UNIVERSITY OF CAMBRIDGE

Physics-Informed Machine Learning for populational inverse problems

Yann Divet

Prof. Mark Girolami
Dr. Arnaud Vadeboncoeur

IIB Project Final Presentation **6th June 2025**

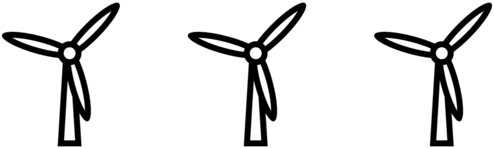
Presentation Outline

Presentation Outline

Motivation

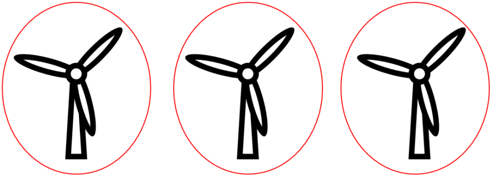
- Inverse problem: Model + Observations $\rightarrow$ Parameters
- Common to build models that use data from different systems
- Compare and contrast ML methods:
  - Hierarchical Bayes
  - Distribution-matching

Hierarchical problem structure



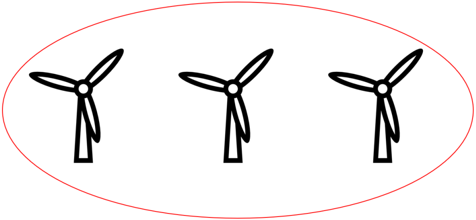
- Population has N = 3 different physical systems
- How do we learn their characteristics?

Fully Pooled Model leads to overfitting



Fully Pooled
(every windmill has separate model)

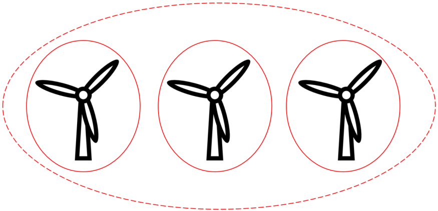
Unpooled Model leads to underfitting



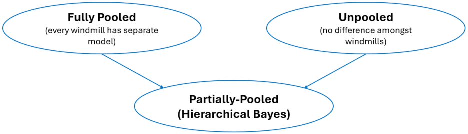
Fully Pooled
(every windmill has separate model)

Unpooled
(no difference amongst windmills)

Partially Pooled (or Hierarchical Bayes) Model



- Solves overfitting and underfitting problems
- Each wind turbine has individual parameters influenced by population distribution



Presentation Outline

Damped Harmonic Oscillator Problem

- Damped Harmonic Oscillator equation:

$$\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + \omega_0^2 x = 0$$

where:

  - x corresponds to the displacement,
  - the undamped angular frequency of the oscillator ω_0,
  - $\gamma = 2\zeta\omega_0$, where ζ is the damping ratio.

Multi-layered structure

$$\mathbf{y}^{(n)} = \mathcal{G}(\mathbf{z}^{(n)}) + \epsilon^{(n)},\ \epsilon \sim \eta := \mathcal{N}(0, \sigma^2 I)$$

$$\mathbf{z}^{(n)} \sim \text{LogNormal}(\boldsymbol{\mu}, \boldsymbol{\tau})^2$$

$$\begin{array}{ccccc}
 & & \mu, \tau & & \\
 & \swarrow & \downarrow & \searrow & \\
z_1 & z_2 & z_3 & \cdots & z_N \\
\downarrow & \downarrow & \downarrow & \cdots & \downarrow \\
y_1 & y_2 & y_3 & \cdots & y_N
\end{array}$$

- Observations
- Populations
- Hyperparameter

Method 1: Hierarchical Bayes

- Explore posterior, an update of our prior beliefs of the hyperparameters and population-level parameters using the likelihood of our observed data:

$$p(\mathbf{z}^{(1:N)}, \boldsymbol{\alpha} | \mathbf{y}^{(1:N)}) \propto \prod_{i=1}^{N} p(\mathbf{y}^{(1:N)} | \mathbf{z}^{(1:N)}) p(\mathbf{z}^{(1:N)} | \boldsymbol{\alpha})$$

- Using Markov Chain Monte Carlo to sample from the posterior distribution

Method 2: Distribution-matching

- Obtain optimal α parameter that minimises objective function:

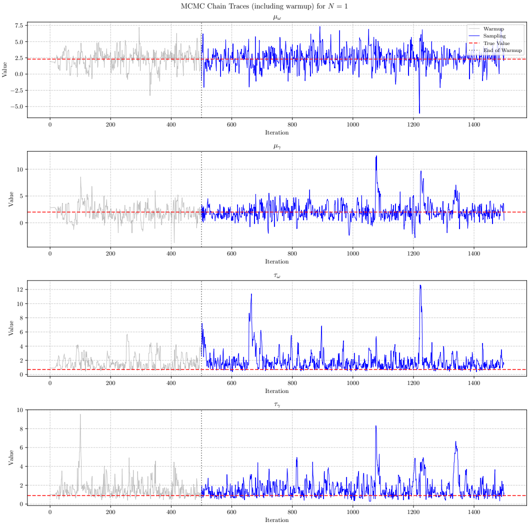
$$\alpha^* = \arg\min_{\alpha} J(\alpha)$$

- Objective = Data-fit term + Regularisation term

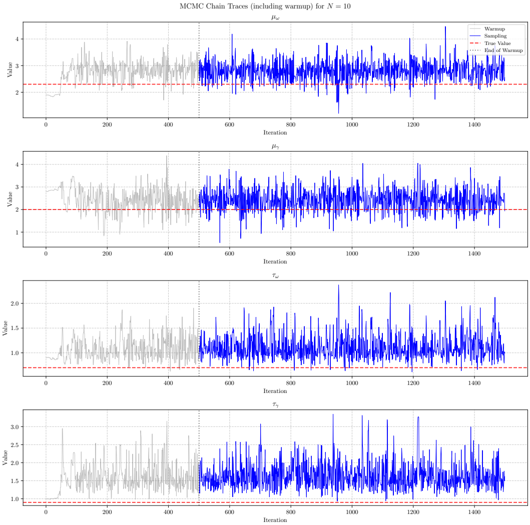
$$J(\alpha) = d_1(\nu, \eta * G_\# P^\alpha(z)) + d_2(P^\alpha(z), P^0(z))$$

Presentation Outline

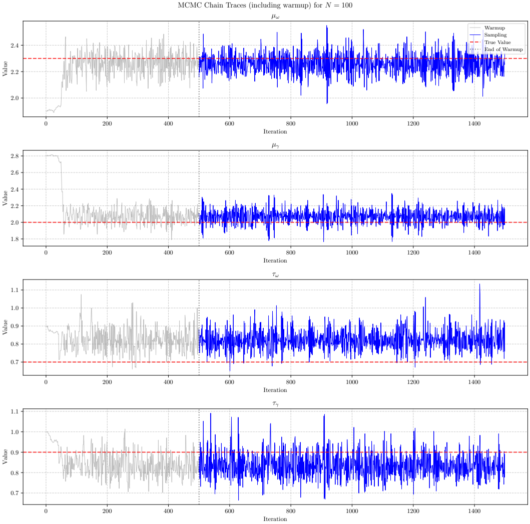
Hierarchical Bayes (N = 1)



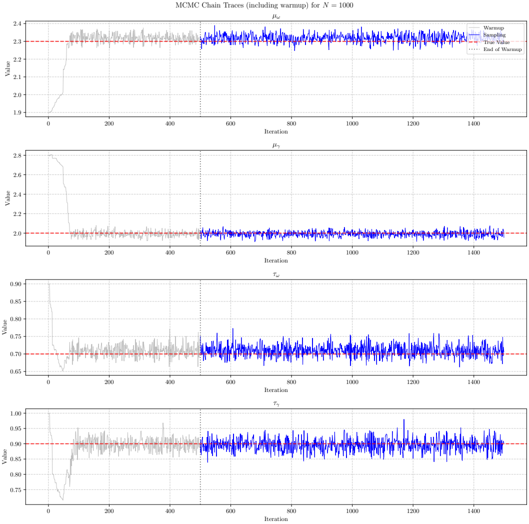
Hierarchical Bayes (N = 10)



