# Adaptive moment closure for parameter inference of biochemical reaction networks

**Christian Schilling**, Sergiy Bogomolov, Jakob Ruess,
Andreas Podelski Thomas A. Henzinger

University of Freiburg, Germany IST Austria, Austria

CMSB 2015 — September 16th, 2015 — Nantes, France

# Road Map

**Problems**

**Parameter inference**

**Likelihood computation**

**Approximation choice**

**Solutions**

**MCMC algorithm**

**Moment closure approximation**

**Adaptive algorithm**

# Road Map

**Problems**

**Solutions**

**Parameter inference**

**MCMC algorithm**

**Likelihood computation**

**Moment closure approximation**

**Approximation choice**

**Adaptive algorithm**

# Road Map

**Problems**

$\boxed{\textbf{Parameter inference}}$

**Likelihood computation**

**Approximation choice**

**Solutions**

**MCMC algorithm**

**Moment closure approximation**

**Adaptive algorithm**

# Parameter inference

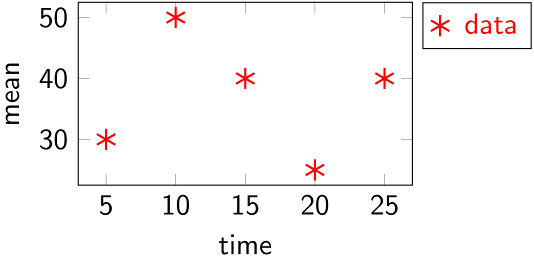
Given

- **Parametric model** of **chemical reaction network** with unknown rate **parameters $\theta$**

$$\nu'_{1,k}X_1 + \cdots + \nu'_{m,k}X_m \xrightarrow{\theta_k} \nu''_{1,k}X_1 + \cdots + \nu''_{m,k}X_m$$

  Here: **continuous-time Markov chain** (CTMC)
  **Model prediction $M(\theta)$** (output): **mean** and **variance**

- **Data $\mu$**
  Here: **mean** and **variance** of population measurements



# Parameter inference

Given

- **Parametric model** of **chemical reaction network** with unknown rate **parameters** $\theta$

$$\nu'_{1,k}X_1 + \cdots + \nu'_{m,k}X_m \xrightarrow{\theta_k} \nu''_{1,k}X_1 + \cdots + \nu''_{m,k}X_m$$

  Here: **continuous-time Markov chain** (CTMC)
  **Model prediction** $M(\theta)$ (output): **mean** and **variance**

- **Data** $\mu$
  Here: **mean** and **variance** of population measurements

Goal (Maximum likelihood estimation problem)

- Find **parameters** $\theta^*$ maximizing **likelihood** $\mathcal{L}$ between **data** and **model prediction**

$$\theta^* = \arg\max_{\theta} \mathcal{L}(\mu, M(\theta))$$

# Parameter inference

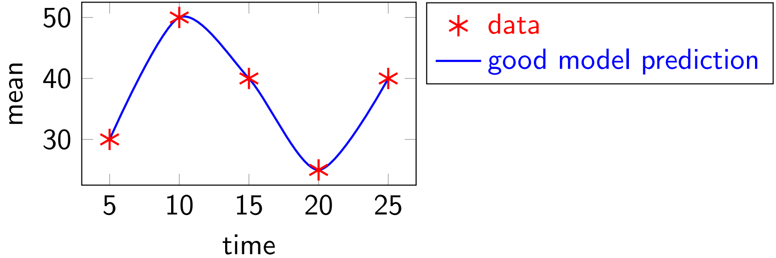
Given

- **Parametric model** of **chemical reaction network** with unknown rate **parameters $\theta$**

$$\nu'_{1,k}X_1 + \cdots + \nu'_{m,k}X_m \xrightarrow{\theta_k} \nu''_{1,k}X_1 + \cdots + \nu''_{m,k}X_m$$

  Here: **continuous-time Markov chain** (CTMC)
  **Model prediction $M(\theta)$** (output): **mean** and **variance**

