11 July 2024

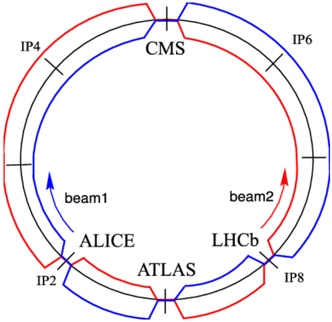
Extrapolation of Time Evolution of Chaos Indicators Applied to Single-Particle Tracking Simulations in Circular Accelerators

Abigail Levison

Supervisors:
Tatiana Pieloni (LPAP)
Massimo Giovannozzi
Carlo Emilio Montanari

Lattice Optimisation for LHC, HL-LHC

- Maximize luminosity, beam stability, and minimize particle losses from high energy beams to prevent equipment damage (HL-LHC beams: 700 MJ)
- Accelerator lattice quality assessed through single-particle tracking
- Ideal scenario: single-particle tracking for 10^8 turns ($\approx$ 10 hours of LHC runtime)



LHC Schematic

Probing Beam Dynamics at 10^8 turns

- Realistic LHC lattice simulations reach 10^6 turns: 10^4 particles takes ≈ 2 days on GPU
- Hénon map: simplified model, allows tracking 10^4 particles up to 10^8 turns in ≈ 2 hours

The Hénon map gives fundamental behaviour expected from a realistic accelerator lattice:

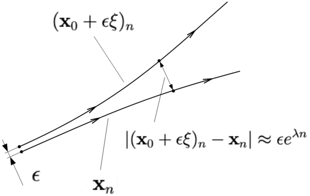
$$\begin{pmatrix} x_{n+1} \\ p_{x,n+1} \\ y_{n+1} \\ p_{y,n+1} \end{pmatrix} = \underbrace{R\left(\omega_{x,n}, \omega_{y,n}\right)}_{\substack{\text{Linear part} \\ \text{Dipoles, Quadrupoles}}} \begin{pmatrix} x_n \\ p_{x,n} \quad +x_n^2 - y_n^2 + \mu\left(x_n^3 - 3x_n y_n^3\right) \\ y_n \\ p_{y,n} \quad \underbrace{-2x_n y_n + \mu\left(y_n^3 - 3y_n x_n^3\right)}_{\substack{\text{Non-linear magnet kick} \\ \text{Applied as } \delta \text{ function}}} \end{pmatrix}$$

Modulation introduced by varying $\omega_{x,y}$ with number of turns n

Chaotic Regions of phase space

Initial separation ϵ between two initial conditions grows over time according to:

$$|(\mathbf{x}_0 + \epsilon\xi)_n - \mathbf{x}_n| \approx \epsilon e^{\lambda n}$$



Lyapunov Exponent λ

$\lambda > 0 \implies$ orbit defined as chaotic
$\lambda = 0 \implies$ orbit defined as regular

Faster identification of chaos $\longrightarrow$ link between chaotic phase-space regions and beam loss dynamics

Lyapunov Exponent

$$\lambda = \lim_{n\to\infty} \lim_{\epsilon\to 0} \frac{1}{n} \ln \frac{\|(\mathbf{x}_0 + \epsilon\xi)_n - \mathbf{x}_n\|}{\epsilon}$$

- Mathematical object defined for $n \to \infty$
- Its value is estimated by means of chaos indicators

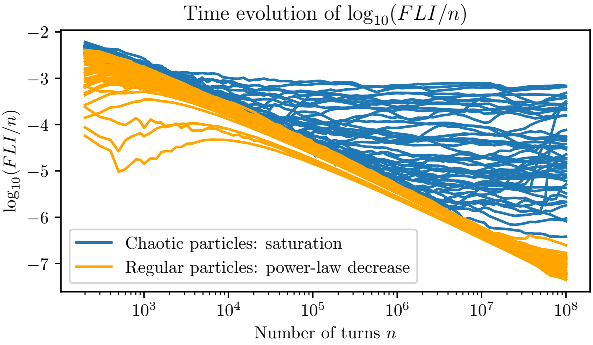
This study considers two (out of many!) chaos indicators:

- Fast Lyapunov Indicator (FLI)
- Reversibility Error Method (REM)

Fast Lyapunov Indicator (FLI)

$$\mathrm{FLI}_n(\mathbf{x}_0, \xi) = \lim_{\epsilon \to 0} \ln \frac{\|(\mathbf{x}_0 + \epsilon\xi)_n - \mathbf{x}_n\|}{\epsilon}$$

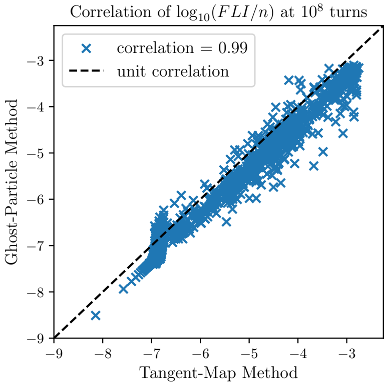
- Provides computation of the Lyapunov exponent for finite time
- FLI/n (time average) converges to the Lyapunov Exponent as $n \to \infty$



Two Ways to Generate the FLI

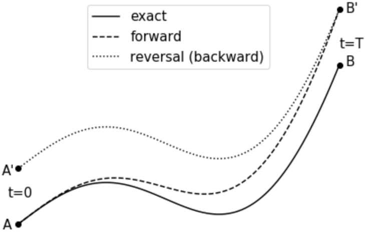
- Tangent-map method: direct analytical computation, only possible for Hénon map
- Ghost-particle method: approximation, possible for any lattice

