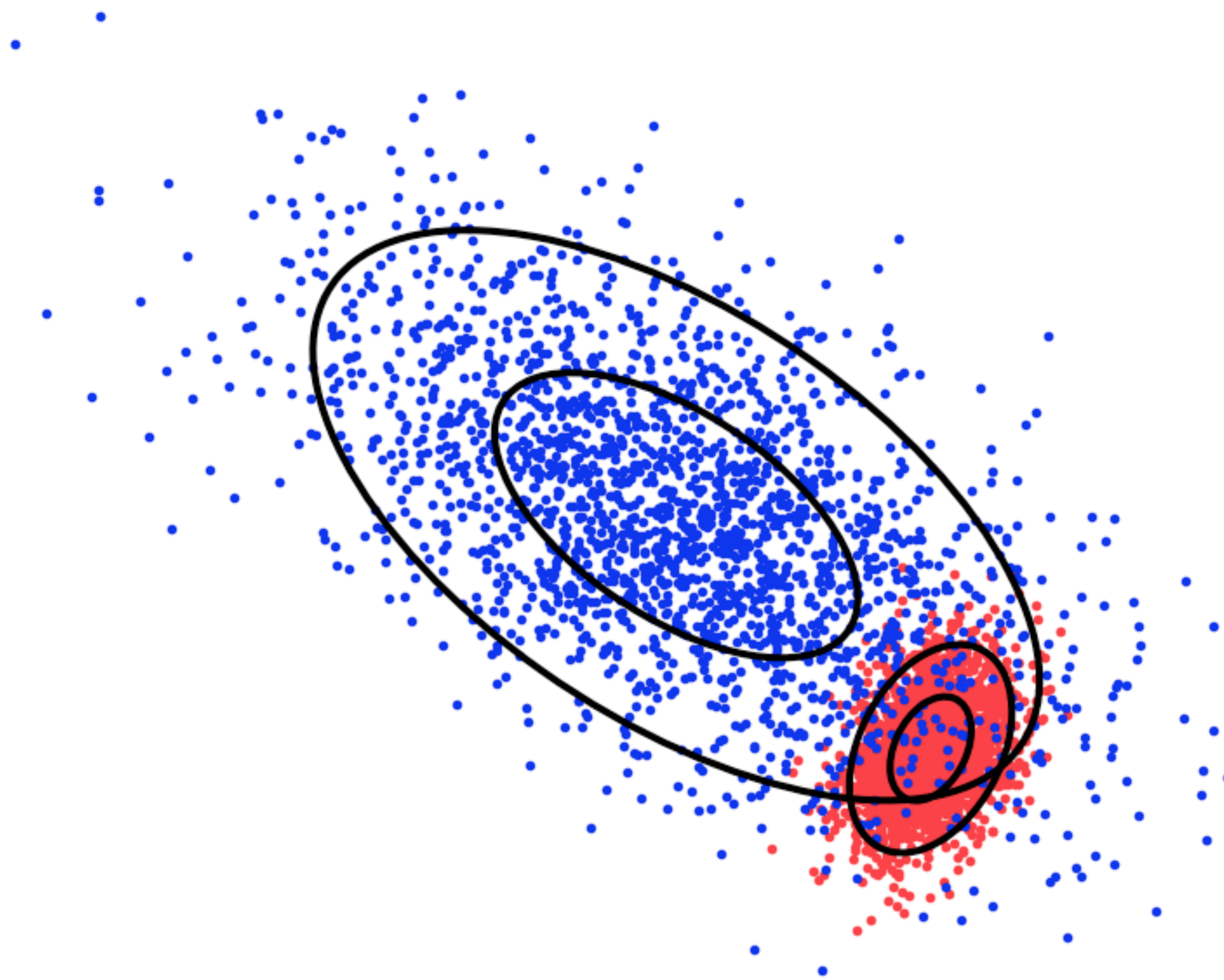
1. Assign points to initial clusters
2. Compute the mean and covariance matrix for each cluster
3. For each bootstrap sample, find most probably permutation of points among clusters
4. Repeat 2. and 3. until the process converges

$$p_1(\vec{x}) > p_2(\vec{x})$$

$$\Rightarrow \frac{1}{\sqrt{|\Sigma_1|}} \mathrm{e}^{-\frac{1}{2}(\vec{x}-\vec{\mu_1})^T \Sigma_1^{-1} (\vec{x}-\vec{\mu_1})} > \frac{1}{\sqrt{|\Sigma_2|}} \mathrm{e}^{-\frac{1}{2}(\vec{x}-\vec{\mu_2})^T \Sigma_2^{-1} (\vec{x}-\vec{\mu_2})},$$

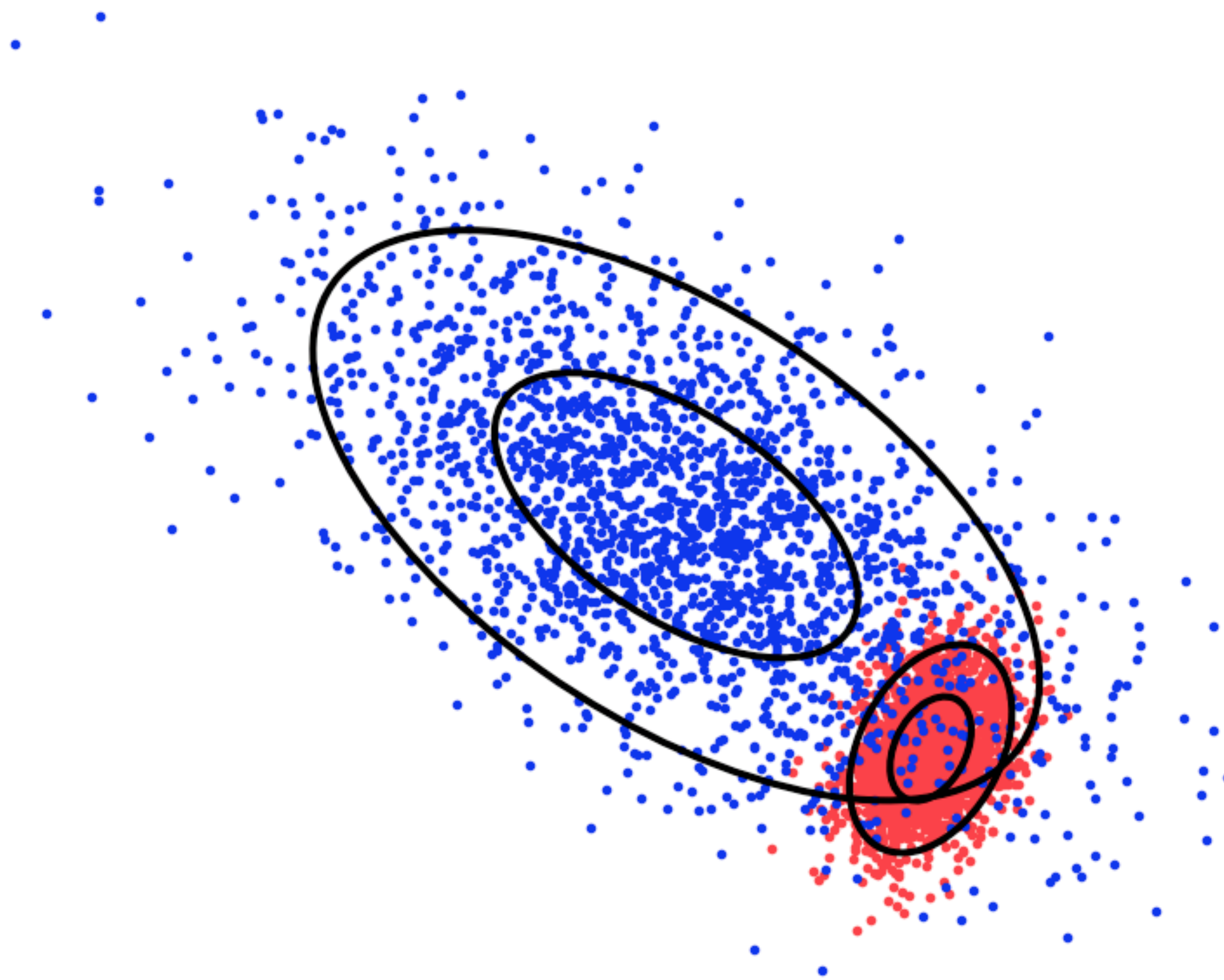
$$\boxed{\Rightarrow d_1(\vec{x}) + \log|\Sigma_1| < d_2(\vec{x}) + \log|\Sigma_2|}$$

K-Means Algorithm



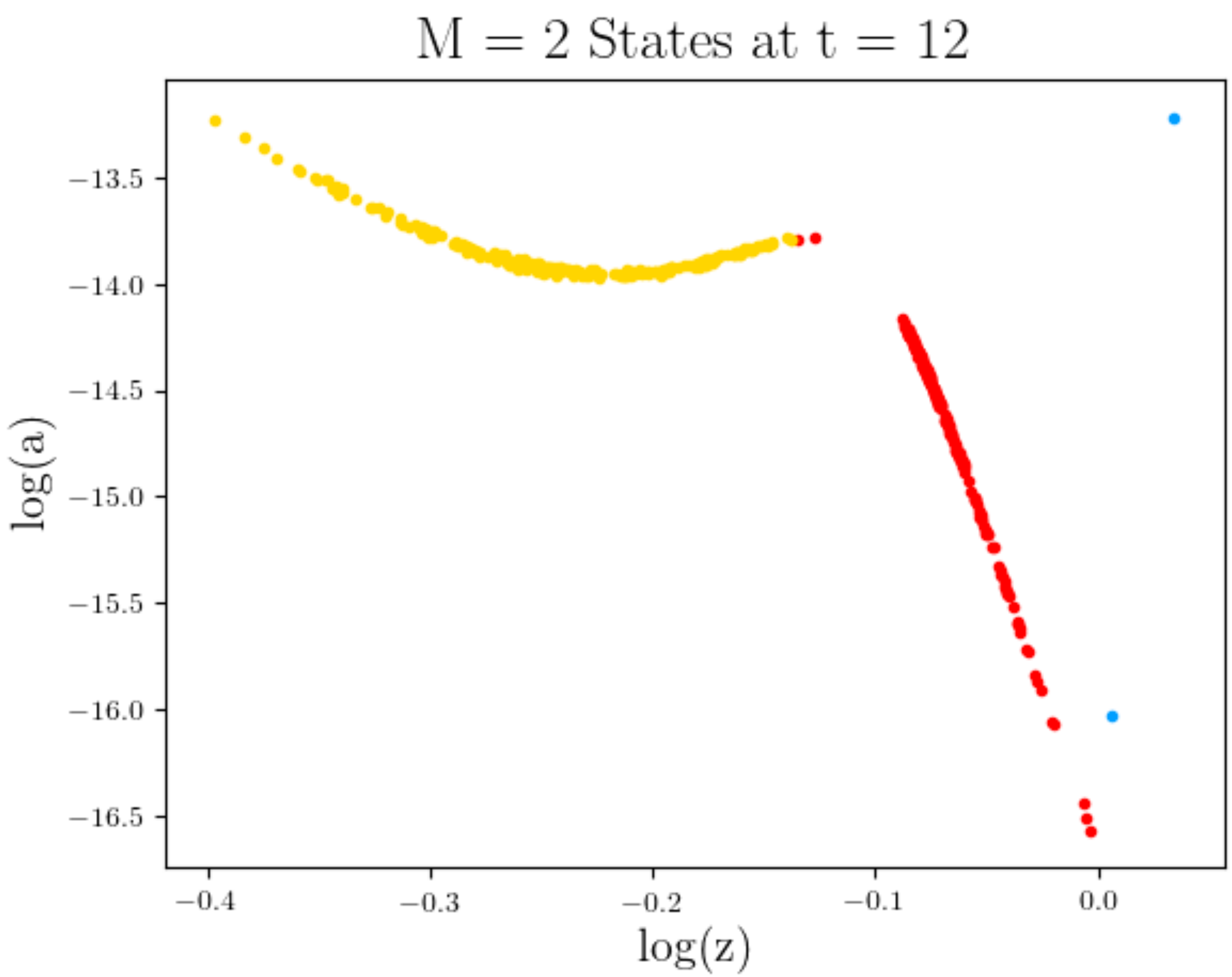
$$\boxed{\Rightarrow d_1(\vec{x}) + \log|\Sigma_1| < d_2(\vec{x}) + \log|\Sigma_2|}$$