- **Data $\mu$**
  Here: **mean** and **variance** of population measurements



# Parameter inference

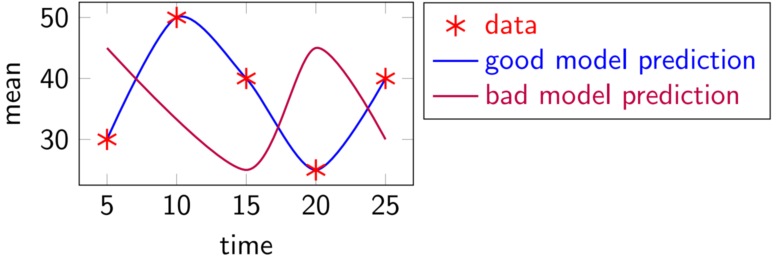
Given

- **Parametric model** of **chemical reaction network** with unknown rate **parameters $\theta$**

$$\nu'_{1,k}X_1 + \cdots + \nu'_{m,k}X_m \xrightarrow{\theta_k} \nu''_{1,k}X_1 + \cdots + \nu''_{m,k}X_m$$

  Here: **continuous-time Markov chain** (CTMC)
  **Model prediction $M(\theta)$** (output): **mean** and **variance**

- **Data $\mu$**
  Here: **mean** and **variance** of population measurements



# Road Map

**Problems**

**Parameter inference**

**Likelihood computation**

**Approximation choice**

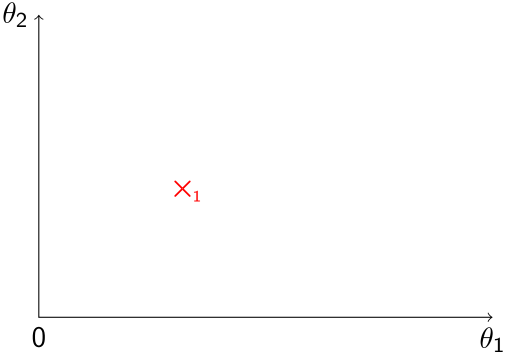
**Solutions**

**MCMC algorithm**

**Moment closure approximation**

**Adaptive algorithm**

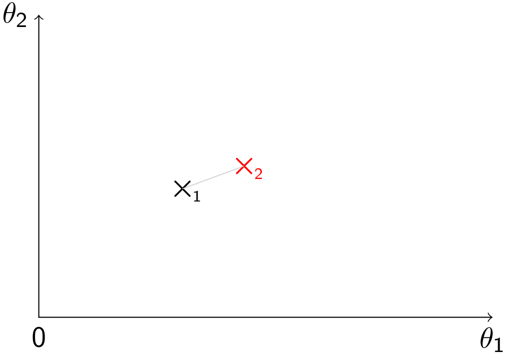
# Markov chain Monte Carlo (MCMC) algorithm




Start at **initial parameters** and repeat:

1. **Compute likelihood** for **current parameters**
2. **Propose new parameters** (based on step 1)

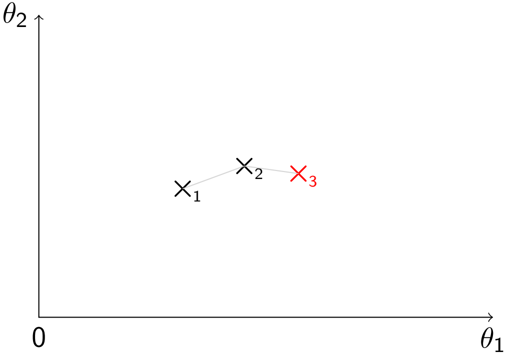
# Markov chain Monte Carlo (MCMC) algorithm




Start at **initial parameters** and repeat:

1. **Compute likelihood** for **current parameters**
2. **Propose** **new parameters** (based on step 1)