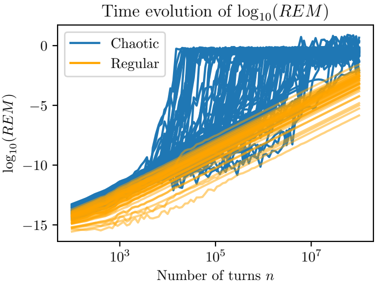
Examination of both methods shows no significant difference in values at 10^8 turns for the two methods



Reversibility Error Method (REM)

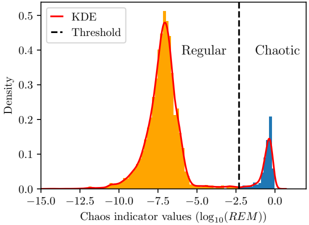
- REM measures particle displacement due to numerical errors to evaluate chaotic behavior
- Regular particles: power law increase
- Chaotic particles: exponential growth, timescale determined by Lyapunov exponent
- Chaotic saturation corresponding to diameter of bounded motion region

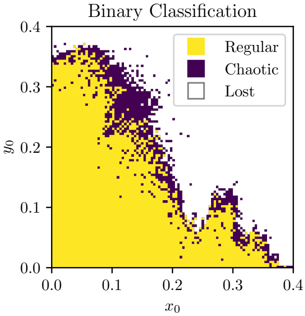




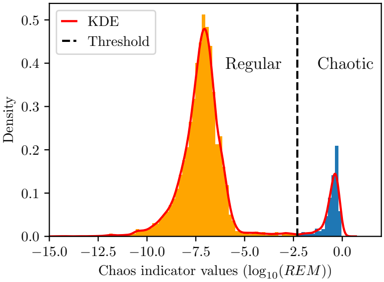
State-of-the-art Performance Analysis

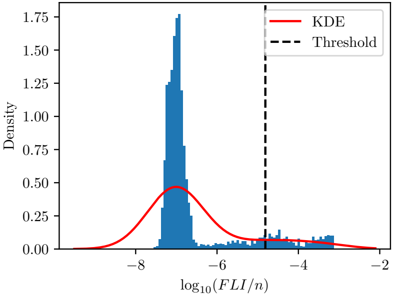
- Ranks chaos indicators according to fast identification of chaos for modulated Hénon map
- Ground truth binary classification at 10^8 turns
- Classifications at smaller numbers of turns provides benchmark classification accuracy
- REM was one of the highest-performing indicators





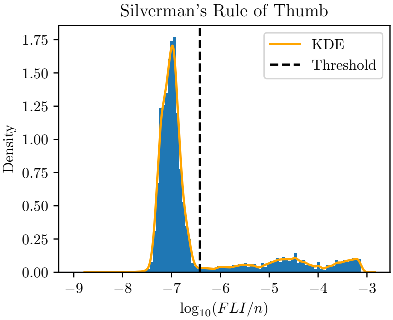
Kernel Density Estimation (KDE) Algorithm





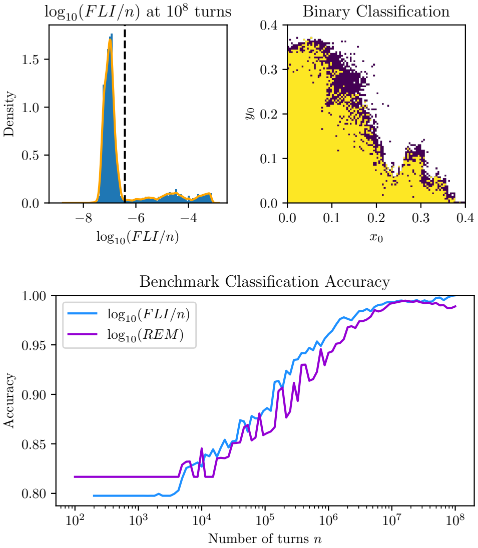
- KDE estimates histogram probability density function
- Threshold set at KDE minimum between peaks
- Algorithm identifies how quickly chaos indicator values separate into bimodal distribution
- Not effective for multimodal distributions

Improvement to the Classification Algorithm



- Original contribution: algorithm adaptation
- Silverman's Rule of Thumb: statistical approach to compute KDE
- Better captures multimodal histogram distributions

Benchmark Classification Accuracy



- FLI classification at 10^8 turns taken as ground truth for FLI and REM in line with state-of-the-art analysis ideology
- FLI performance comparable to REM
- REM does not reach unit accuracy at 10^8 turns due to different binary classification from FLI at 10^8 turns

Can the classification performance of chaos indicators be increased through extrapolation techniques?