K-Means Algorithm

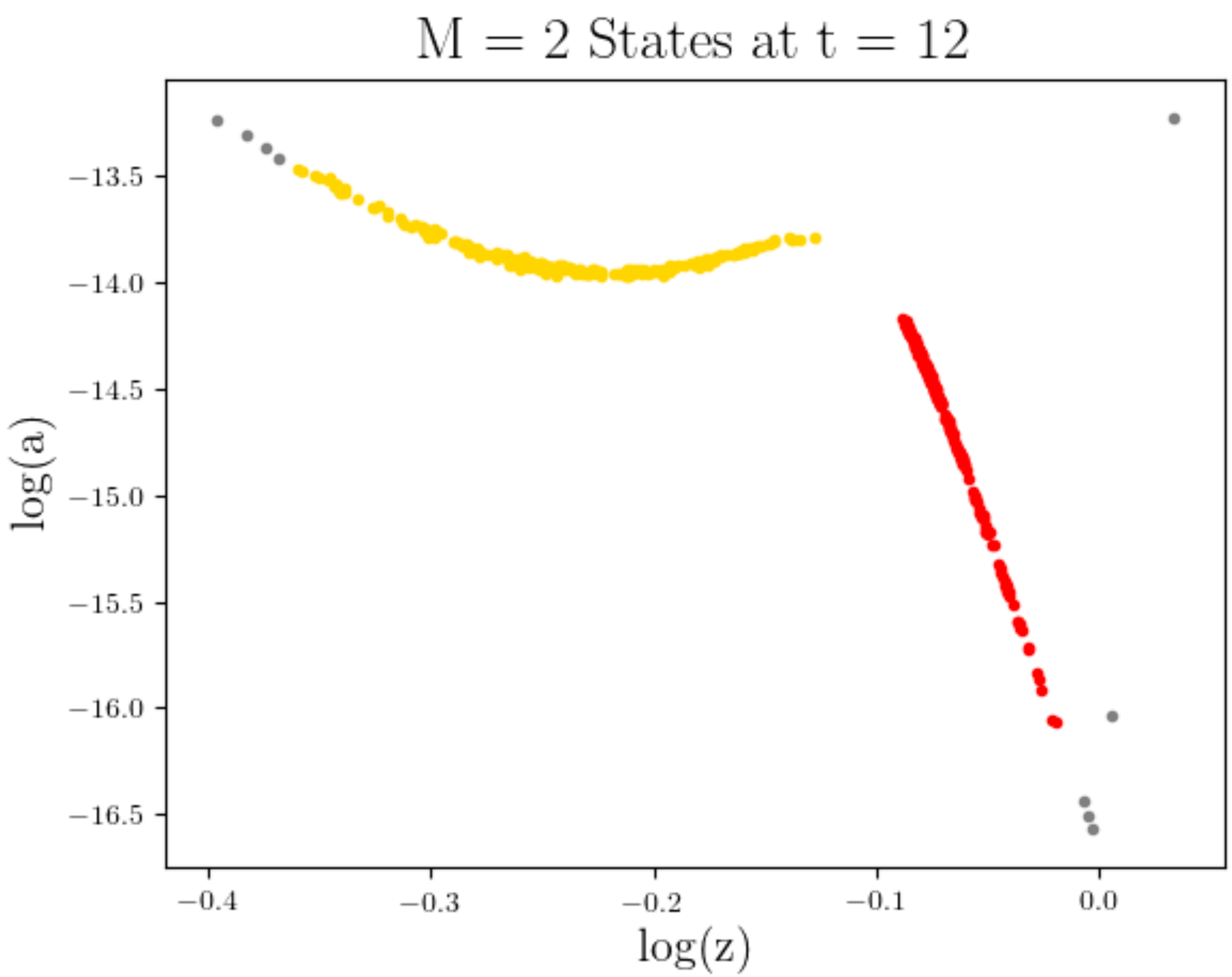


$$0 = p_0 + p_1 z + p_2 z^2 \qquad z_m = \mathrm{e}^{-E_m}$$

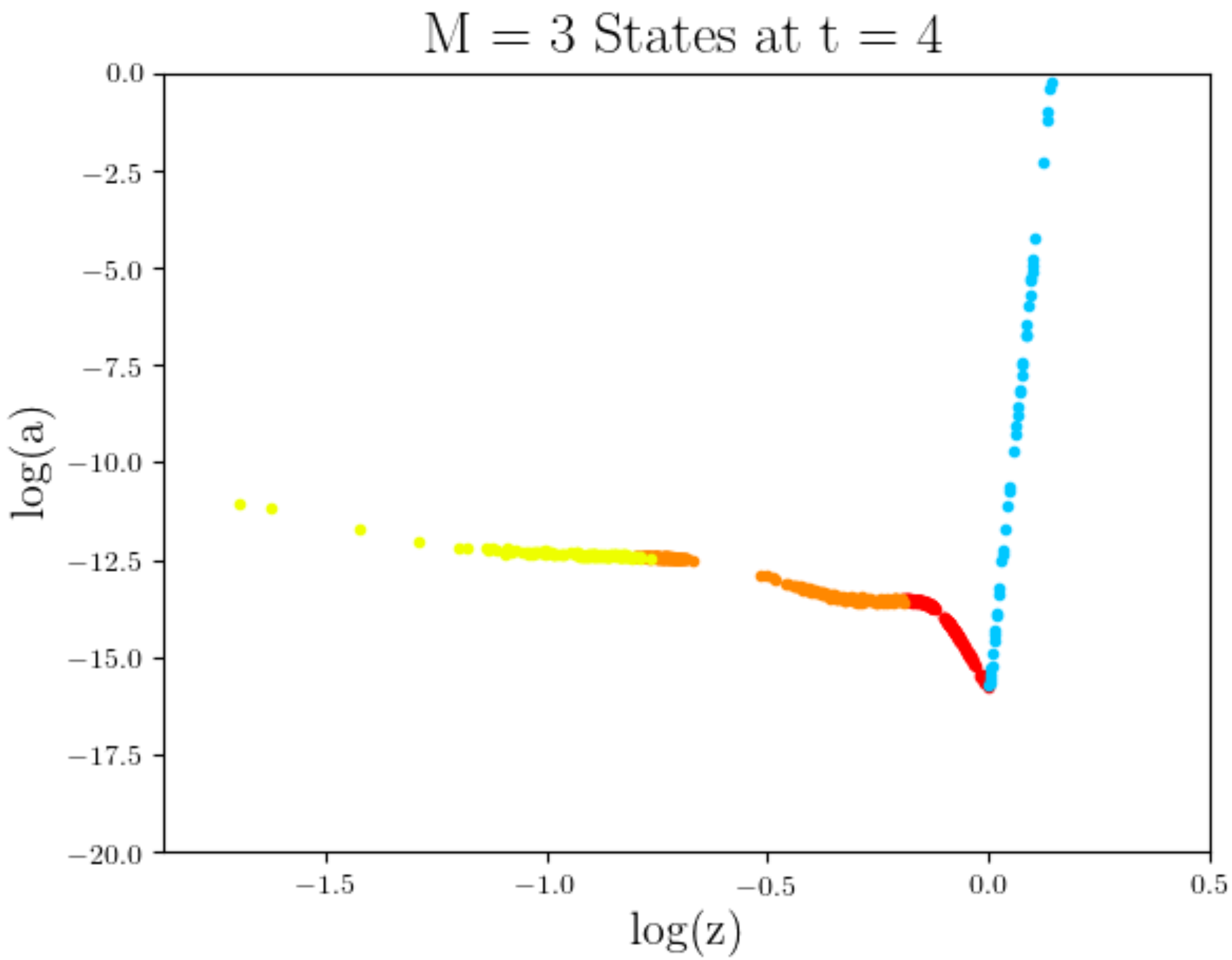
Clustering M = 2, Initial



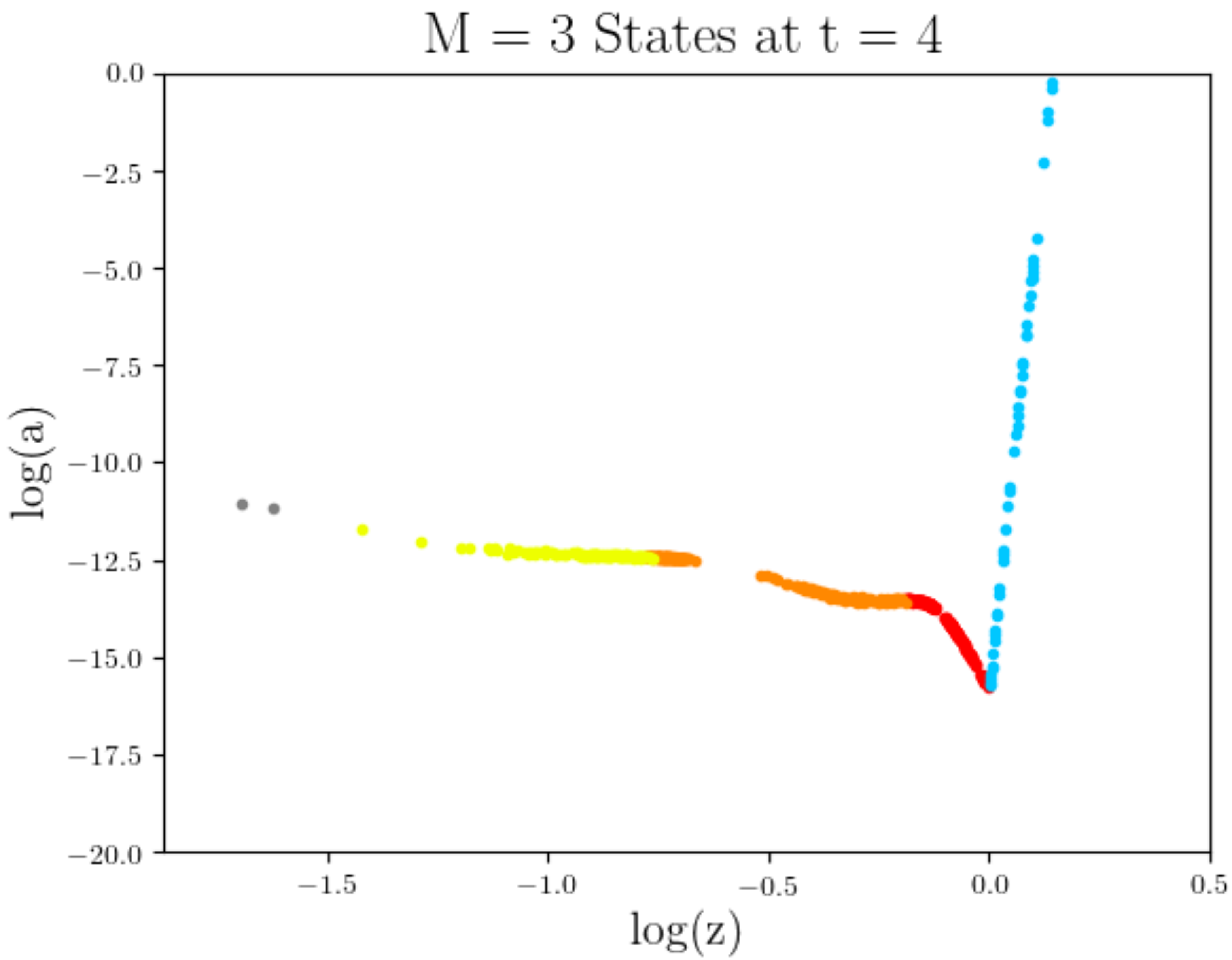
Clustering M = 2, 1 iteration



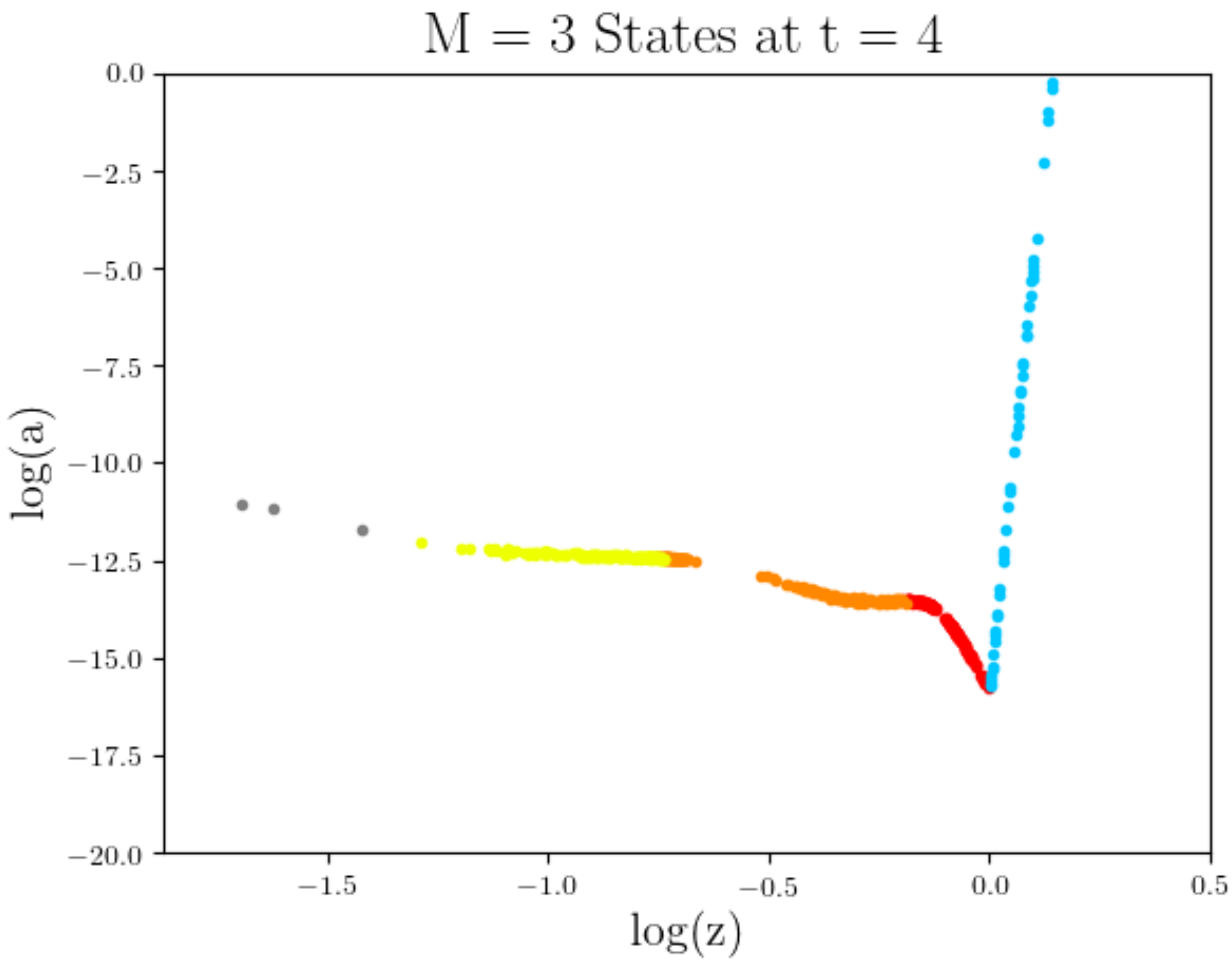