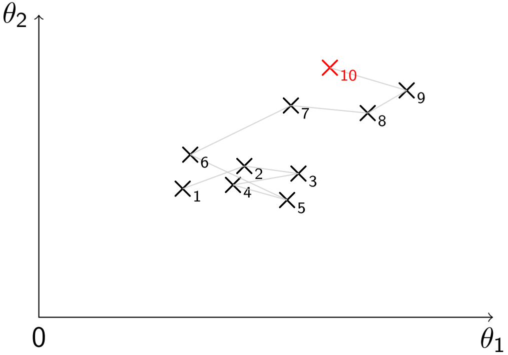
# Markov chain Monte Carlo (MCMC) algorithm




Start at **initial parameters** and repeat:

1. **Compute likelihood** for **current parameters**
2. **Propose new parameters** (based on step 1)

# Markov chain Monte Carlo (MCMC) algorithm




Start at **initial parameters** and repeat:

1. **Compute likelihood** for **current parameters**
2. **Propose new parameters** (based on step 1)

# Road Map

**Problems**

**Parameter inference**

**Likelihood computation**

**Approximation choice**

**Solutions**

**MCMC algorithm**

**Moment closure approximation**

**Adaptive algorithm**

# Likelihood computation

**Likelihood** given via **chemical master equation** (CME):

$$\frac{\partial}{\partial t} p(x,t) = -p(x,t) \sum_{k=1}^{K} a_k(x,\theta) + \sum_{k=1}^{K} p(x-\nu_k,t) a_k(x-\nu_k,\theta)$$

But: Usually **impossible to compute directly**

Approximations

1. Apply a **stochastic simulation algorithm** (SSA)
2. Solve a system of **moment equations** (ODEs)

# Likelihood computation

**Likelihood** given via **chemical master equation** (CME):

$$\frac{\partial}{\partial t}p(x,t) = -p(x,t)\sum_{k=1}^{K} a_k(x,\theta) + \sum_{k=1}^{K} p(x-\nu_k,t)a_k(x-\nu_k,\theta)$$

But: Usually **impossible to compute directly**

## Approximations

1. Apply a **stochastic simulation algorithm** (SSA)
2. Solve a system of **moment equations** (ODEs)

# Road Map

**Problems**

**Parameter inference**

**Likelihood computation**

**Approximation choice**

**Solutions**

**MCMC algorithm**

**Moment closure approximation**

**Adaptive algorithm**

# Moment closure approximation (MCA)

We can derive a **system of moment equations** from the CME

$$\frac{\partial}{\partial t}\boldsymbol{\nu(t)} = A(\theta)\boldsymbol{\nu(t)} + B(\theta)\bar{\boldsymbol{\nu}}\boldsymbol{(t)}$$

where $\bar{\boldsymbol{\nu}}\boldsymbol{(t)}$ depends on moments of higher order

Apply **moment closure approximation** (MCA)

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{\nu}}\boldsymbol{(t)} = A(\theta)\tilde{\boldsymbol{\nu}}\boldsymbol{(t)} + B(\theta)\varphi(\tilde{\boldsymbol{\nu}}\boldsymbol{(t)})$$

where $\tilde{\boldsymbol{\nu}}\boldsymbol{(t)}$ approximates $\boldsymbol{\nu(t)}$

**Precision** of the approximation and **computational complexity** to solve the ODE system determined by the **closure order** and the **choice of function** $\varphi$

# Moment closure approximation (MCA)

We can derive a **system of moment equations** from the CME

$$\frac{\partial}{\partial t}\boldsymbol{\nu(t)} = A(\theta)\boldsymbol{\nu(t)} + B(\theta)\boldsymbol{\bar{\nu}(t)}$$

where $\boldsymbol{\bar{\nu}(t)}$ depends on moments of higher order

Apply **moment closure approximation** (MCA)

$$\frac{\partial}{\partial t}\boldsymbol{\tilde{\nu}(t)} = A(\theta)\boldsymbol{\tilde{\nu}(t)} + B(\theta)\varphi(\boldsymbol{\tilde{\nu}(t)})$$

where $\boldsymbol{\tilde{\nu}(t)}$ approximates $\boldsymbol{\nu(t)}$

**Precision** of the approximation and **computational complexity** to solve the ODE system determined by the **closure order** and the **choice of function** $\varphi$