ARIMA model: established extrapolation tool capable of extrapolating trends using minimal set of free parameters

Linear combination of an autoregressive (AR) and/or moving average (MA) model fit to a *stationary* time series:

- Mean μ and white noise a with variance σ_a^2
- Data is differenced d times until stationary:

$$z'_t = z_t - z_{t-1}$$

ARIMA model components

Auto-regressive (AR) model order p:

$$z_t = \mu + a_t + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} \,,$$

Moving Average (MA) model order q:

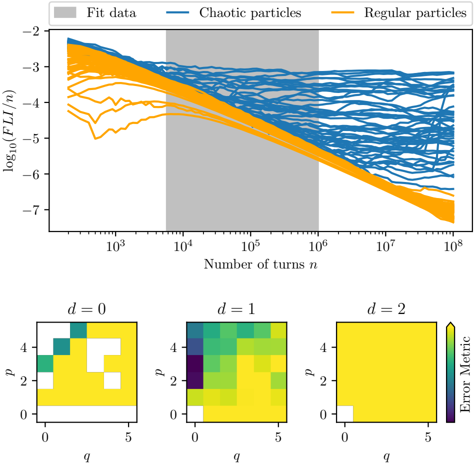
$$z_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \,,$$

Combined ARIMA model order (p, d, q)

$$z_t = \mu + a_t \underbrace{+\phi_1 z_{t-1} + \ldots + \phi_p z_{t-p}}_{\text{AR component}} \underbrace{-\theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}}_{\text{MA component}}$$

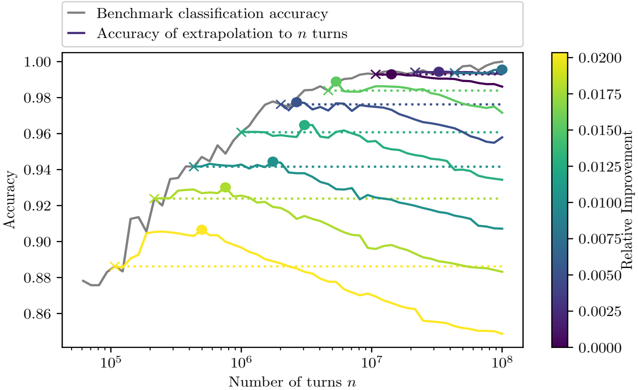
Optimised parameters: μ, σ_a^2, ϕ, θ

ARIMA parameter scan to find optimal order (p, d, q)



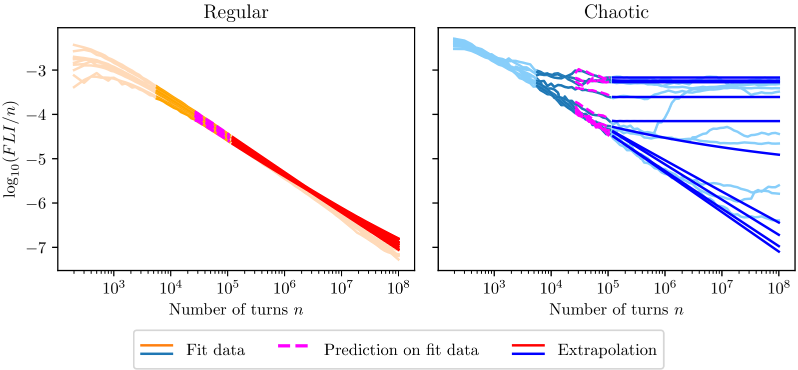
- Fit ARIMA model of order (p, d, q) to time evolution where trend manifests, up to 10^6 turns
- Extrapolate to 10^8 turns; evaluate absolute error at 10^8 turns
- Repeat process for 100 initial conditions in sample, sum up the absolute error

Extrapolation performance after providing data up to various numbers of turns



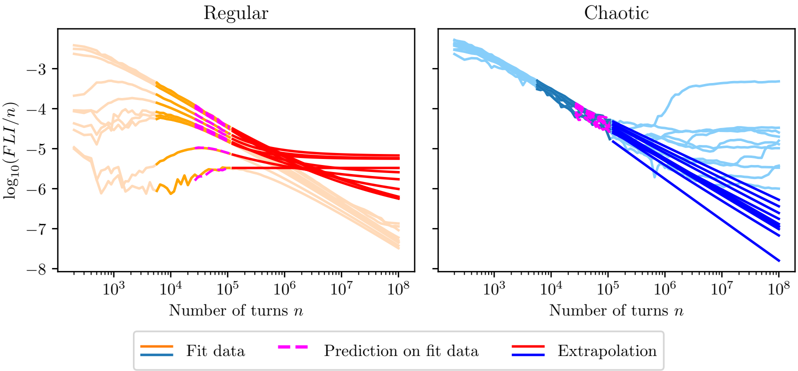
- Initial increase in accuracy, but subsequent decrease
- Extrapolation accuracy never exceeds benchmark accuracy

Extrapolation from 10^5 turns appears promising ...



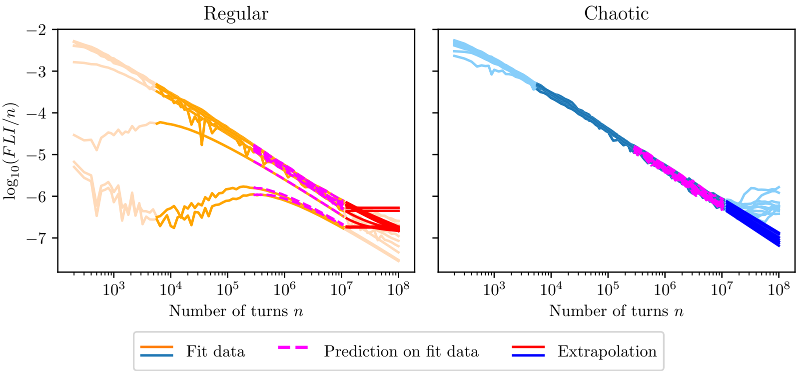
- ARIMA models appear capable of extrapolating power law decrease trend

Initial improvement but subsequent decrease in accuracy?



