Accelerating a multiscale continuum-particle fluid dynamics model with on-the-fly machine learning

David Stephenson James Kermode Duncan Lockerby

EPSRC
Engineering and Physical Sciences Research Council

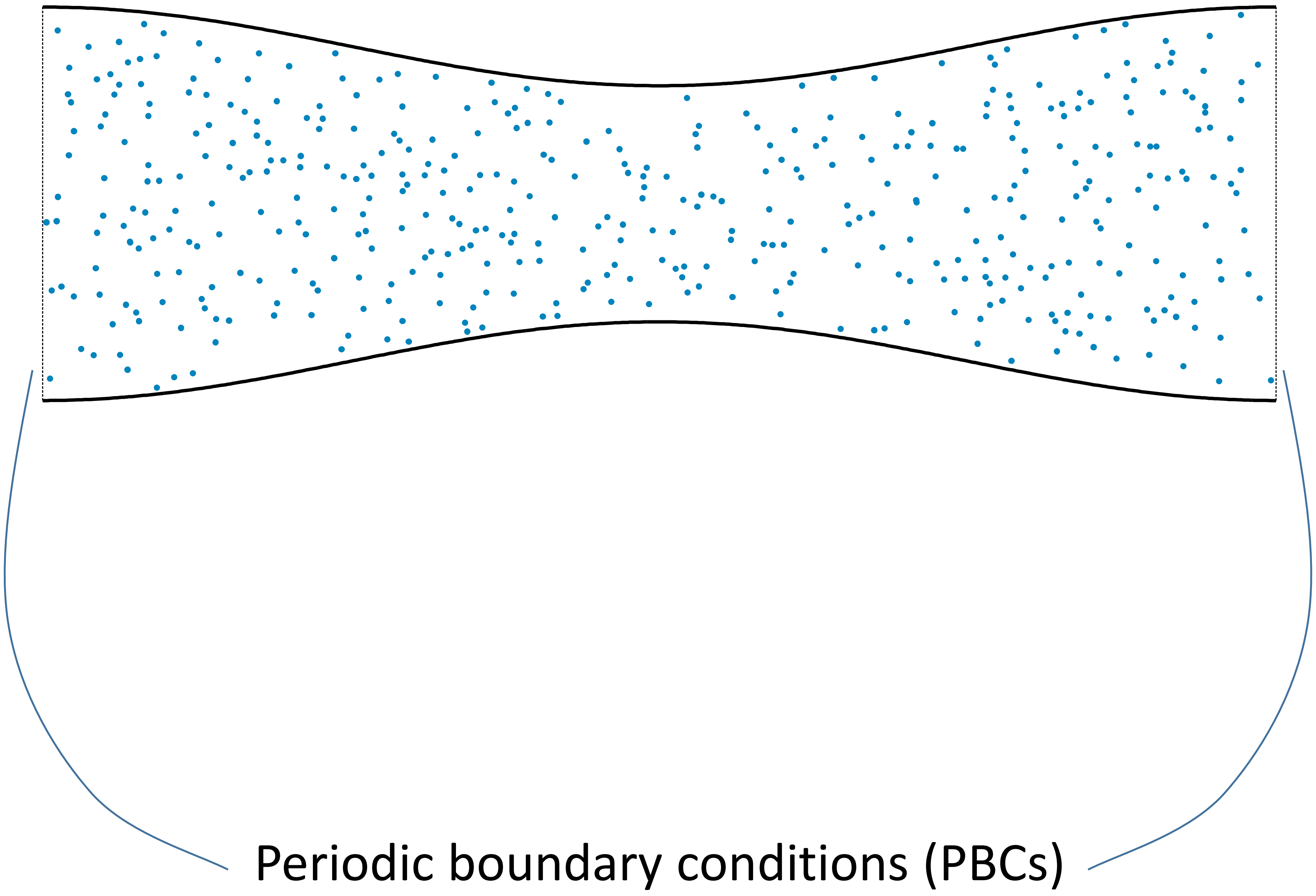
Micro and Nano Flows for Engineering

WARWICK
SCHOOL OF ENGINEERING

The problem with hybrid methods…

…too many similar and repetitious simulations.