# Moment closure approximation (MCA)

We can derive a **system of moment equations** from the CME

$$\frac{\partial}{\partial t}\boldsymbol{\nu(t)} = A(\theta)\boldsymbol{\nu(t)} + B(\theta)\boldsymbol{\bar{\nu}(t)}$$

where $\boldsymbol{\bar{\nu}(t)}$ depends on moments of higher order

Apply **moment closure approximation** (MCA)

$$\frac{\partial}{\partial t}\boldsymbol{\tilde{\nu}(t)} = A(\theta)\boldsymbol{\tilde{\nu}(t)} + B(\theta)\varphi(\boldsymbol{\tilde{\nu}(t)})$$

where $\boldsymbol{\tilde{\nu}(t)}$ approximates $\boldsymbol{\nu(t)}$

**Precision** of the approximation and **computational complexity** to solve the ODE system determined by the **closure order** and the **choice of function** $\varphi$

# Road Map

**Problems**

**Solutions**

**Parameter inference**

**MCMC algorithm**

**Likelihood computation**

**Moment closure approximation**

**Approximation choice**

**Adaptive algorithm**

# Approximation choice

## Observation 1

The **approximation precision** depends on the **model**

Consider an evaluation of three **MCAs** in three **parameter regions**

$(+ = \text{good}, - = \text{bad})$

| MCA | region 1 | region 2 | region 3 |
|---|---|---|---|
| $MC_1$ | + | − | − |
| $MC_2$ | + | + | − |
| $MC_3$ | + | − | + |

## Observation 2

The **approximation precision** depends on the **parameter region**

# Approximation choice

Observation 1

The **approximation precision** depends on the **model**

Consider an evaluation of three **MCAs** in three **parameter regions**

$(+ = \text{good}, - = \text{bad})$

| MCA | region 1 | region 2 | region 3 |
|---|---|---|---|
| $MC_1$ | $+$ | $-$ | $-$ |
| $MC_2$ | $+$ | $+$ | $-$ |
| $MC_3$ | $+$ | $-$ | $+$ |

Observation 2

The **approximation precision** depends on the **parameter region**

# Approximation choice

## Observation 1

The **approximation precision** depends on the **model**

Consider an evaluation of three **MCAs** in three **parameter regions**

$$(+ = \text{good}, - = \text{bad})$$

| MCA | region 1 | region 2 | region 3 |
| --- | --- | --- | --- |
| $MC_1$ | $+$ | $-$ | $-$ |
| $MC_2$ | $+$ | $+$ | $-$ |
| $MC_3$ | $+$ | $-$ | $+$ |

## Observation 2

The **approximation precision** depends on the **parameter region**

# Road Map

| Problems | Solutions |
| --- | --- |
| Parameter inference | |
| | MCMC algorithm |
| Likelihood computation | |
| | Moment closure approximation |
| Approximation choice | |
| | $\boxed{\text{Adaptive algorithm}}$ |

# Adaptive algorithm for MCA choice

## Iterative MCMC algorithm (recalled)

Start at **initial parameters** and repeat:

1. **Compute likelihood** for **current parameters**
2. **Propose new parameters** (based on step 1)

## Proposed approach for step 1 (MCA choice)

Use the **most precise MCA** for some **particular parameters** $\theta$ for **all likelihood computations** in the **parameter region** of $\theta$

## Rationale

If an MCA is the **most precise** approximation available **at some point**, it should be a **precise** approximation in the **region** around

# Adaptive algorithm for MCA choice

## Iterative MCMC algorithm (recalled)

Start at **initial parameters** and repeat:

1. **Compute likelihood** for **current parameters**
2. **Propose** **new parameters** (based on step 1)

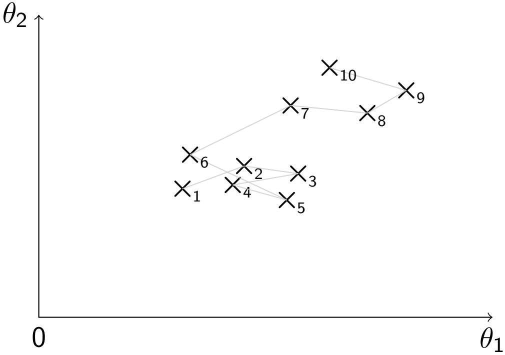
## Proposed approach for step 1 (MCA choice)