- Extrapolations have not deviated enough from actual trend to result in incorrect classifications at $10^{5.70}$ turns
- Various features of FLI emphasised just enough by ARIMA model at $10^{5.70}$ turns to increase extrapolation accuracy

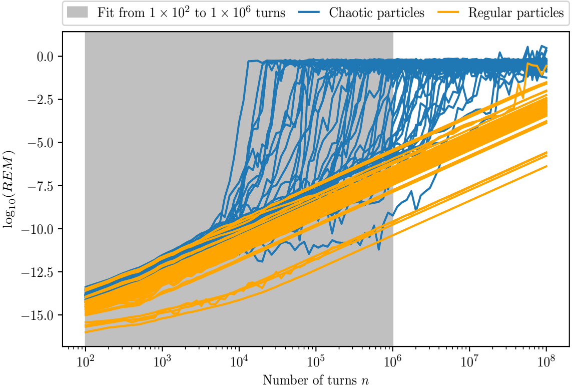
Providing data up to higher numbers of turns



Extrapolation accuracy immediately decreases:

- Overfitting for regular initial conditions
- ARIMA failure to predict chaotic saturation

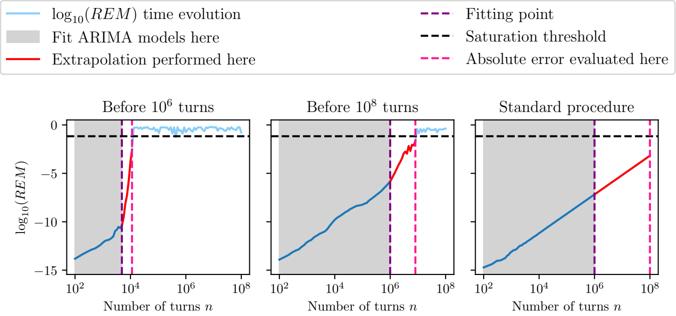
ARIMA extrapolation of REM



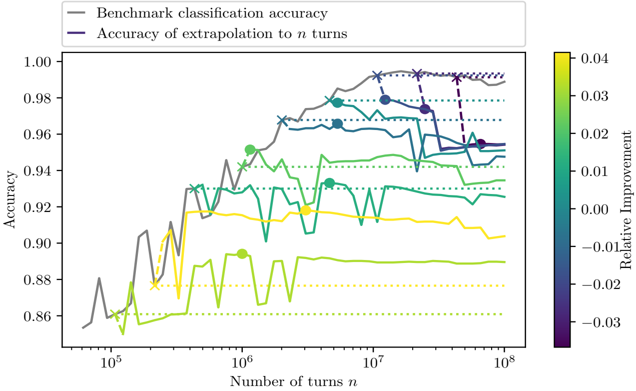
REM requires adaptations to fitting and extrapolation to avoid fitting ARIMA models to pure noise for saturated data

REM adaptation to parameter scan

- Saturation before 10^6 turns: fit 80% of data before saturation, extrapolate up to saturation
- Saturation before 10^8 turns: fit up to 10^6 turns, extrapolate up to saturation
- No saturation before 10^8 turns: fit up to 10^6 turns, extrapolate to 10^8 turns

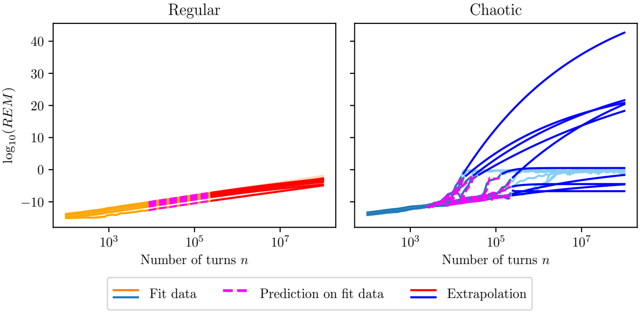


REM extrapolation results



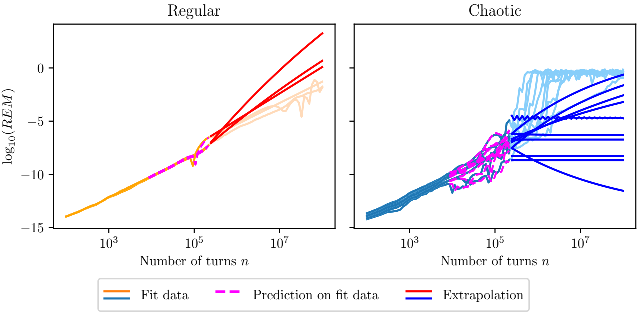
Benchmark never exceeded, but improvement in initial accuracy observed when extrapolating from $\approx 10^5$ turns to 10^8 turns

REM extrapolation from 2×10^5 turns



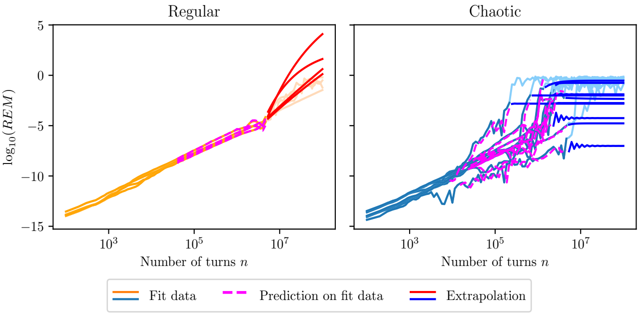
- Trend not exactly captured by ARIMA models
- Produces optimal final values for regular particles, but chaotic particles may level off too early