Consider a converging-diverging nanochannel:

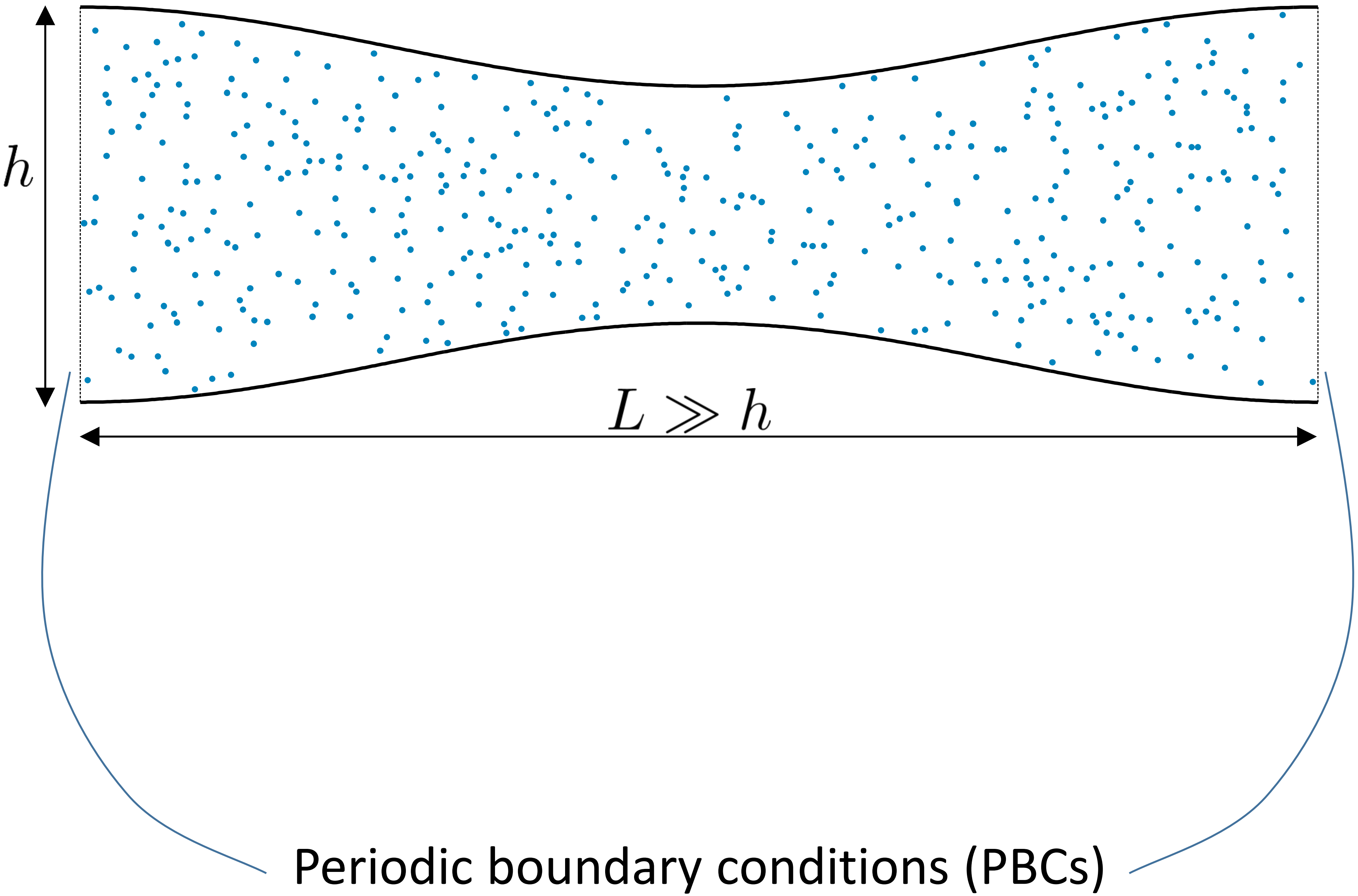


The problem with hybrid methods…

…too many similar and repetitious simulations.