Use the **most precise MCA** for some **particular parameters** $\boldsymbol{\theta}$ for **all likelihood computations** in the **parameter region** of $\boldsymbol{\theta}$

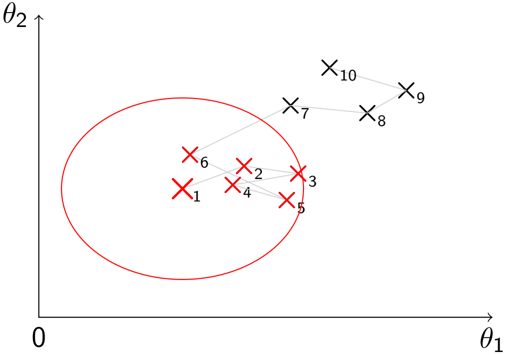
## Rationale

If an MCA is the **most precise** approximation available **at some point**, it should be a **precise** approximation in the **region** around

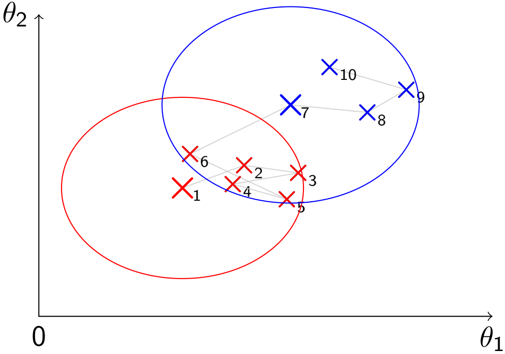
# Re-evaluation during search



# Re-evaluation during search



# Re-evaluation during search



# Pseudocode

### Input

parametric CTMC, data, MCAs, initial parameter choice

1. Run **simulation** (SSA) and **evaluate all MCAs**
   Fix the **most precise MCA** and the **region**
2. **Compute likelihood** for **current parameters**
   using the **current MCA**
3. **Propose new parameters**
4. If **new parameters** are still in the **current region**
   go to step 2 (keep **MCA**)
   else
   go to step 1 (re-evaluate **MCAs**)

# Pseudocode

### Input

parametric CTMC, data, MCAs, initial parameter choice

1. Run **simulation** (SSA) and **evaluate all MCAs**
   Fix the **most precise MCA** and the **region**
2. **Compute likelihood** for **current parameters**
   using the **current MCA**
3. **Propose new parameters**
4. If **new parameters** are still in the **current region**
   go to step 2 (keep **MCA**)
   else
   go to step 1 (re-evaluate **MCAs**)

# Pseudocode

## Input

parametric CTMC, data, MCAs, initial parameter choice

1. Run **simulation** (SSA) and **evaluate all MCAs**
   Fix the **most precise MCA** and the **region**
2. **Compute likelihood** for **current parameters**
   using the **current MCA**
3. **Propose new parameters**
4. If **new parameters** are still in the **current region**
   go to step 2 (keep **MCA**)
   else
   go to step 1 (re-evaluate **MCAs**)

# Pseudocode

## Input

parametric CTMC, data, MCAs, initial parameter choice

1. Run **simulation** (SSA) and **evaluate all MCAs**
   Fix the **most precise MCA** and the **region**
2. **Compute likelihood** for **current parameters**
   using the **current MCA**
3. **Propose new parameters**
4. If **new parameters** are still in the **current region**
   go to step 2 (keep **MCA**)
   else
   go to step 1 (re-evaluate **MCAs**)

# Pseudocode

## Input

parametric CTMC, data, MCAs, initial parameter choice

1. Run **simulation** (SSA) and **evaluate all MCAs**
   Fix the **most precise MCA** and the **region**
2. **Compute likelihood** for **current parameters**
   using the **current MCA**
3. **Propose new parameters**
4. If **new parameters** are still in the **current region**
   go to step 2 (keep **MCA**)
   else
   go to step 1 (re-evaluate **MCAs**)