Clustering M = 3, Initial



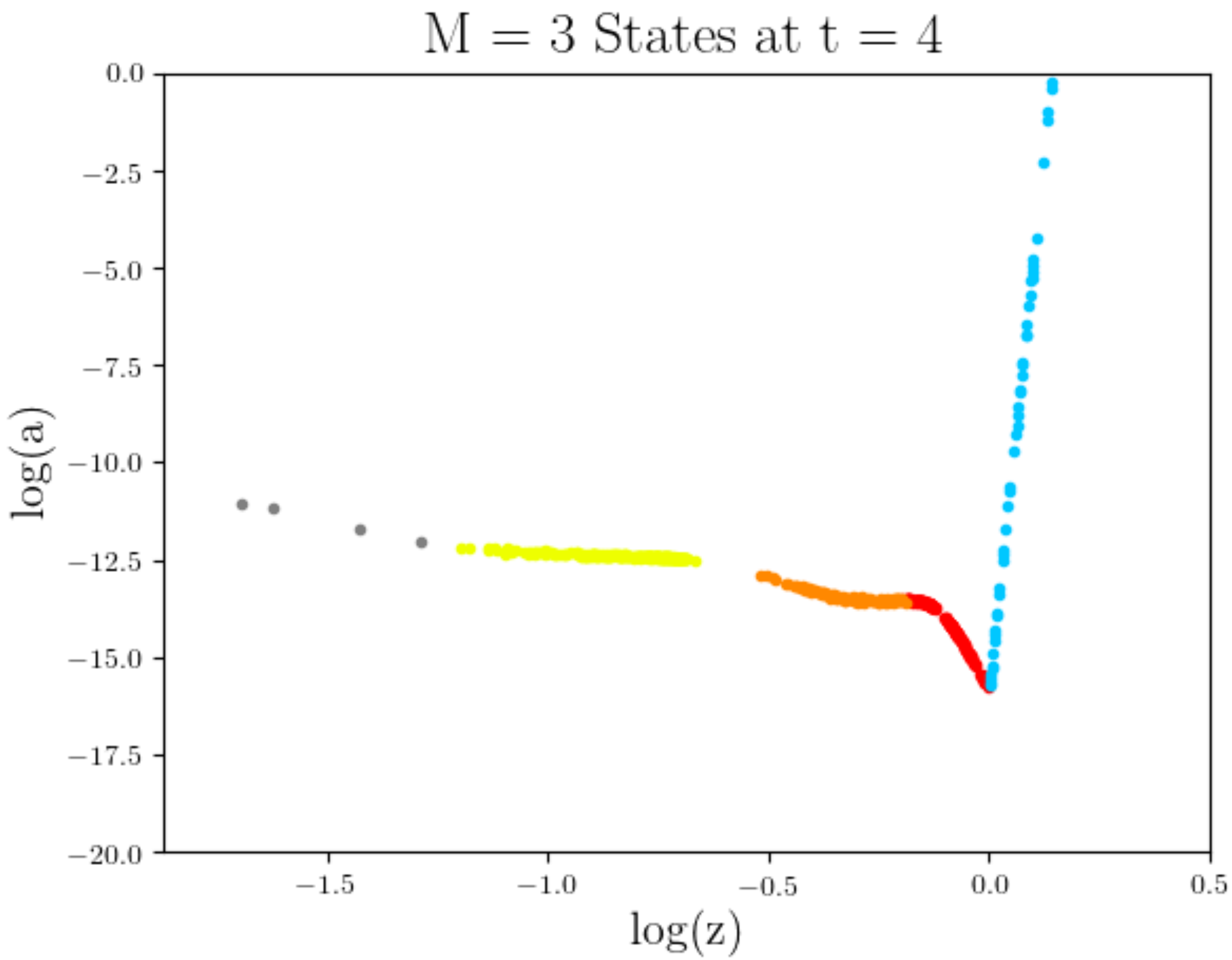
Clustering M = 3, 1 iteration



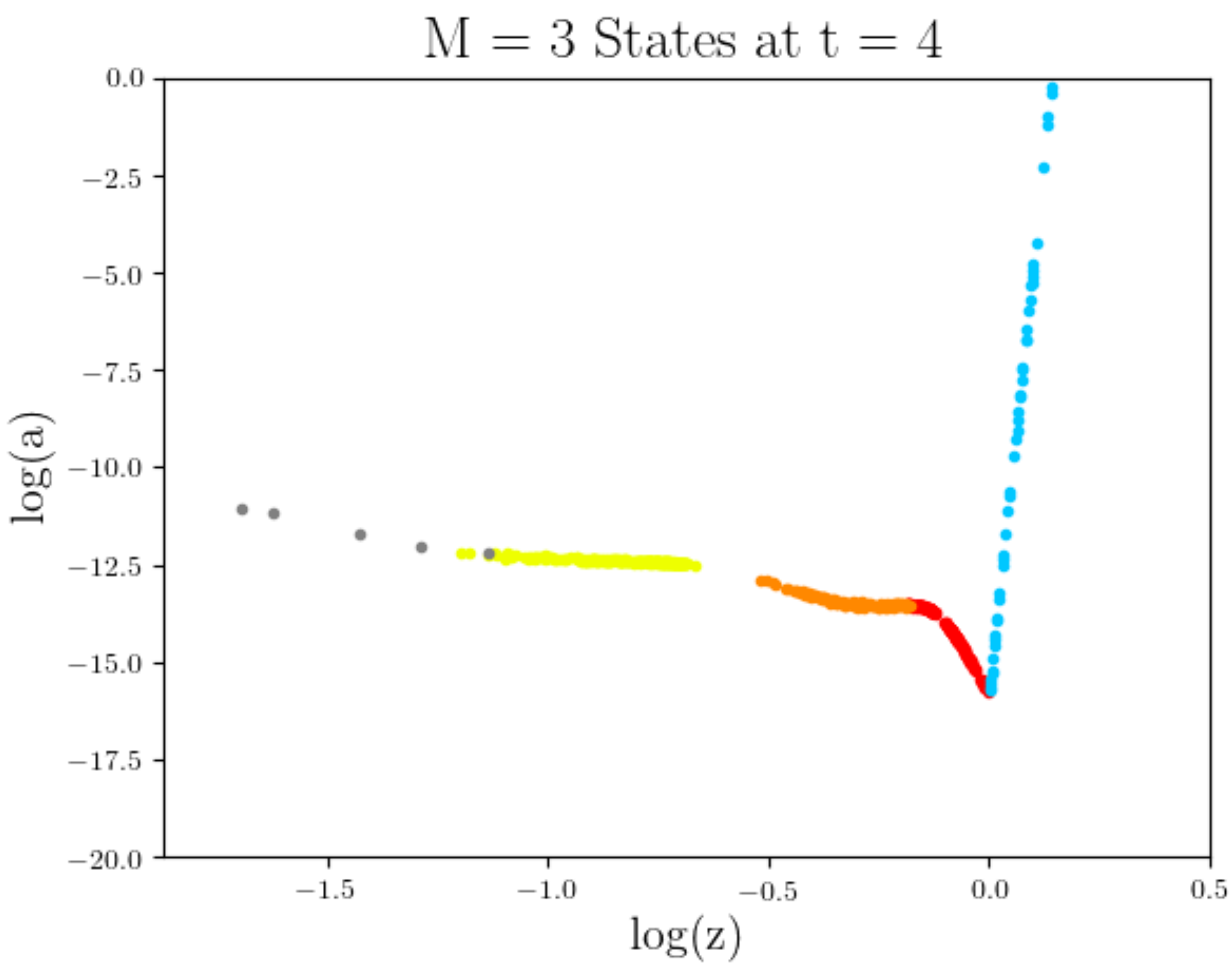
Clustering M = 3, 2 iterations



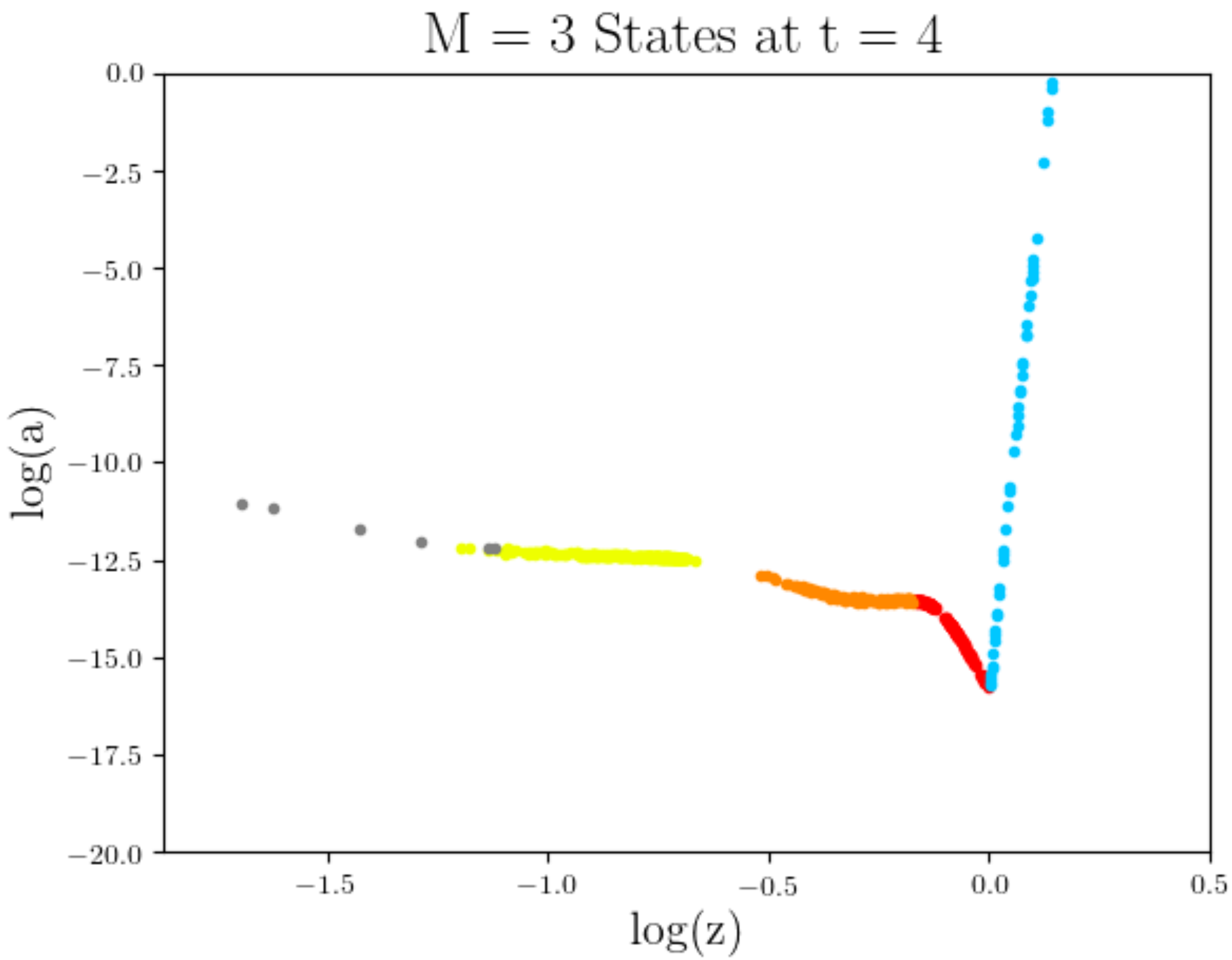
Clustering M = 3, 3 iterations



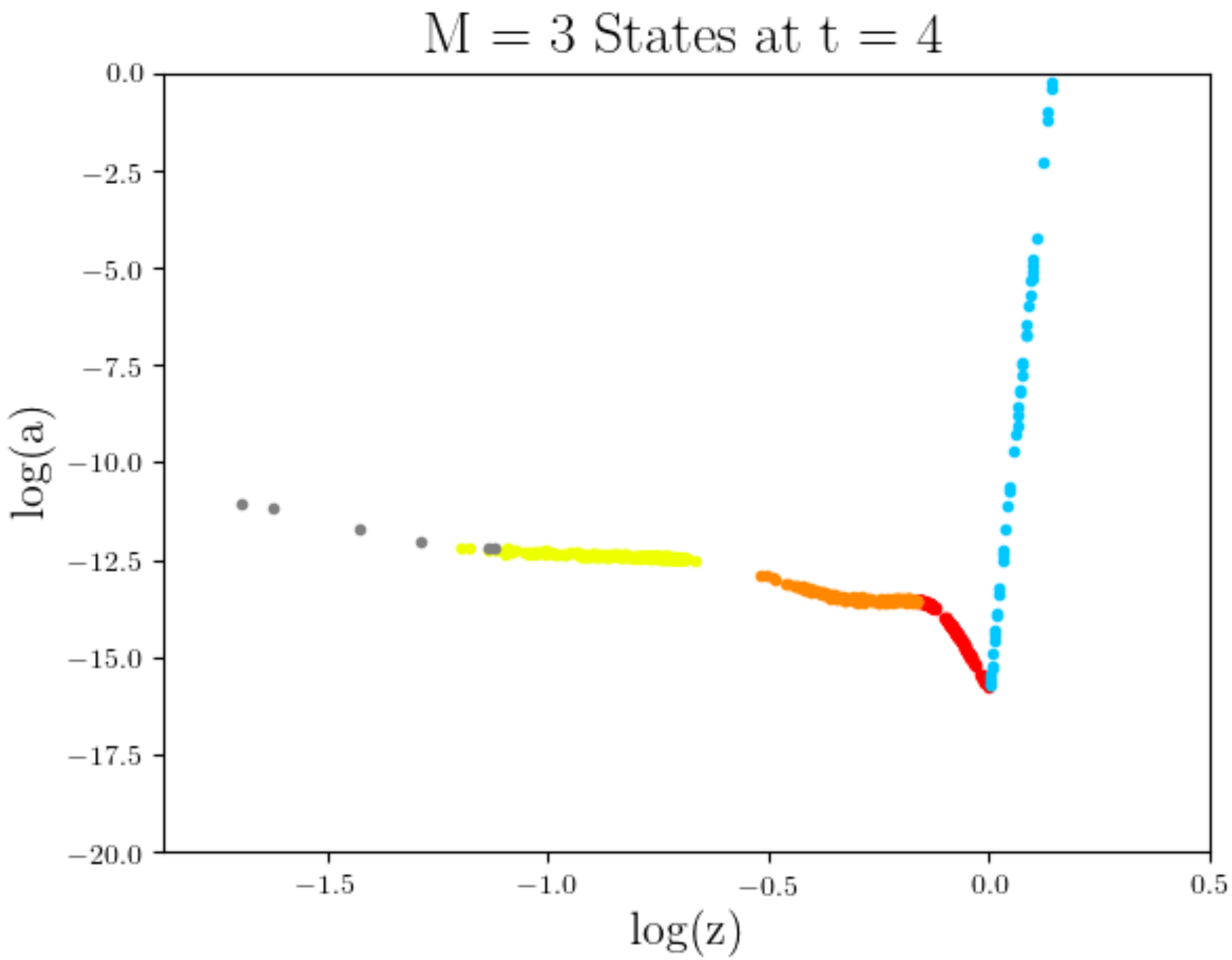
Clustering M = 3, 4 iterations