Consider a converging-diverging nanochannel:

- Channel height

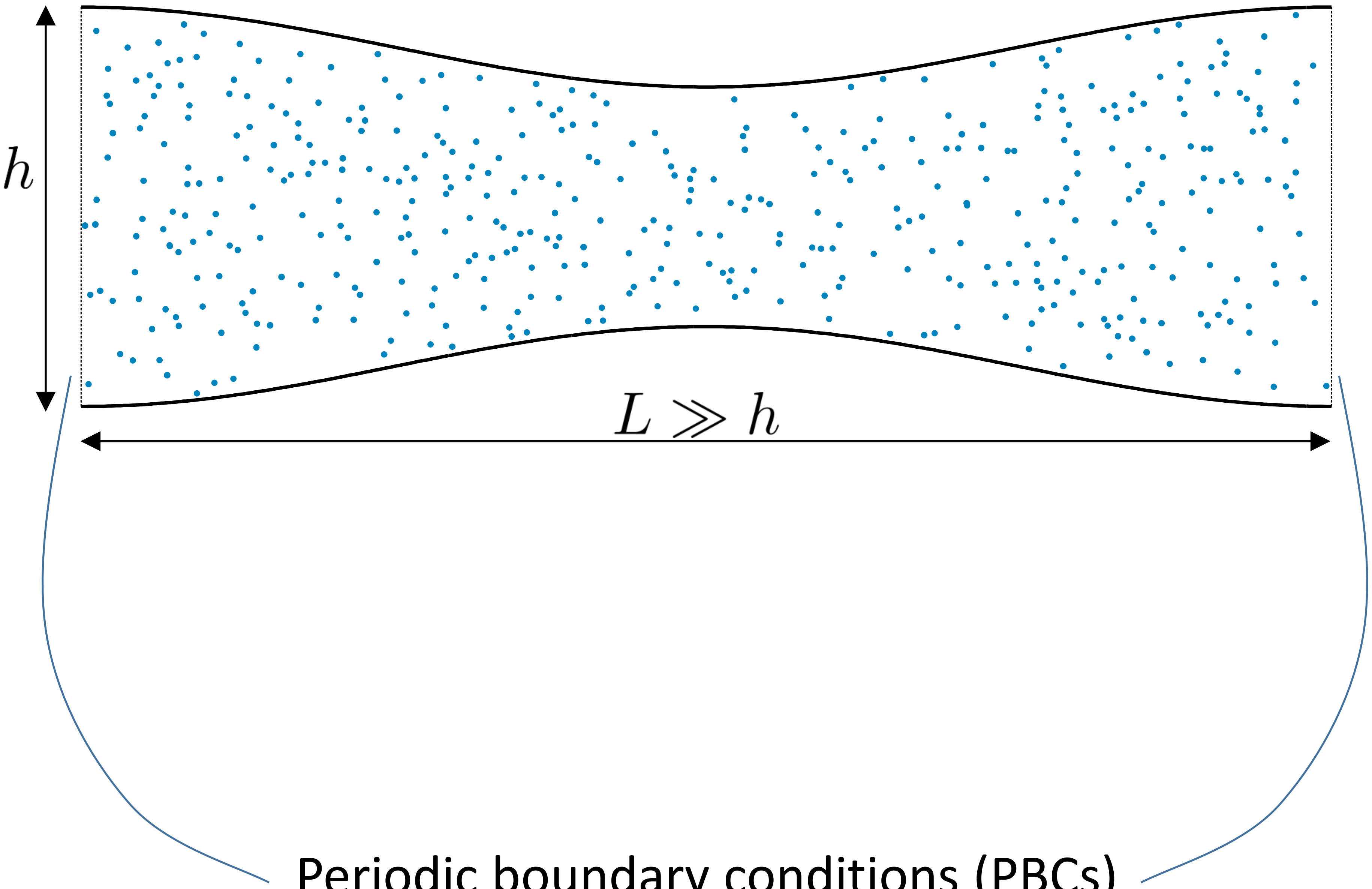


The problem with hybrid methods…

…too many similar and repetitious simulations.