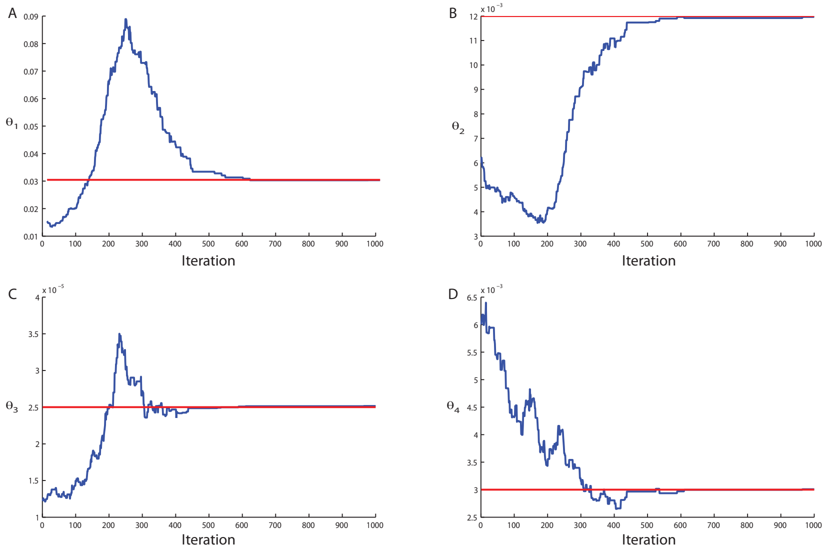
# Evaluation

## Model

- 4 parameters $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$
- 11 MCA choices[1]
- MCMC algorithm run for 1,000 steps
  - One run for the adaptive algorithm
  - One run for each MCA fixed

[1] derivative matching, zero cumulants, zero variance, low dispersion; each of order 2–4 (except low dispersion)

# Adaptive algorithm run for $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$




— true parameters, — proposed parameters

# Evaluation

## Model

- 4 parameters $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$
- 11 MCA choices[1]
- MCMC algorithm run for 1,000 steps
  - One run for the adaptive algorithm
  - One run for each MCA fixed

## Adaptive algorithm statistics

- 23 **MCA evaluations**
- 19 times a **new MCA** was chosen
- 7 **different MCAs** were chosen in total

[1] derivative matching, zero cumulants, zero variance, low dispersion; each of order 2–4 (except low dispersion)

| approximation | relative distance to true parameters [%] | | | | MCA use [%] |
| --- | --- | --- | --- | --- | --- |
| | $\theta_1$ | $\theta_2$ | $\theta_3$ | $\theta_4$ | |
| adaptive | 0.44 | 0.31 | **0.65** | **0.29** | |
| dm2 | 4.45 | 2.74 | 2.68 | 4.32 | 15 |
| zc2 | 11.02 | 6.11 | 3.23 | 2.93 | 10 |
| zv2 | 281.09 | 74.85 | 45.72 | 76.29 | 0 |
| dm3 | 2.54 | 1.23 | 1.85 | 3.55 | 10 |
| zc3 | 9.72 | 4.80 | 0.86 | 2.87 | 5 |
| zv3 | 285.55 | 79.96 | 49.01 | 83.41 | 0 |
| ld3 | 9.08 | 4.30 | 6.75 | 9.63 | 0 |
| dm4 | 3.43 | 1.33 | 4.17 | 9.54 | 15 |
| zc4 | **0.35** | **0.19** | 3.77 | 9.29 | 35 |
| zv4 | 292.60 | 78.89 | 46.60 | 71.90 | 0 |
| ld4 | 14.44 | 3.80 | 12.31 | 28.06 | 10 |