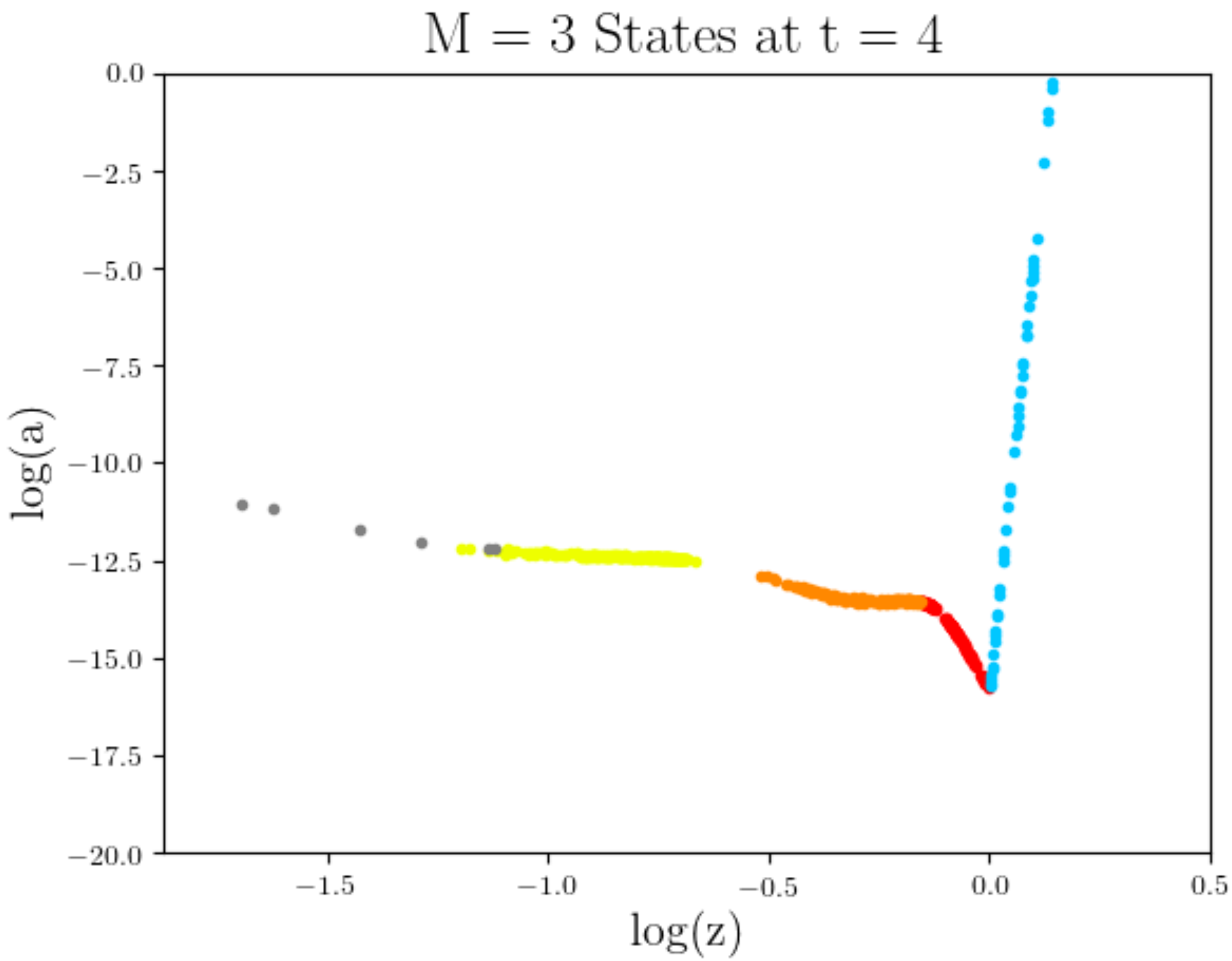
Clustering M = 3, 5 iterations



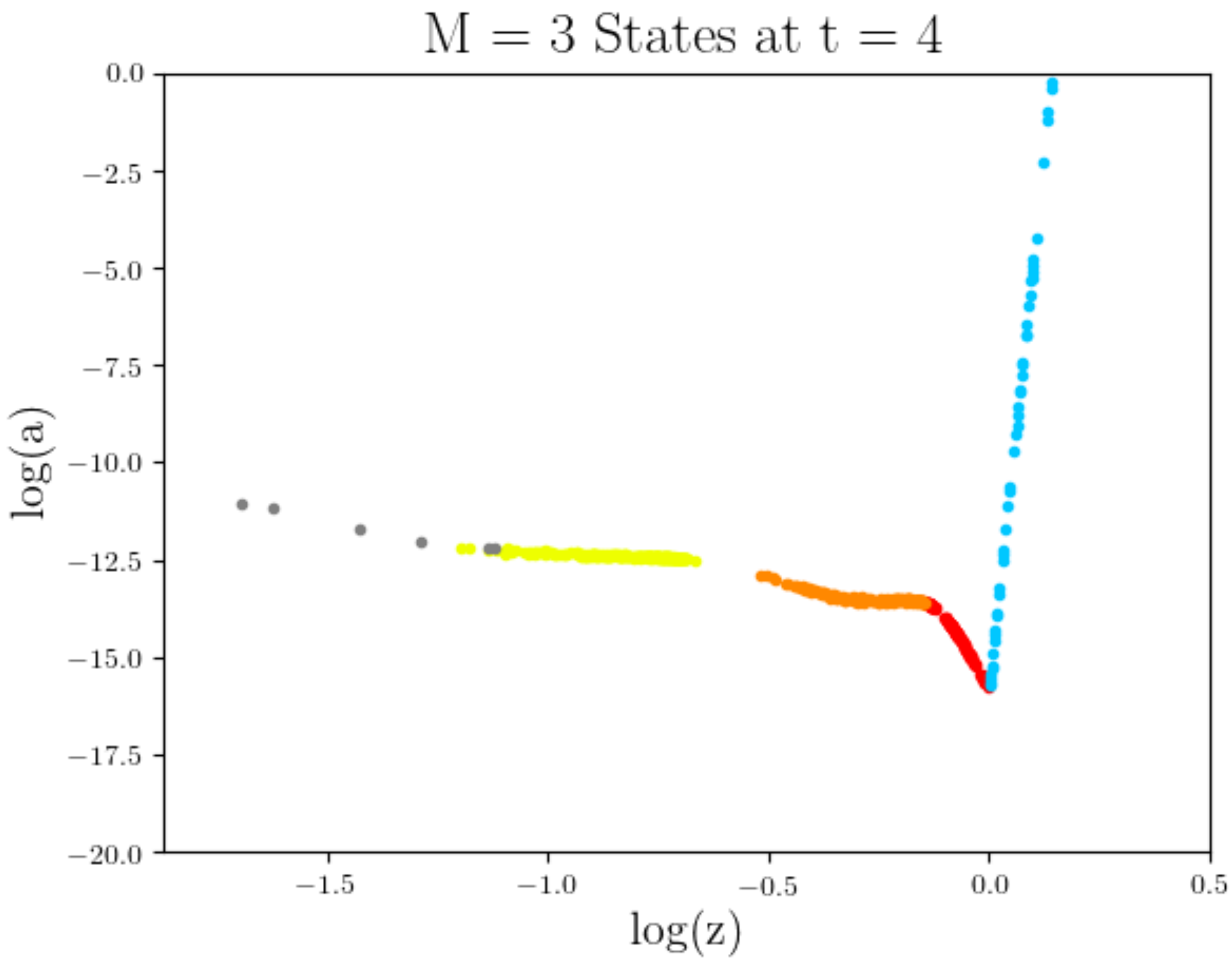
Clustering M = 3, 6 iterations



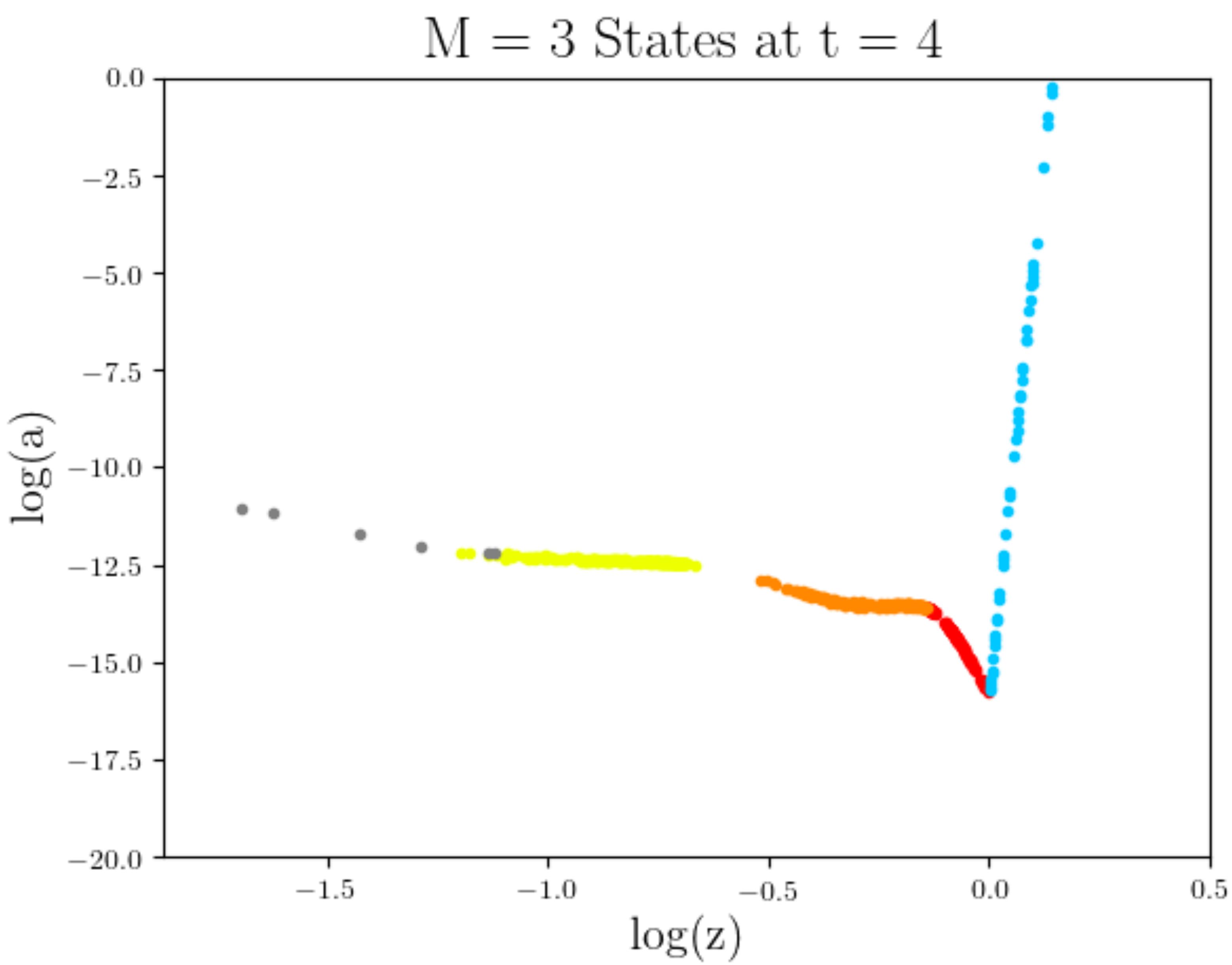
Clustering M = 3, 7 iterations



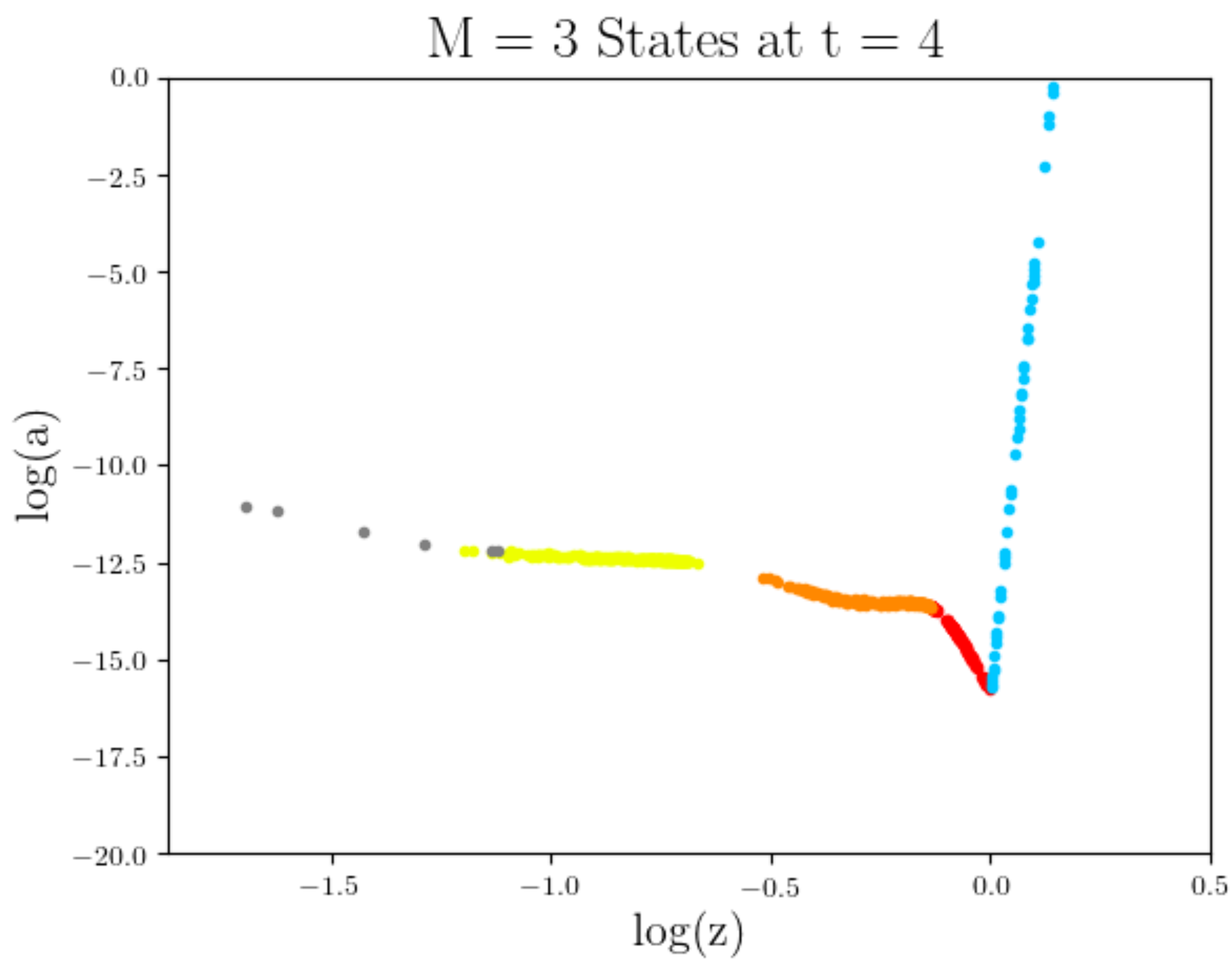
Clustering M = 3, 8 iterations