Consider a converging-diverging nanochannel:

- Channel height
- Density

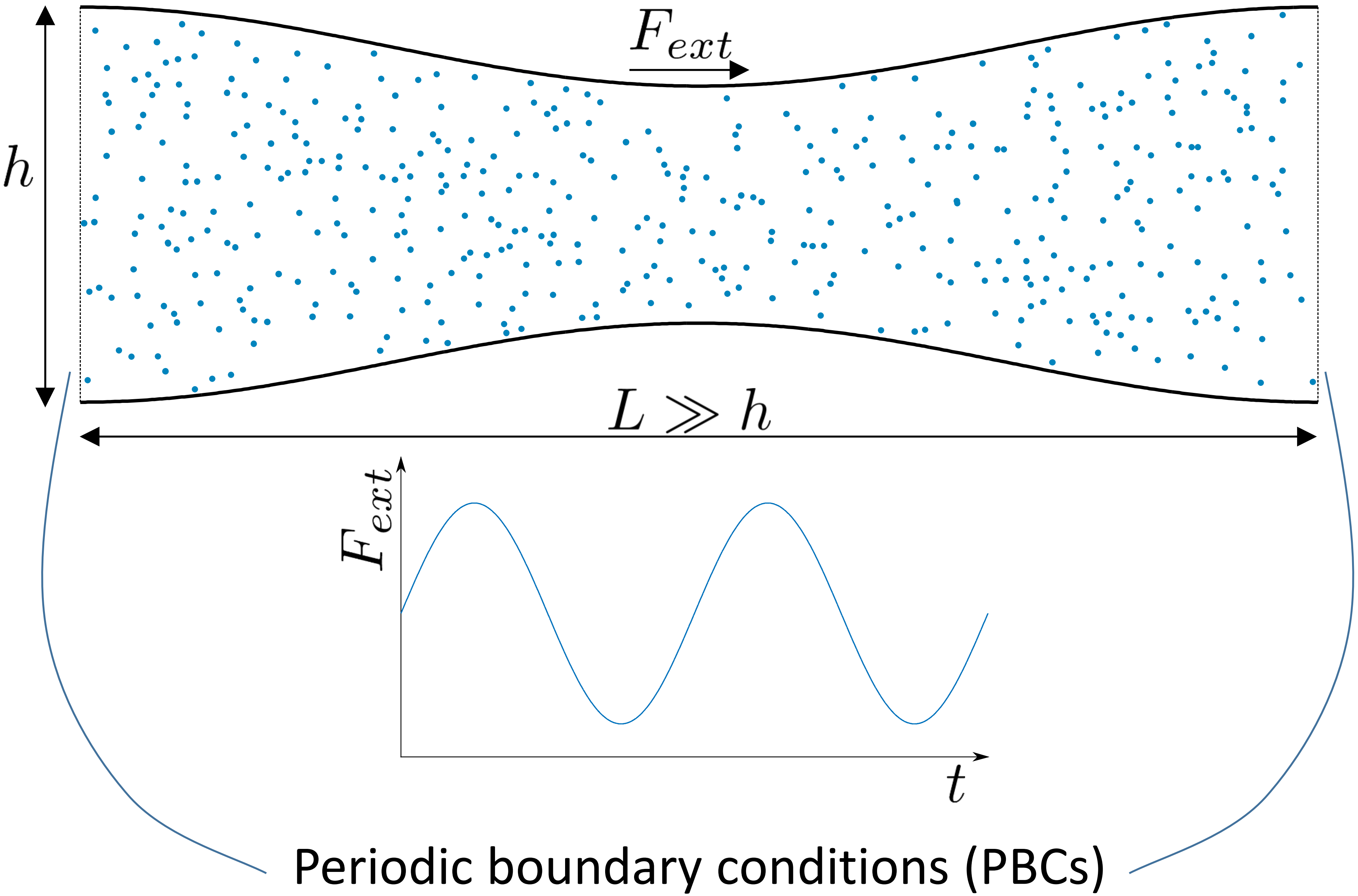


The problem with hybrid methods...

…too many similar and repetitious simulations.