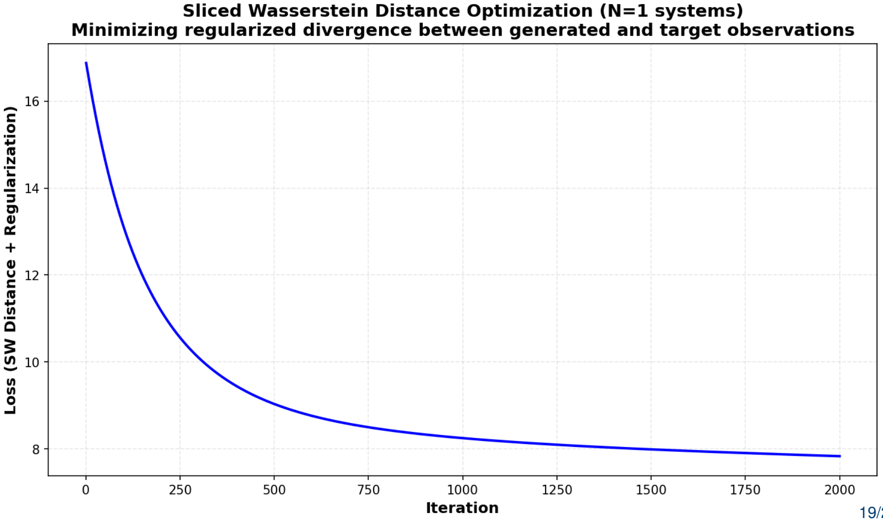
Hierarchical Bayes (N = 100)



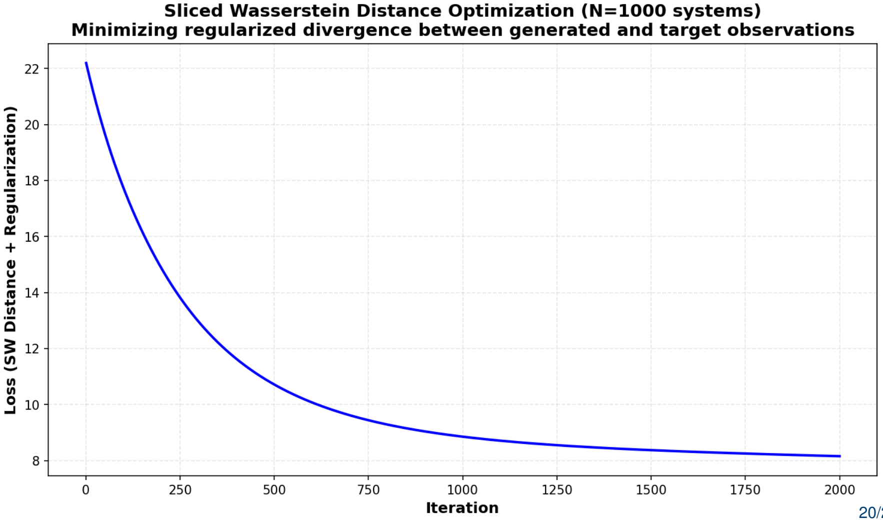
Hierarchical Bayes (N = 1000)



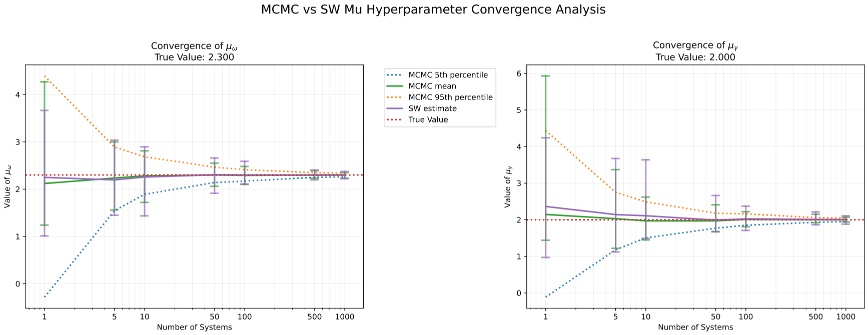
Distribution-matching (N = 1)



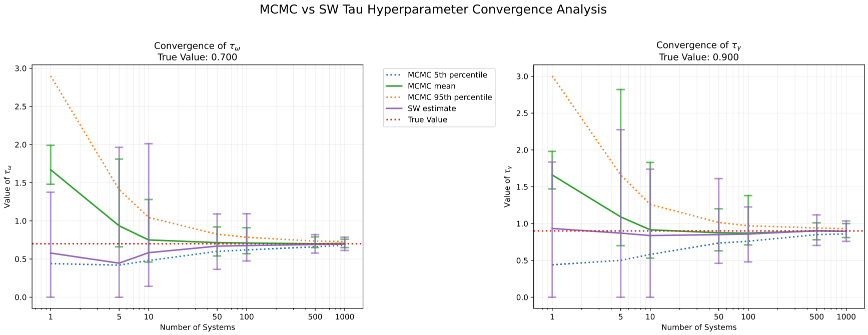
Distribution-matching (N = 1000)