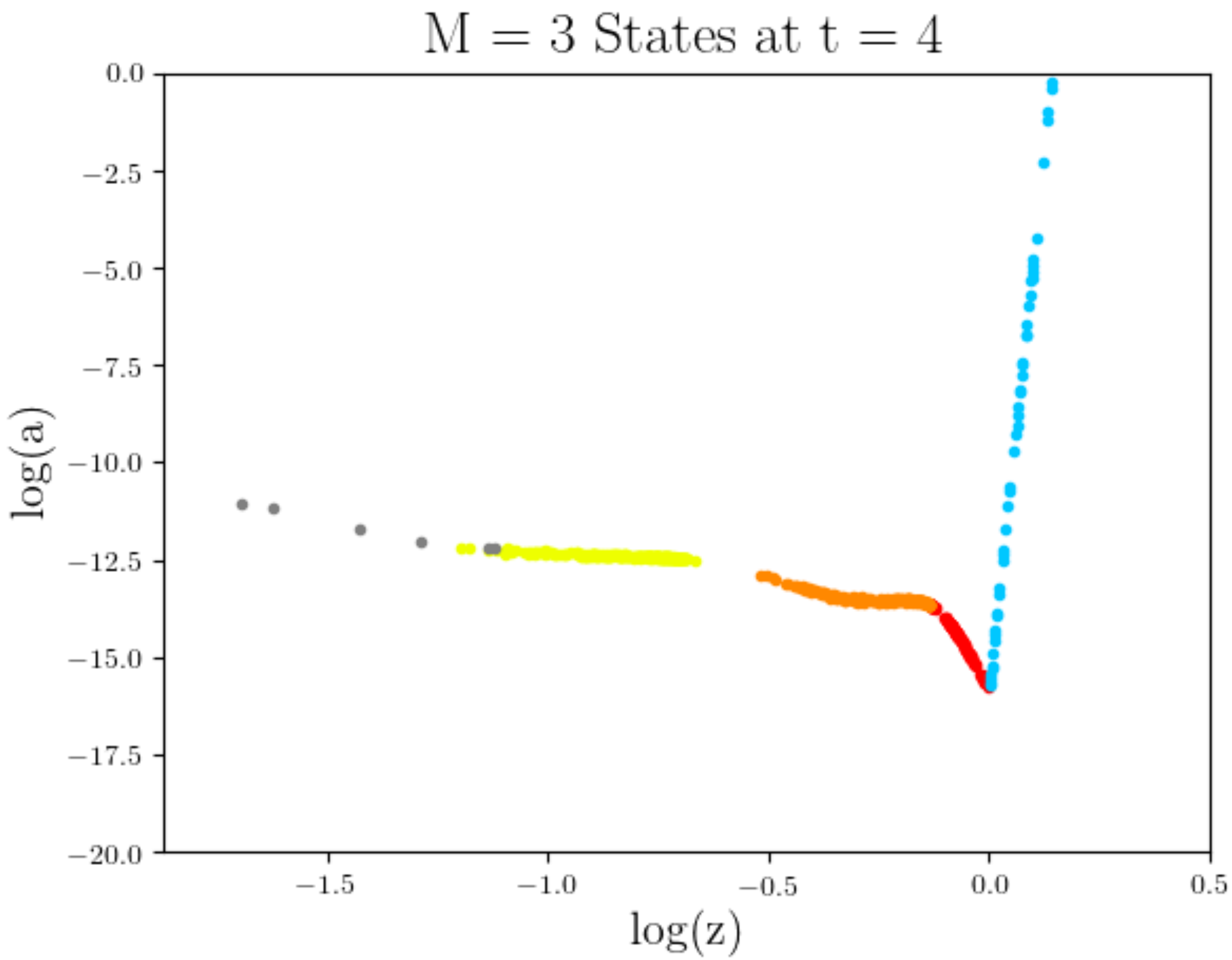
Clustering M = 3, 9 iterations



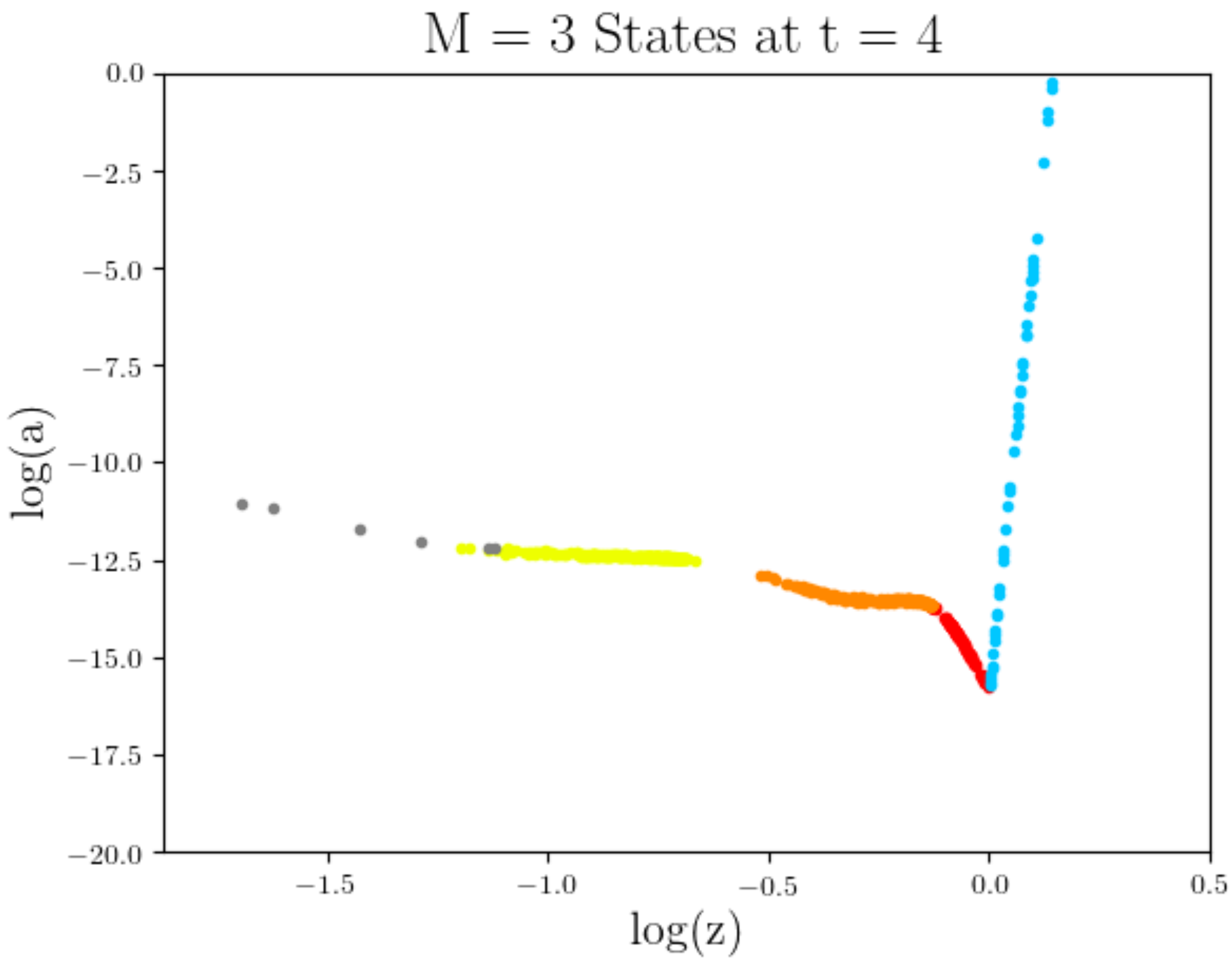
Clustering M = 3, 10 iterations



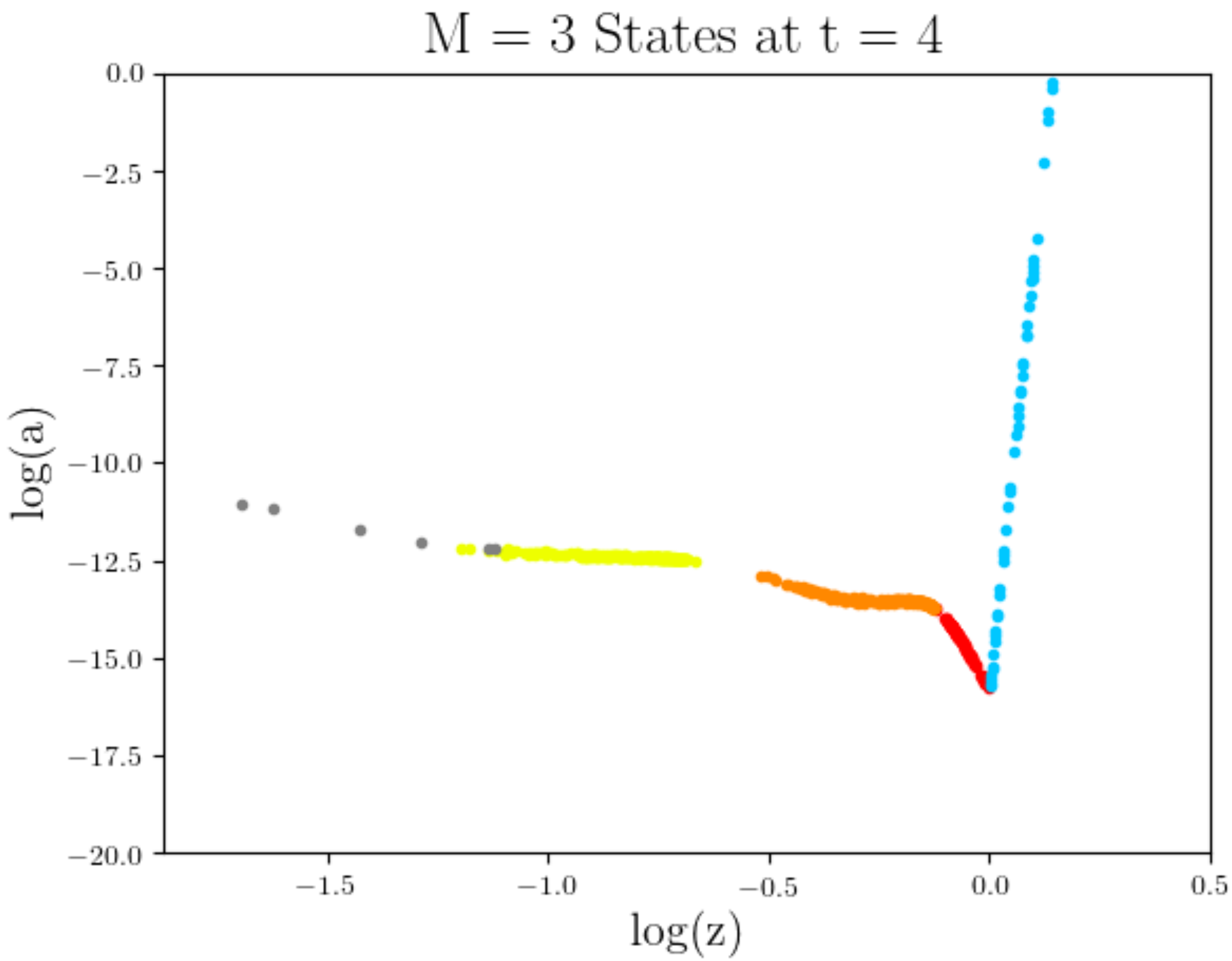
Clustering M = 3, 11 iterations



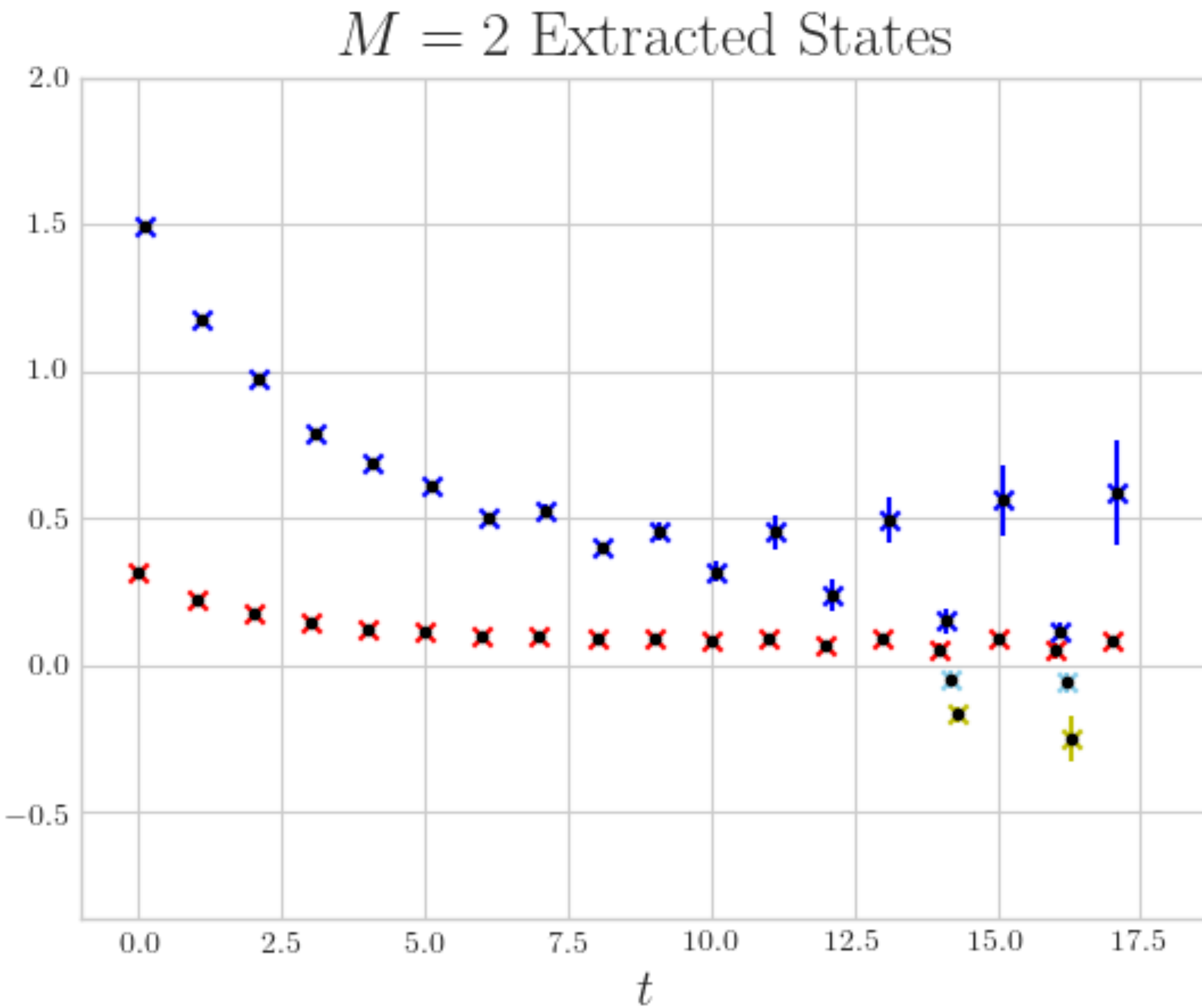
Clustering M = 3, 12 iterations



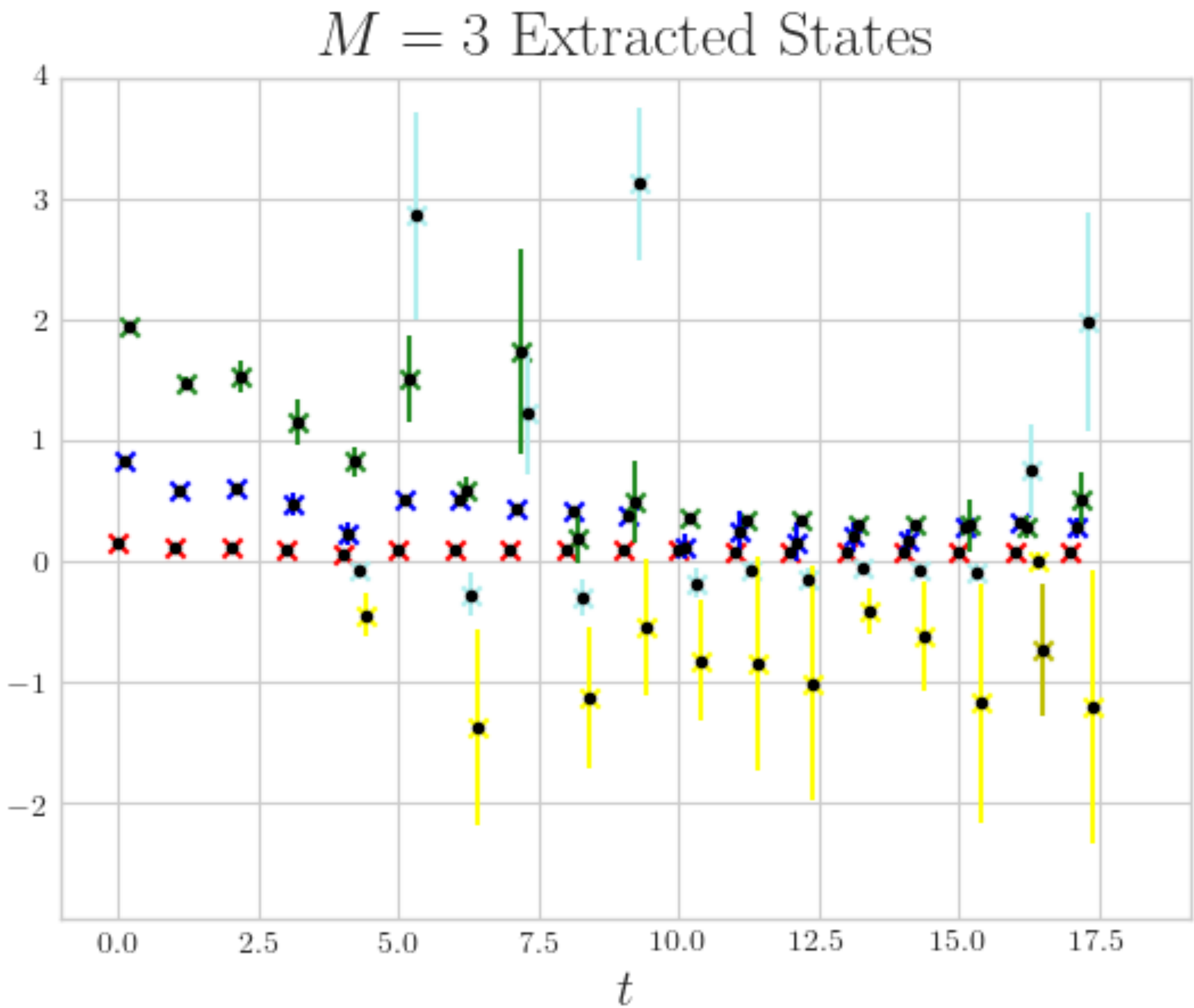