Incorrect classifications after extrapolation from 2×10^5 turns



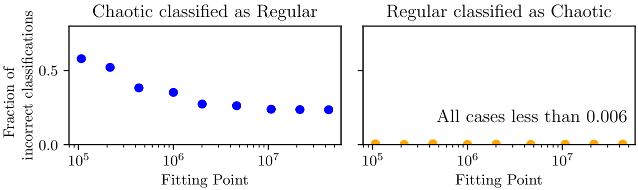
- Extrapolation continues in direction of fluctuations
- Overfitting evident for chaotic particles

Immediate decrease in accuracy when providing data up to high numbers of turns



- Some regular initial conditions show REM saturation before 10^8 turns, so the extrapolation leads to a chaotic classification
- ARIMA model focuses too much on fluctuations rather than the trend, resulting in further decrease in accuracy

Fractions of incorrectly classified initial conditions



- High fractions of chaotic classifications as regular due to levelling off too early and overfitting
- ARIMA models effective at predicting regular REM values at 10^8 turns even if trend is not captured

Conclusions

Reproduction of benchmark classification accuracy:

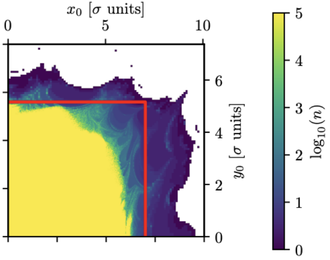
- FLI performance comparable to REM when using Silverman's Rule of Thumb for classifications

ARIMA extrapolations of $\log_{10}(\mathrm{FLI}/n)$ and $\log_{10}(\mathrm{REM})$

- Initial improvement in accuracy observed when extrapolating from $\approx 10^5$ turns for FLI and REM, but subsequent decrease due to ARIMA misunderstanding of trends
- Extrapolation accuracy never exceeds benchmark accuracy for FLI or REM
- Considerations from this study can be used for future investigations of extrapolation techniques using more parameters

Assessing Accelerator Lattice Quality

- Ideal scenario: single-particle tracking for 10^8 turns (≈ 10 hours of LHC runtime)
- Determine volume in phase space where particles have bounded orbits
- Realistic LHC lattice models contain thousands of advanced elements
- Current simulations with realistic lattices consider up to 10^6 turns



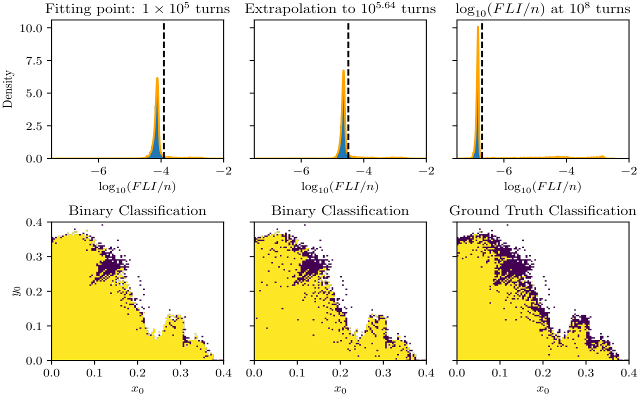
Survival plot for HL-LHC lattice at 10^5 turns

Modulated Hénon Map

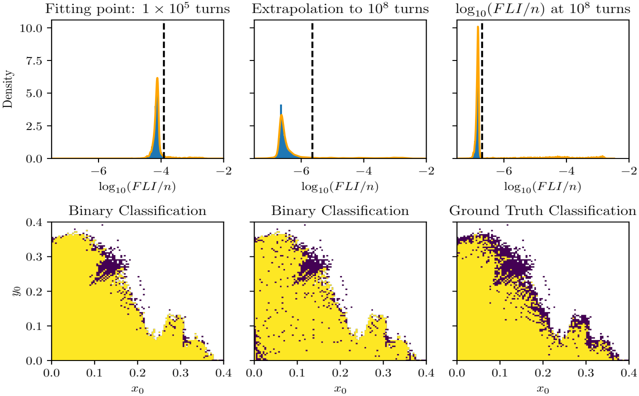
Modulation introduced by varying ω with time, inserting realistic SPS tune-ripple phenomena

$$\omega_{x,y,n} = \omega_{x,y,0}\left(1 + \varepsilon \sum_{k=1}^{m} \varepsilon_k \cos\left(\Omega_k n\right)\right)$$

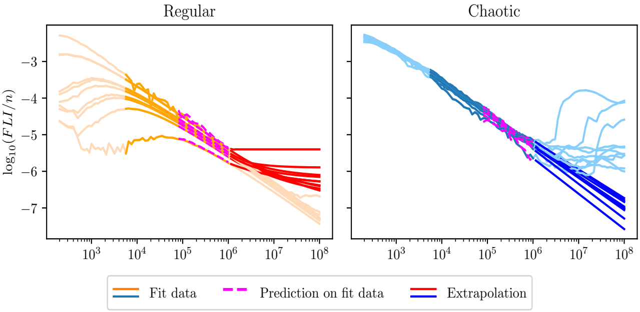
FLI Extrapolation from 10^5 turns



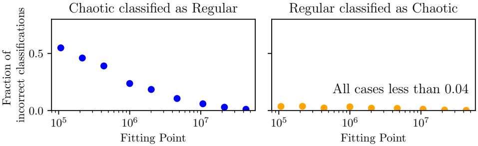
FLI Extrapolation from 10^5 turns

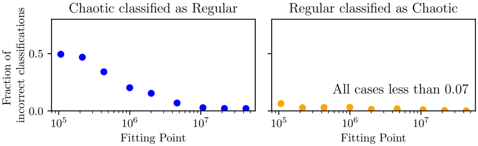


Very similar extrapolation results



Fractions of incorrectly classified initial conditions





- Tangent map method has more chaotic classified as regular
- Ghost particle method has more regular classified as chaotic