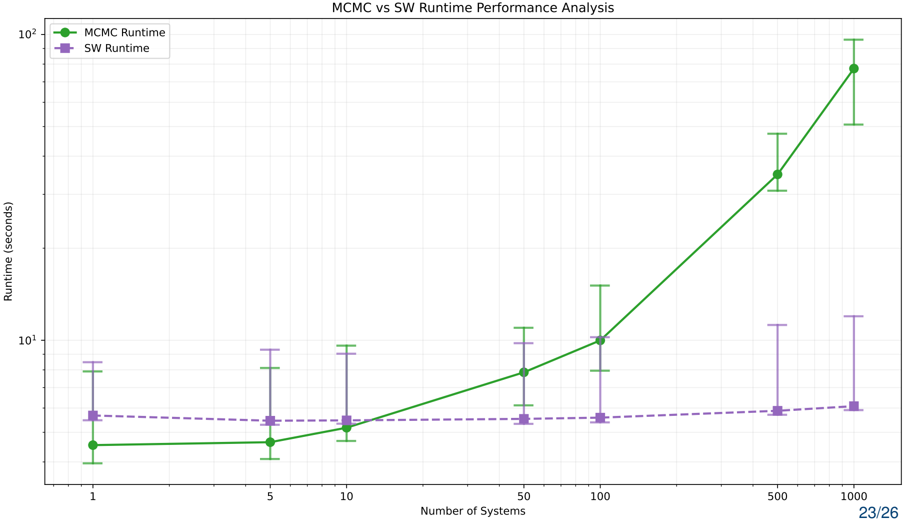
Accuracy Comparison (Hyperparameters mean)



Accuracy Comparison (Hyperparameters std)



Runtime Comparison



Presentation Outline

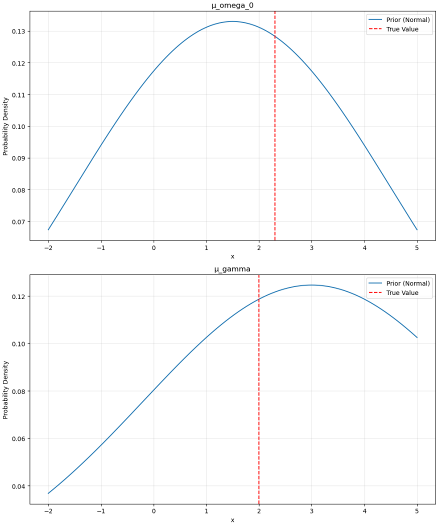
Summary

- Hierarchical Bayes: uncertainty quantification
- Distribution-matching: better computational efficiency
- **HBM when number of physical systems is smaller**
- **DM when number of physical systems is larger**
- Generalise methodology to other elastic-dynamic problems

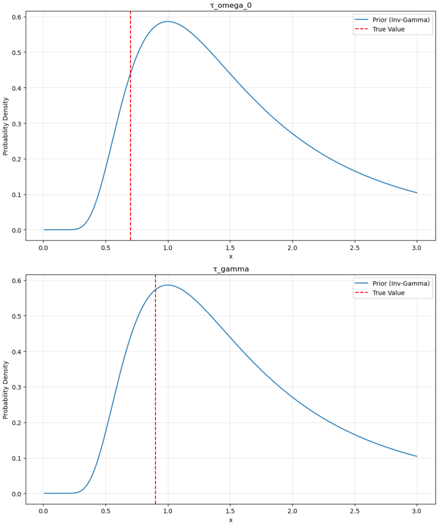
Main References

- A. Gelman, J. B. Carlin, H. S. Stern, & D. B. Rubin. *Bayesian Data Analysis*. 2013.
- A. Gelman. *Multilevel (Hierarchical) Modeling: What It Can and Cannot Do*. 2006.
- M. Betancourt, M. Girolami. *Hamiltonian Monte Carlo for Hierarchical Models*. 2013.
- M. Hoffman, A. Gelman, et al. *The No-U-Turn sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo*. 2014.
- O. Deniz Akyildiz, M. Girolami, A. M. Stuart & A. Vadeboncoeur. *Efficient Prior Calibration From Indirect Data*. 2024.