Clustering M = 3, 13 iterations



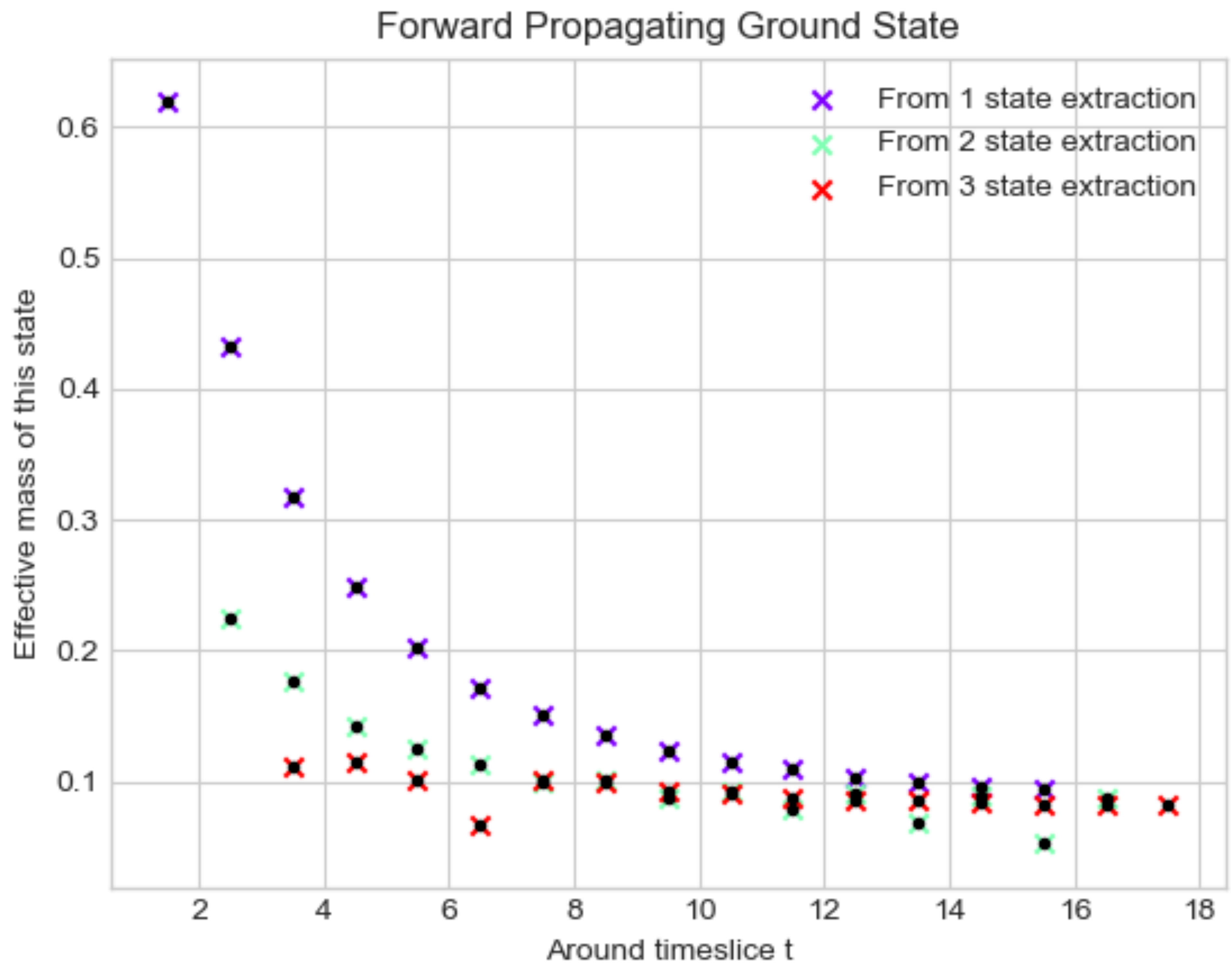
Preliminary Results



Preliminary Results

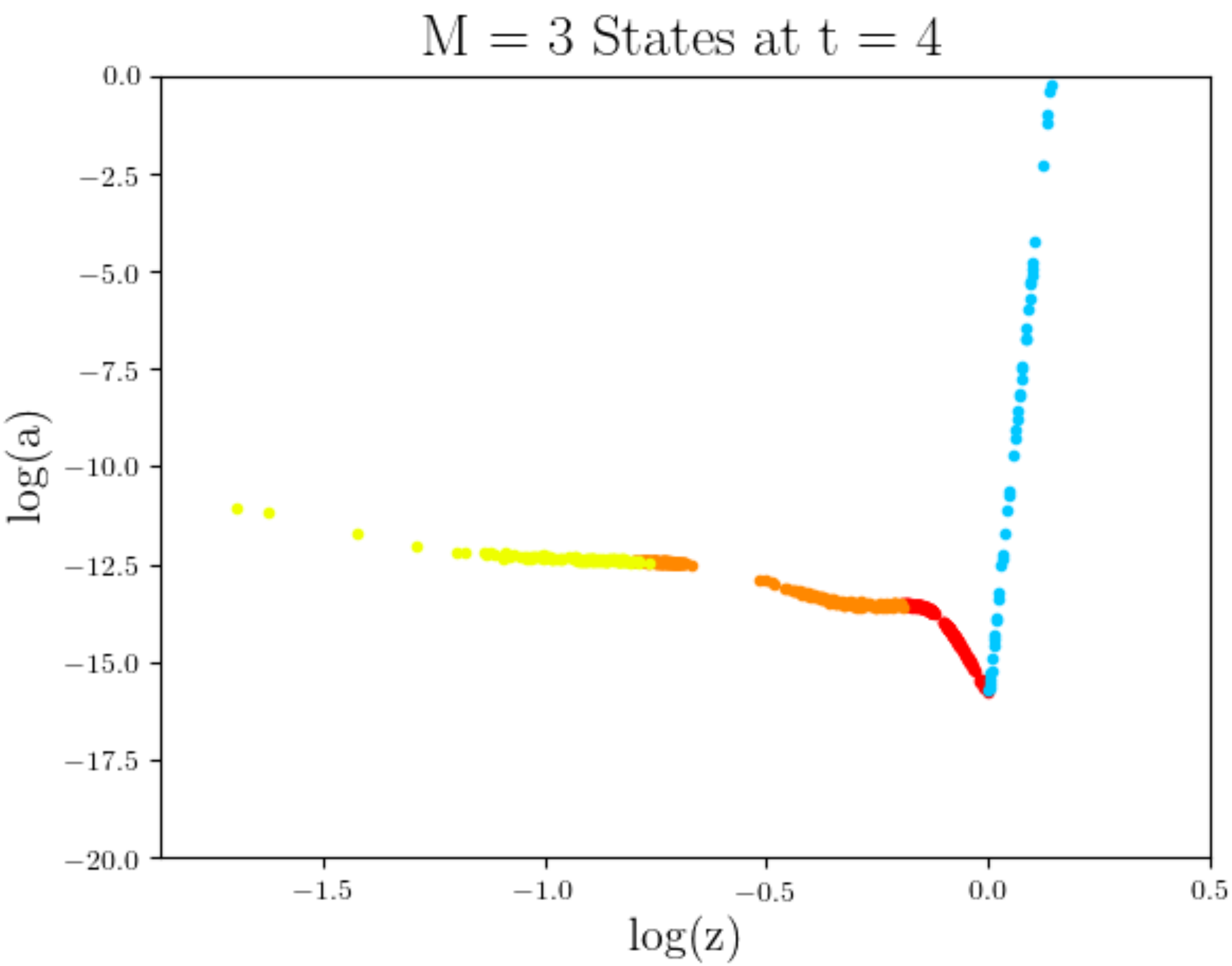


Preliminary Results



Future Work

- **Non-Gaussian clusters - try a new distance metric**
  -Don't worry about noisier time slices