Kernel Density Estimation (KDE)

Statistical method to estimate probability density function of histogram distribution

$$f_h(x) = \frac{1}{n} = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right),$$

- Manually selected bandwidth h
- Assumed kernel function K (typically Gaussian)

Silverman's Rule of Thumb:

$$h = 0.9 \times \min(\sigma, \frac{\text{IQR}}{1.34}) \times n^{-1/5}$$

- Statistical approach to calculate h
- Best suited to unimodal Gaussian distributions

Aims of this Thesis

Reproduction of state-of-the-art results:

- Track chaos indicators' evolution using modulated Hénon map
- Establish ground-truth binary classification of regular or chaotic behavior

Original Contribution

- Refine the KDE-based classification algorithm
- Compare methods to generate FLI values
- Implement ARIMA models to extrapolate chaos indicators' evolution to 10^8 turns
- Compare accuracy and computational efficiency of tracking with extrapolation to simple tracking

Linear Response

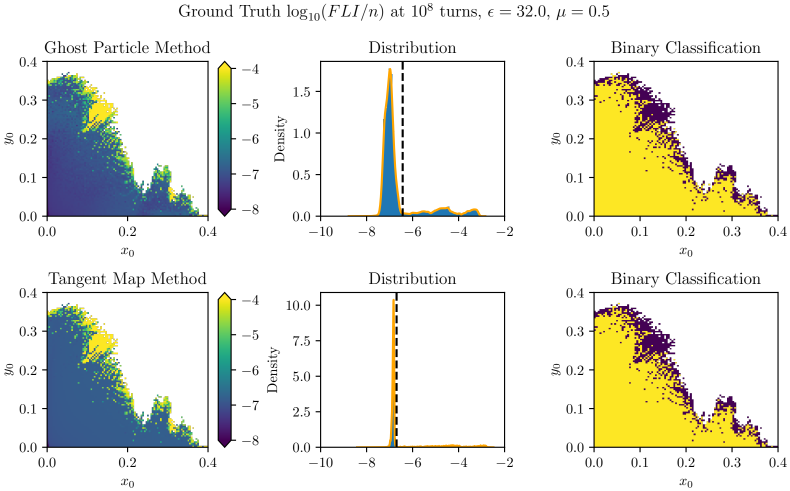
$$\boldsymbol{\Xi}_n(\mathbf{x}) = \lim_{\epsilon \to 0} \frac{\mathbf{y}_n - \mathbf{x}_n}{\epsilon}$$

Two ways to calculate the linear response in single-particle tracking simulations:

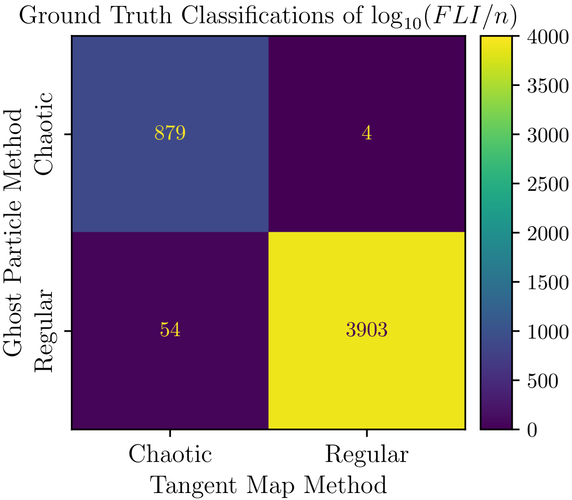
- Tangent-map Method (direct analytical computation)
- Ghost-particle Method (approximation)

These result in slightly different $\log_{10}(\mathrm{FLI}/n)$ distributions at 10^8 turns

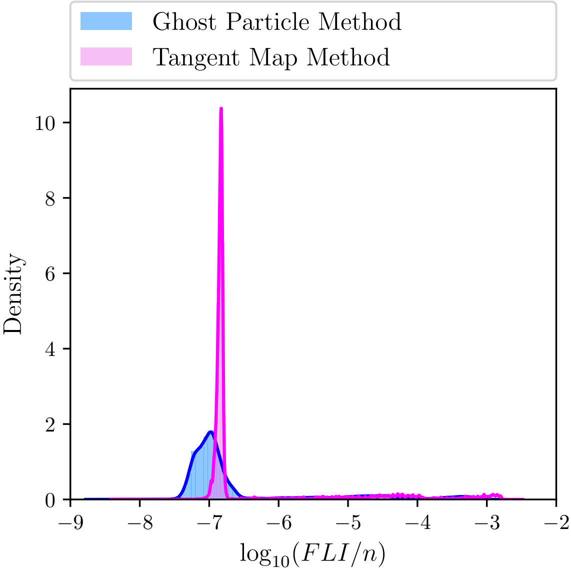
$\log_{10}(\text{FLI}/n)$ distribution comparison