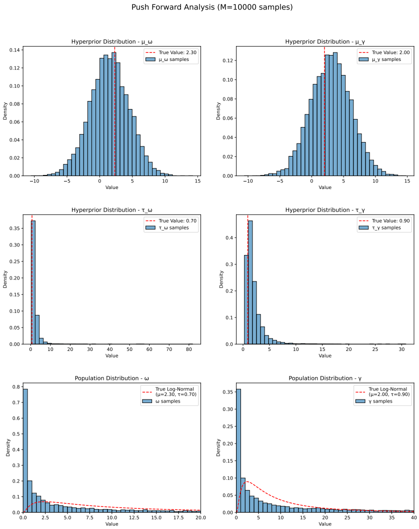
Hyperpriors (mean)



Hyperpriors (std)



Pushforward - Prior on populational level

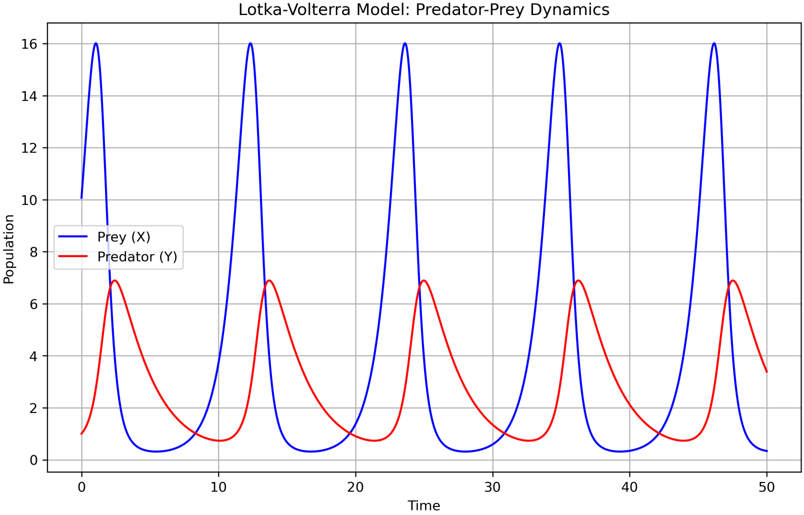


Lotka-Volterra Equation

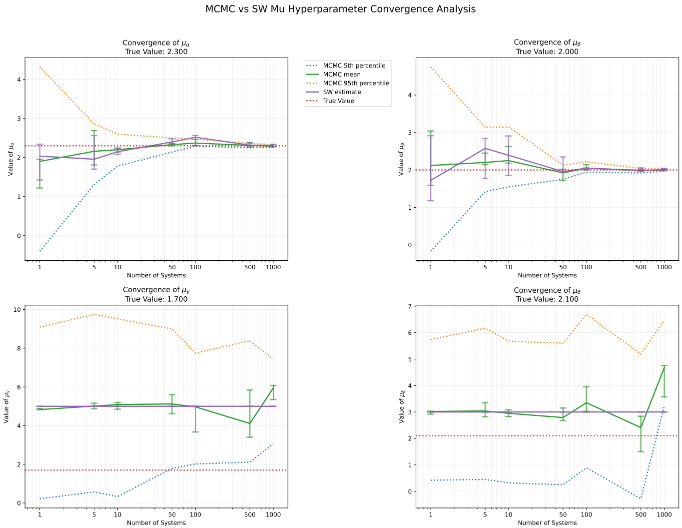
$$\begin{aligned} \frac{dx}{dt} &= \alpha x - \beta xy \\ \frac{dy}{dt} &= \delta xy - \gamma y \end{aligned} \tag{1}$$

Example with $\alpha = 1.1$, $\beta = 0.4$, $\gamma = 0.1$, $\delta = 0.4$ on next slide