Future Work

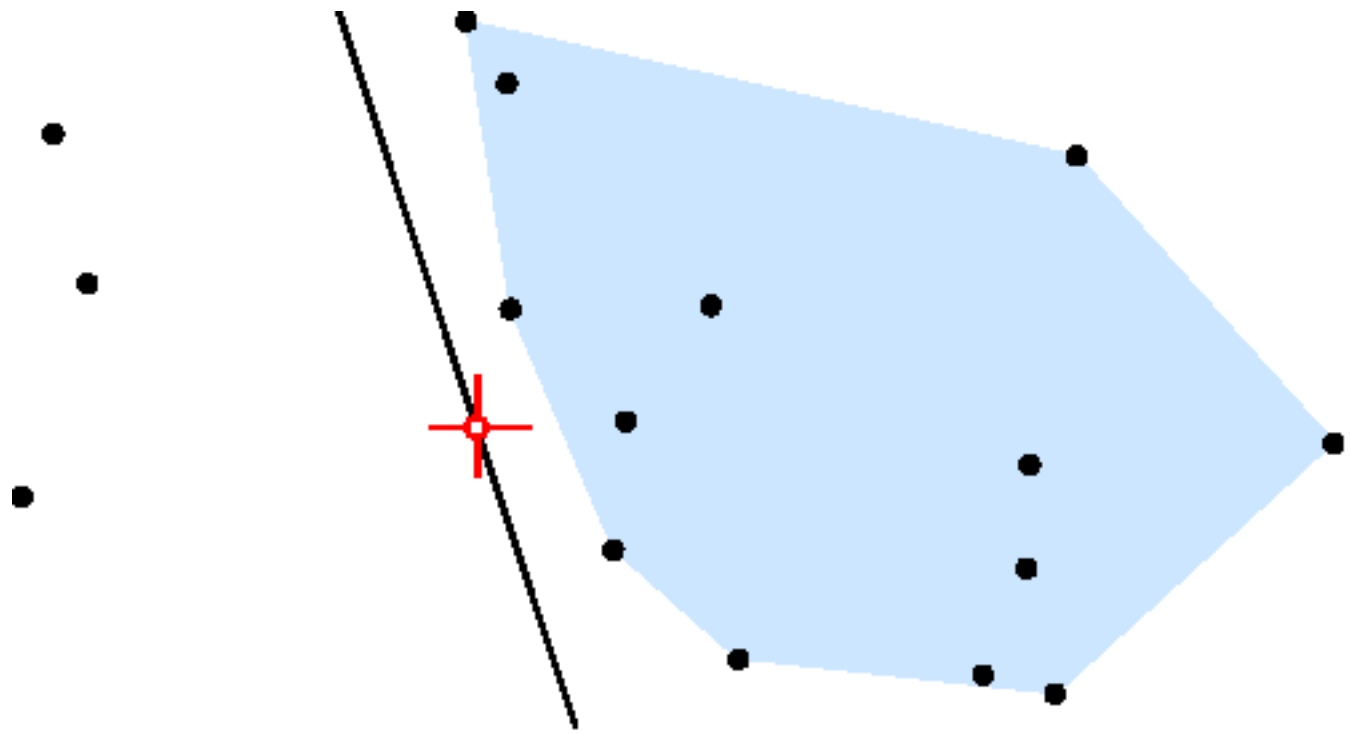
- **Non-Gaussian clusters - try a new distance metric**
  -Don't worry about noisier time slices

- **All data in "stencil" - extract** $M = T/2$ **states**
  -Remove all operator bias

$$0 = \begin{vmatrix} \underbrace{\begin{matrix} y_0 & y_1 & \cdots & y_{M-2} & y_{M-1} \\ y_1 & y_2 & \cdots & y_{M-1} & y_M \\ y_2 & y_3 & \cdots & y_M & y_{M+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_M & y_{M+1} & \cdots & y_{2M-2} & y_{2M-1} \end{matrix}}_{\text{Hankel matrix}} & \begin{matrix} 1 \\ z \\ z^2 \\ \vdots \\ z^M \end{matrix} \end{vmatrix}$$

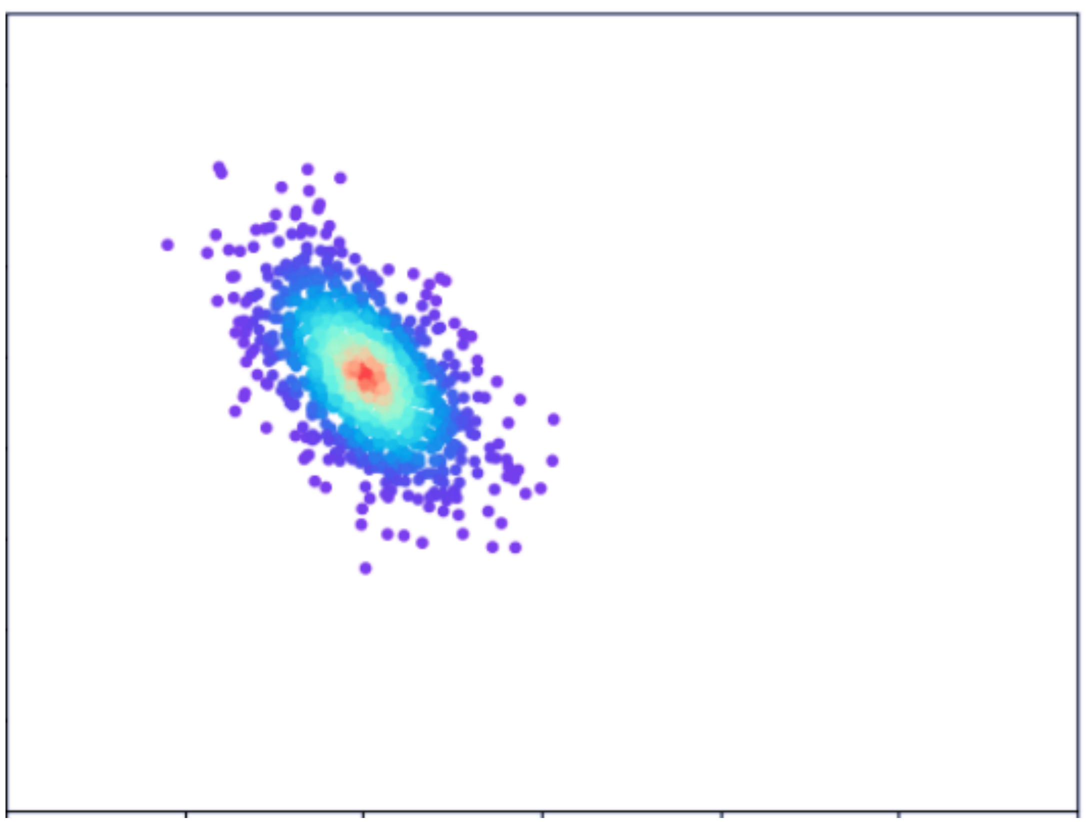
Possible solution: Tukey Depth

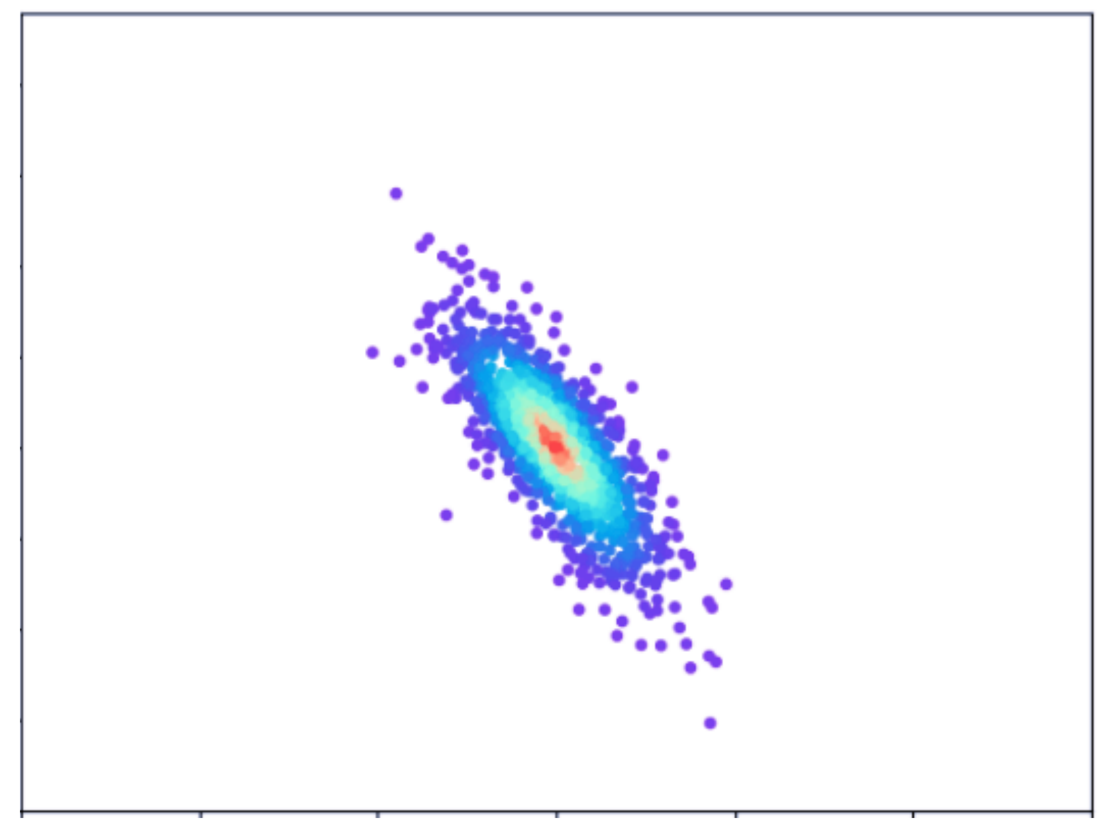
Multidimensional generalization of *percentiles*