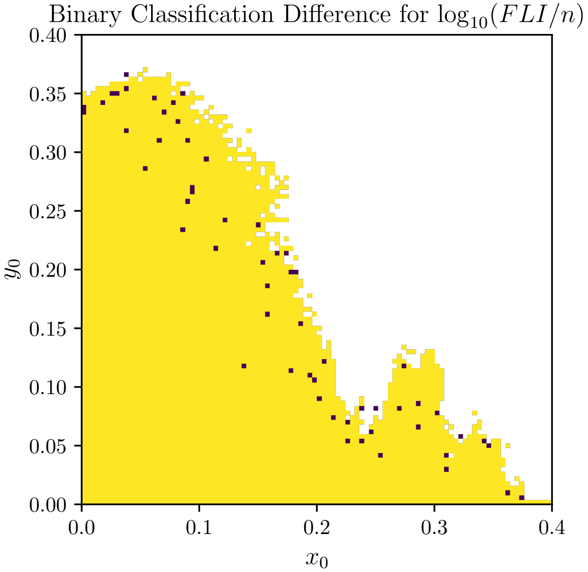
$\log_{10}(\mathrm{FLI}/n)$ distribution comparison



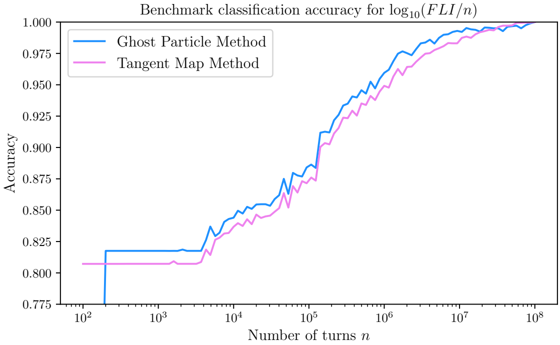
$\log_{10}(\mathrm{FLI}/n)$ distribution comparison



$\log_{10}(\mathrm{FLI}/n)$ distribution comparison

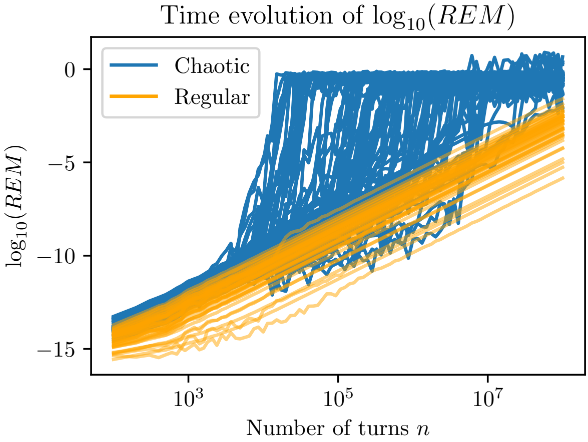


Benchmark Classification Accuracy for FLI



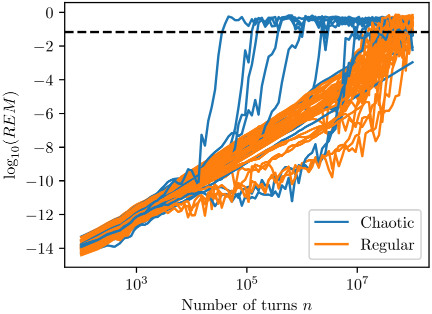
- Ghost-particle and tangent-map methods yield comparable benchmark accuracy
- Accuracy increases at 10^5 turns compared to 10^6 turns using standard algorithm

REM ground truth



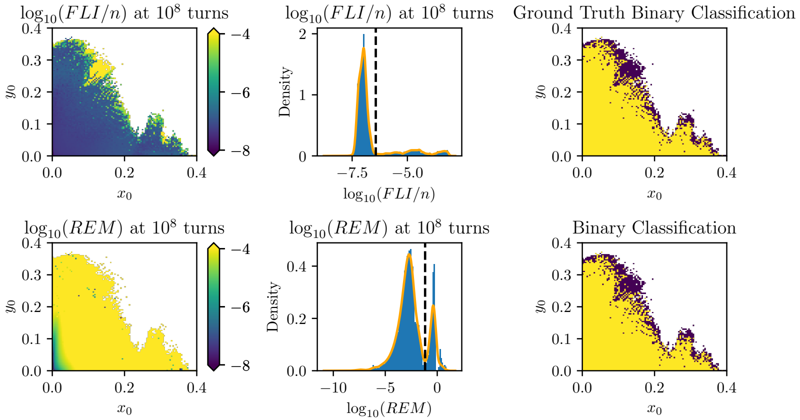
- Saturation of REM values corresponds to maximum radius of bounded orbits
- Power-law increase catches up to saturation at 10^8 turns

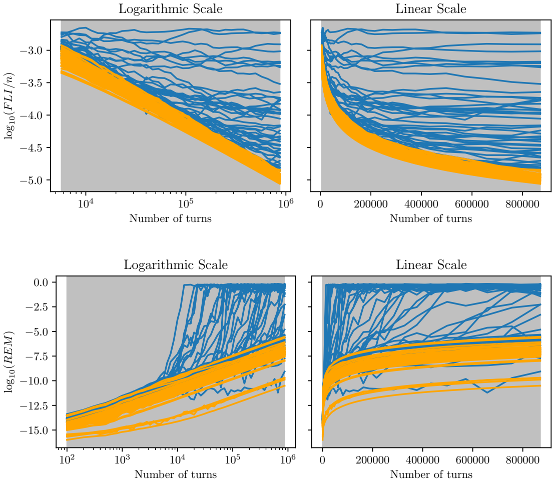
$\log_{10}(\text{REM})$ vs $\log_{10}(\text{FLI}/n)$ at 10^8 turns