| approximation | relative distance to true parameters [%] $\theta_1$ | $\theta_2$ | $\theta_3$ | $\theta_4$ | MCA use [%] |
|---|---|---|---|---|---|
| adaptive | 0.44 | 0.31 | **0.65** | **0.29** | |
| dm2 | 4.45 | 2.74 | 2.68 | 4.32 | 15 |
| zc2 | 11.02 | 6.11 | 3.23 | 2.93 | 10 |
| zv2 | 281.09 | 74.85 | 45.72 | 76.29 | 0 |
| dm3 | 2.54 | 1.23 | 1.85 | 3.55 | 10 |
| zc3 | 9.72 | 4.80 | 0.86 | 2.87 | 5 |
| zv3 | 285.55 | 79.96 | 49.01 | 83.41 | 0 |
| ld3 | 9.08 | 4.30 | 6.75 | 9.63 | 0 |
| dm4 | 3.43 | 1.33 | 4.17 | 9.54 | 15 |
| zc4 | **0.35** | **0.19** | 3.77 | 9.29 | 35 |
| zv4 | 292.60 | 78.89 | 46.60 | 71.90 | 0 |
| ld4 | 14.44 | 3.80 | 12.31 | 28.06 | 10 |

| approximation | relative distance to true parameters [%] | | | | MCA use [%] |
|---|---|---|---|---|---|
| | $\theta_1$ | $\theta_2$ | $\theta_3$ | $\theta_4$ | |
| adaptive | 0.44 | 0.31 | **0.65** | **0.29** | |
| dm2 | 4.45 | 2.74 | 2.68 | 4.32 | 15 |
| zc2 | 11.02 | 6.11 | 3.23 | 2.93 | 10 |
| ▶ zv2 | 281.09 | 74.85 | 45.72 | 76.29 | 0 |
| dm3 | 2.54 | 1.23 | 1.85 | 3.55 | 10 |
| zc3 | 9.72 | 4.80 | 0.86 | 2.87 | 5 |
| zv3 | 285.55 | 79.96 | 49.01 | 83.41 | 0 |
| ld3 | 9.08 | 4.30 | 6.75 | 9.63 | 0 |
| dm4 | 3.43 | 1.33 | 4.17 | 9.54 | 15 |
| zc4 | **0.35** | **0.19** | 3.77 | 9.29 | 35 |
| zv4 | 292.60 | 78.89 | 46.60 | 71.90 | 0 |
| ld4 | 14.44 | 3.80 | 12.31 | 28.06 | 10 |

| approximation | relative distance to true parameters [%] | | | | MCA use [%] |
|---|---|---|---|---|---|
| | $\theta_1$ | $\theta_2$ | $\theta_3$ | $\theta_4$ | |
| adaptive | 0.44 | 0.31 | **0.65** | **0.29** | |
| dm2 | 4.45 | 2.74 | 2.68 | 4.32 | 15 |
| zc2 | 11.02 | 6.11 | 3.23 | 2.93 | 10 |
| zv2 | 281.09 | 74.85 | 45.72 | 76.29 | 0 |
| dm3 | 2.54 | 1.23 | 1.85 | 3.55 | 10 |
| zc3 | 9.72 | 4.80 | 0.86 | 2.87 | 5 |
| zv3 | 285.55 | 79.96 | 49.01 | 83.41 | 0 |
| ld3 | 9.08 | 4.30 | 6.75 | 9.63 | 0 |
| dm4 | 3.43 | 1.33 | 4.17 | 9.54 | 15 |
| ▶ zc4 | **0.35** | **0.19** | 3.77 | 9.29 | 35 |
| zv4 | 292.60 | 78.89 | 46.60 | 71.90 | 0 |
| ld4 | 14.44 | 3.80 | 12.31 | 28.06 | 10 |

# Region criterion influence

**SSA approach**

**fixed MCA approach**

Observation

- Extreme **region criteria** result in the monolithic approaches
- **Trade-off** between **precision** and **efficiency**

# Region criterion influence


## Observation

- Extreme **region criteria** result in the monolithic approaches
- **Trade-off** between **precision** and **efficiency**

# Conclusion

- **Parameter inference** problem, solved via **MCMC method**
- **Likelihood computation** problem, solved via approximation by **simulation** or **moment closure**
- **Adaptive algorithm** automatically chooses a **good approximation** for the **current parameter region**
- **Leverages** the strengths of two different approaches