Possible solution: Tukey Depth

Multidimensional generalization of *percentiles*

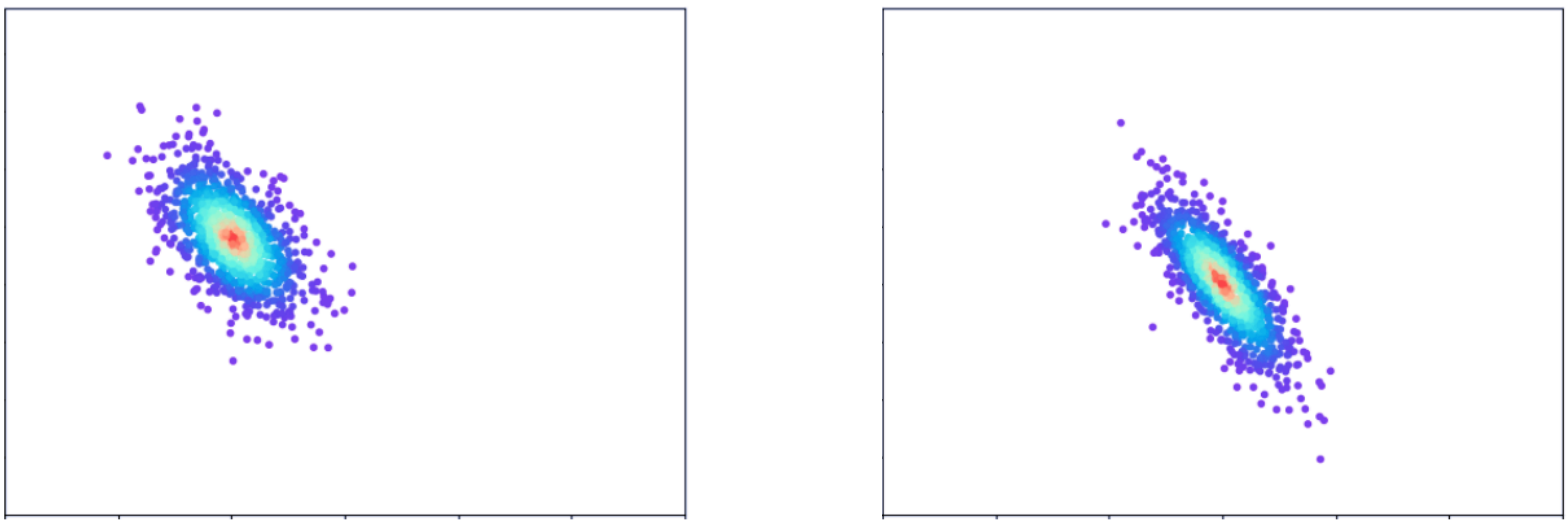




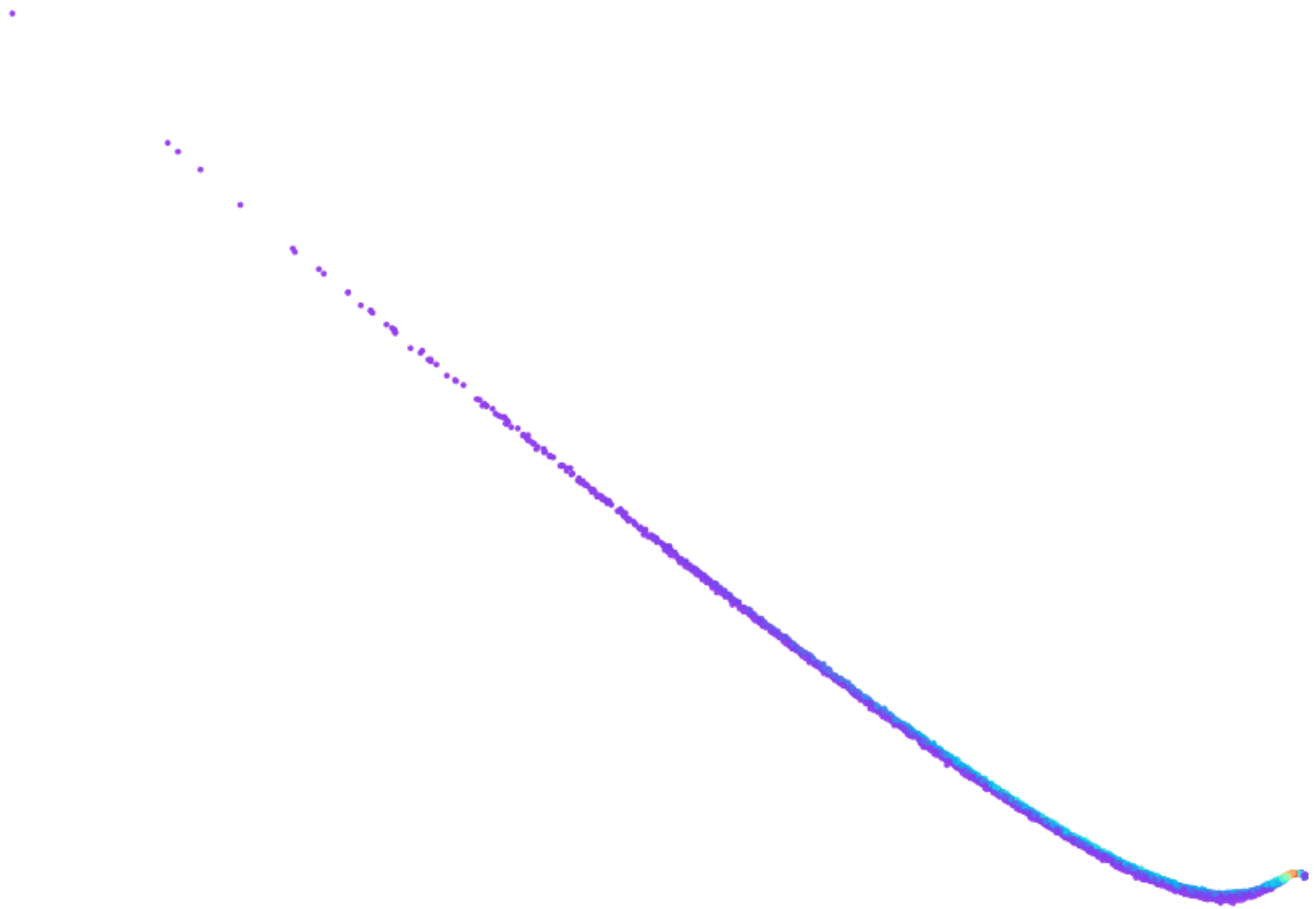
Possible solution: Tukey Depth

Multidimensional generalization of *percentiles*

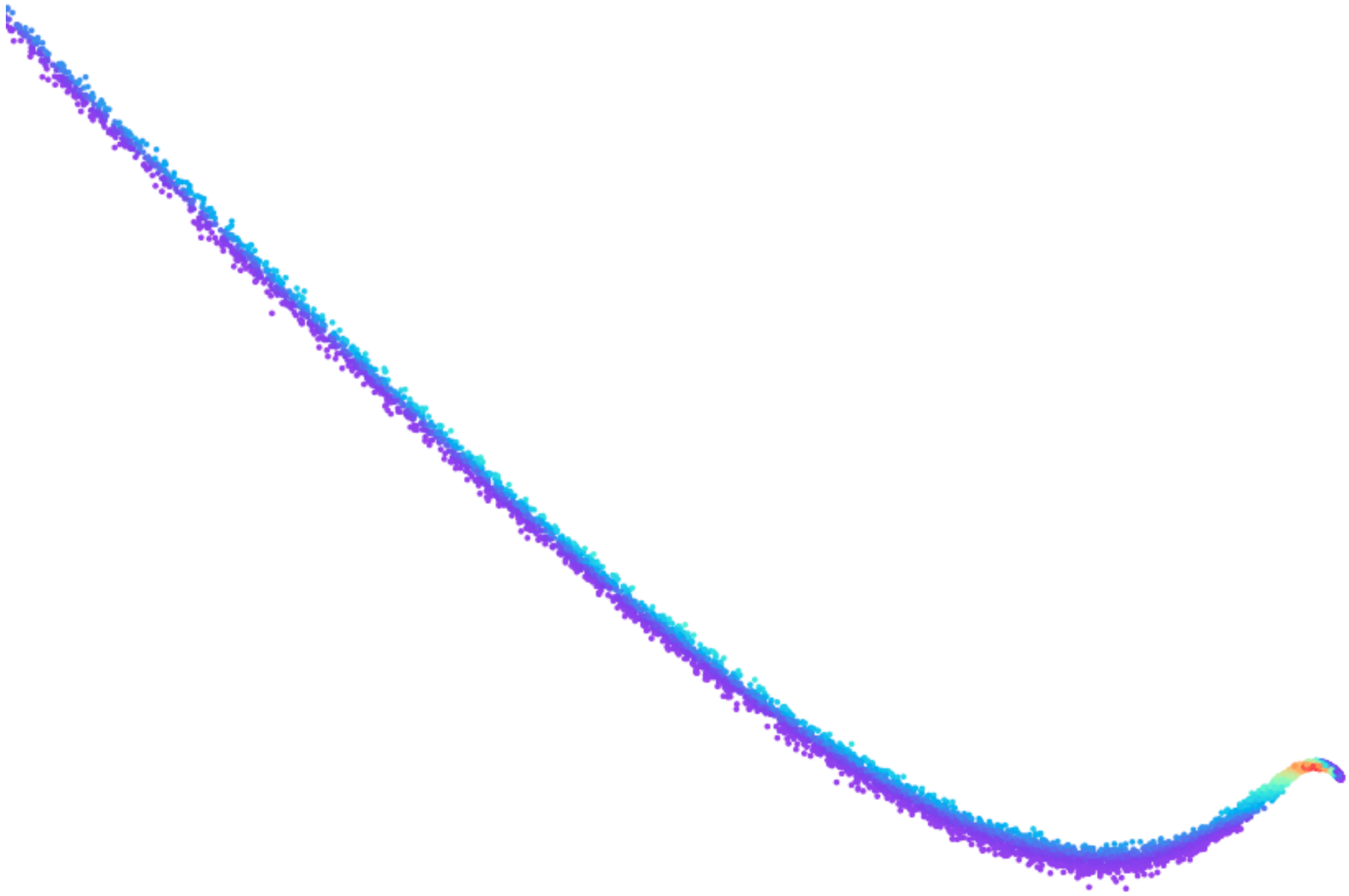
Non-parametric statistic!



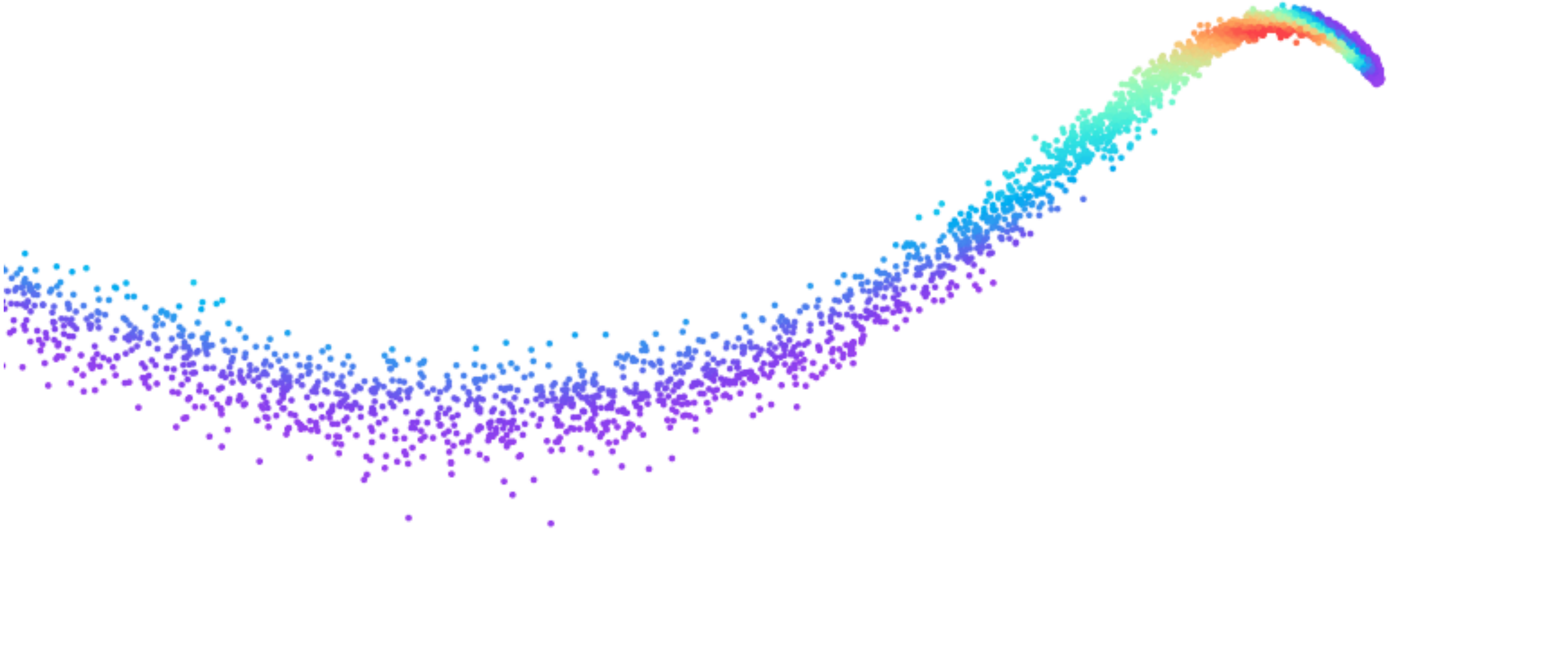
Tukey Depth Example with non-Gaussian Data



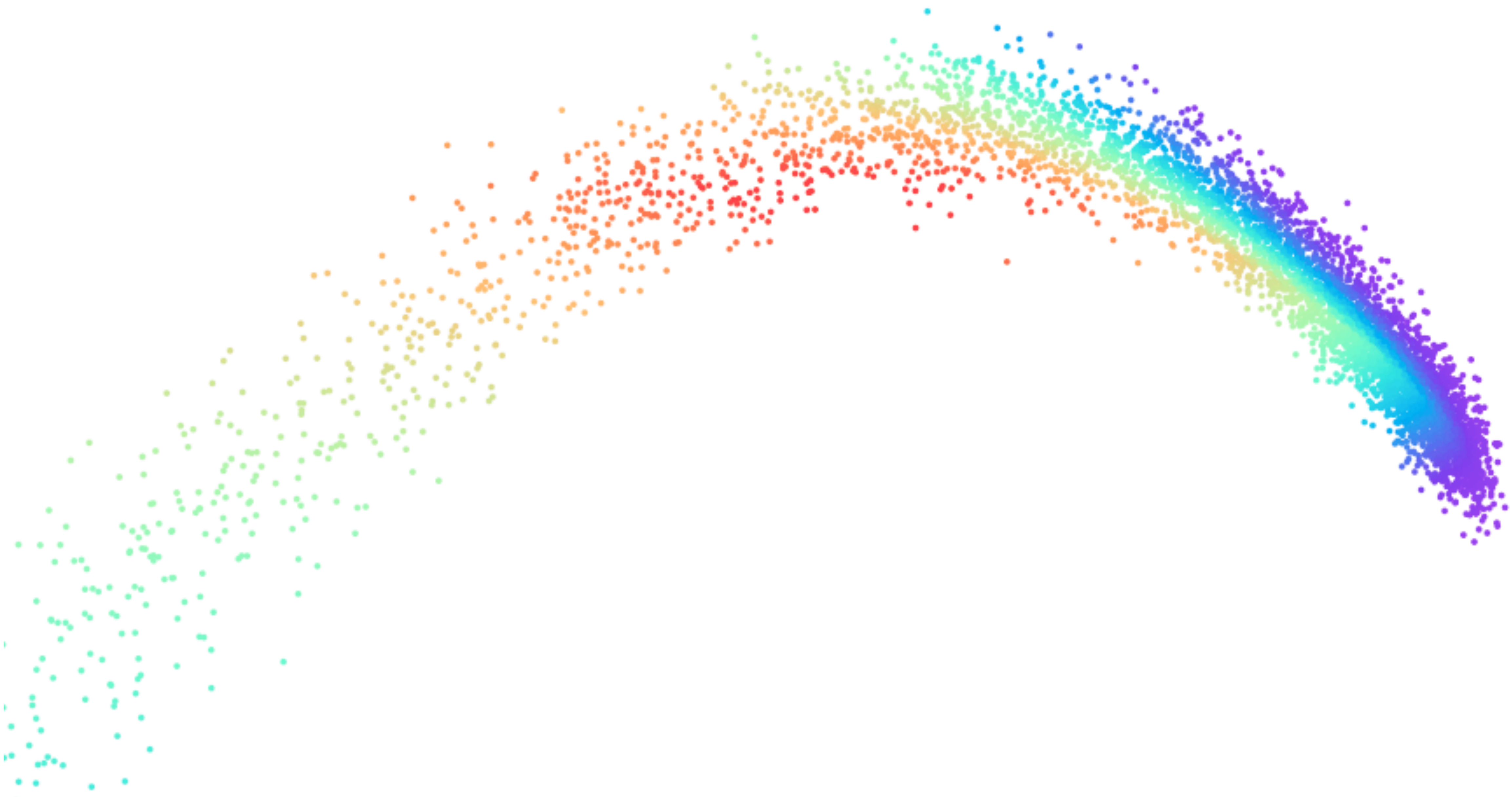
Tukey Depth Example with non-Gaussian Data



Tukey Depth Example with non-Gaussian Data



Tukey Depth Example with non-Gaussian Data



Conclusion

- **Extracting excited states is important for lattice QCD and BSM lattice**
- **Standard method involves fitting to exponentials and has many know problems**
- **Prony's method may be a better approach if we can identify clusters**
- **We will need a better clustering algorithm to account for weirdly shaped clusters**

Thanks!