- Using REM histogram for ground truth could misclassify saturated particles as regular
- Using FLI for ground truth reduces misclassifications

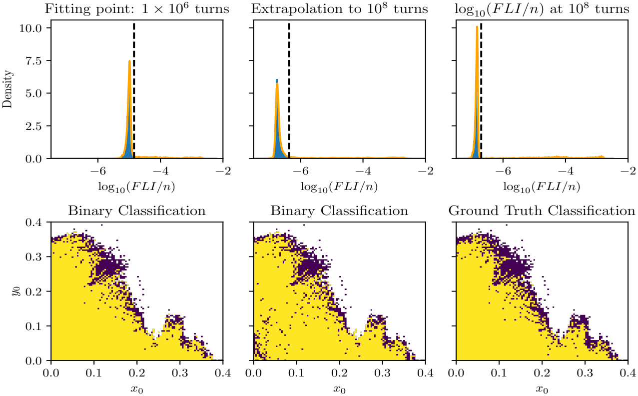
$\log_{10}(\text{REM})$ vs $\log_{10}(\text{FLI}/n)$ at 10^8 turns



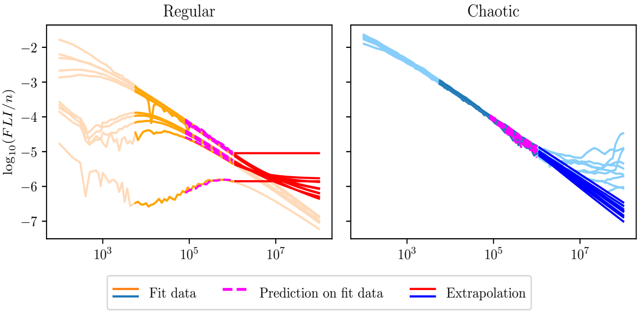


Using logarithmic scale clarifies trend for ARIMA models and reduces number of data points

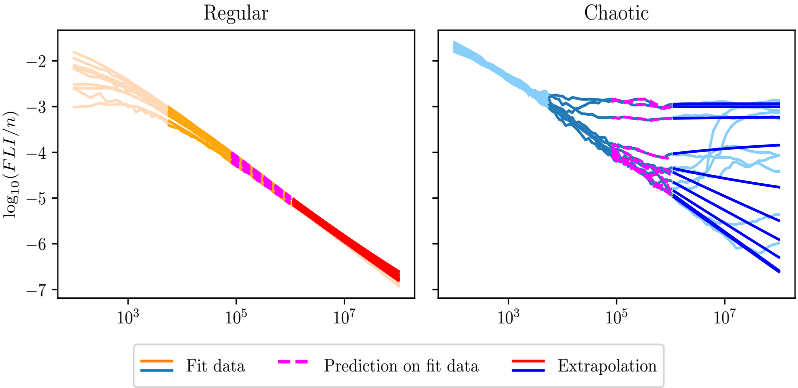
FLI Extrapolation from 10^6 turns



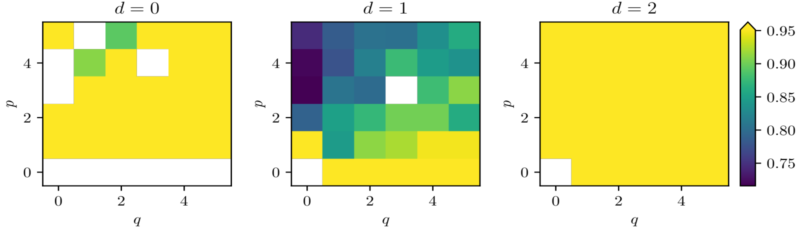
FLI Extrapolation from 10^6 turns



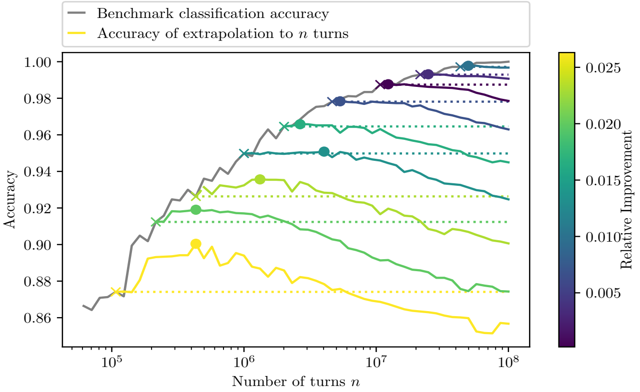
FLI Extrapolation from 10^6 turns



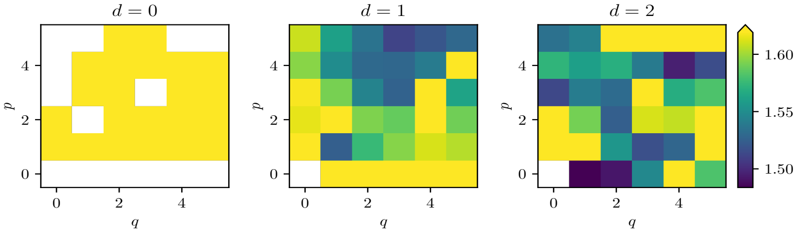
FLI generated with Tangent Map method



FLI generated with Tangent Map method



REM parameter scan



REM extrapolation from 2×10^5 turns