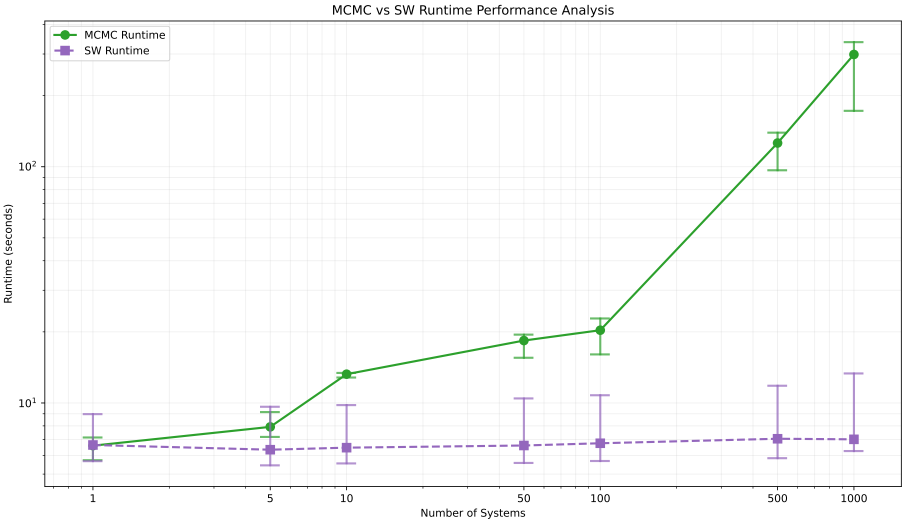
Lotka-Volterra Model



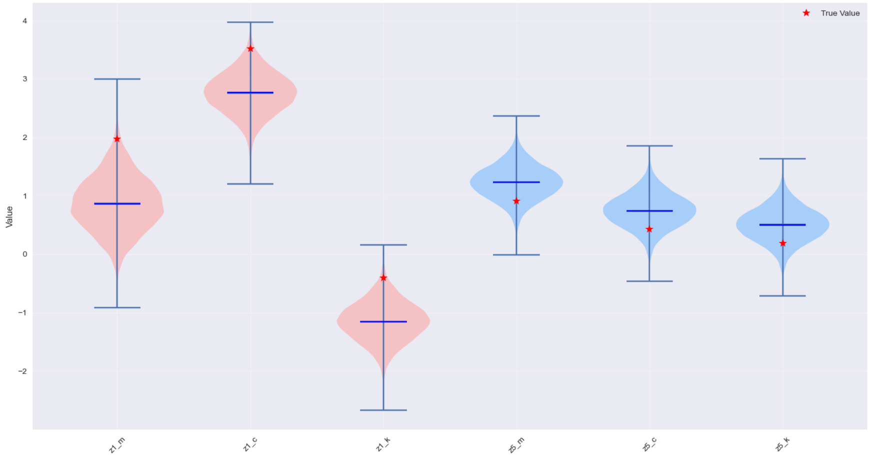
Lotka-Volterra - Accuracy



Lotka-Volterra - Runtime



Distribution of populational parameters (HBM)



Analogy with other systems

Translational mechanical	**Rotational mechanical**	**Series RLC circuit**	**Parallel RLC circuit**
Position x	Angle θ	Charge q	Flux linkage φ
Velocity $\frac{dx}{dt}$	Angular velocity $\frac{d\theta}{dt}$	Current $\frac{dq}{dt}$	Voltage $\frac{d\varphi}{dt}$
Mass m	Moment of inertia I	Inductance L	Capacitance C
Momentum p	Angular momentum L	Flux linkage φ	Charge q
Spring constant k	Torsion constant μ	Elastance $1/C$	Magnetic reluctance $1/L$
Damping c	Rotational friction Γ	Resistance R	Conductance $G = 1/R$
Drive force $F(t)$	Drive torque $\tau(t)$	Voltage v	Current i
Undamped resonant frequency f_n:			
$\frac{1}{2\pi}\sqrt{\frac{k}{m}}$	$\frac{1}{2\pi}\sqrt{\frac{\mu}{I}}$	$\frac{1}{2\pi}\sqrt{\frac{1}{LC}}$	$\frac{1}{2\pi}\sqrt{\frac{1}{LC}}$
Damping ratio ζ:			
$\frac{c}{2}\sqrt{\frac{1}{km}}$	$\frac{\Gamma}{2}\sqrt{\frac{1}{I\mu}}$	$\frac{R}{2}\sqrt{\frac{C}{L}}$	$\frac{G}{2}\sqrt{\frac{L}{C}}$
Differential equation:			
$m\ddot{x} + c\dot{x} + kx = F$	$I\ddot{\theta} + \Gamma\dot{\theta} + \mu\theta = \tau$	$L\ddot{q} + R\dot{q} + q/C = v$	$C\ddot{\varphi} + G\dot{\varphi} + \varphi/L = i$