Consider a converging-diverging nanochannel:

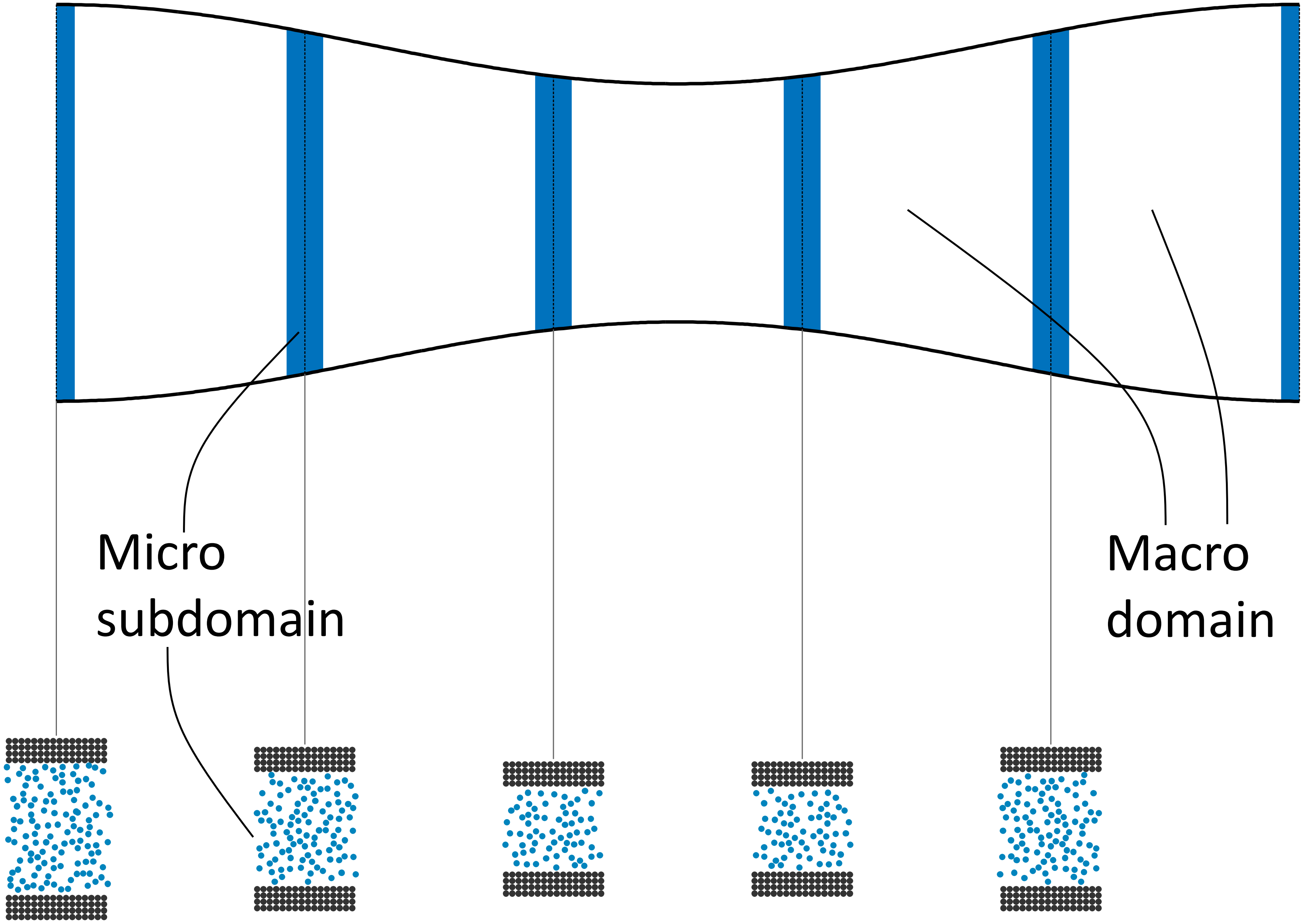
- Channel height
- Density
- Forcing



Hybrid method

We split the macro domain into micro subdomains.

These individual periodic subdomains are simulated using molecular dynamics (MD).



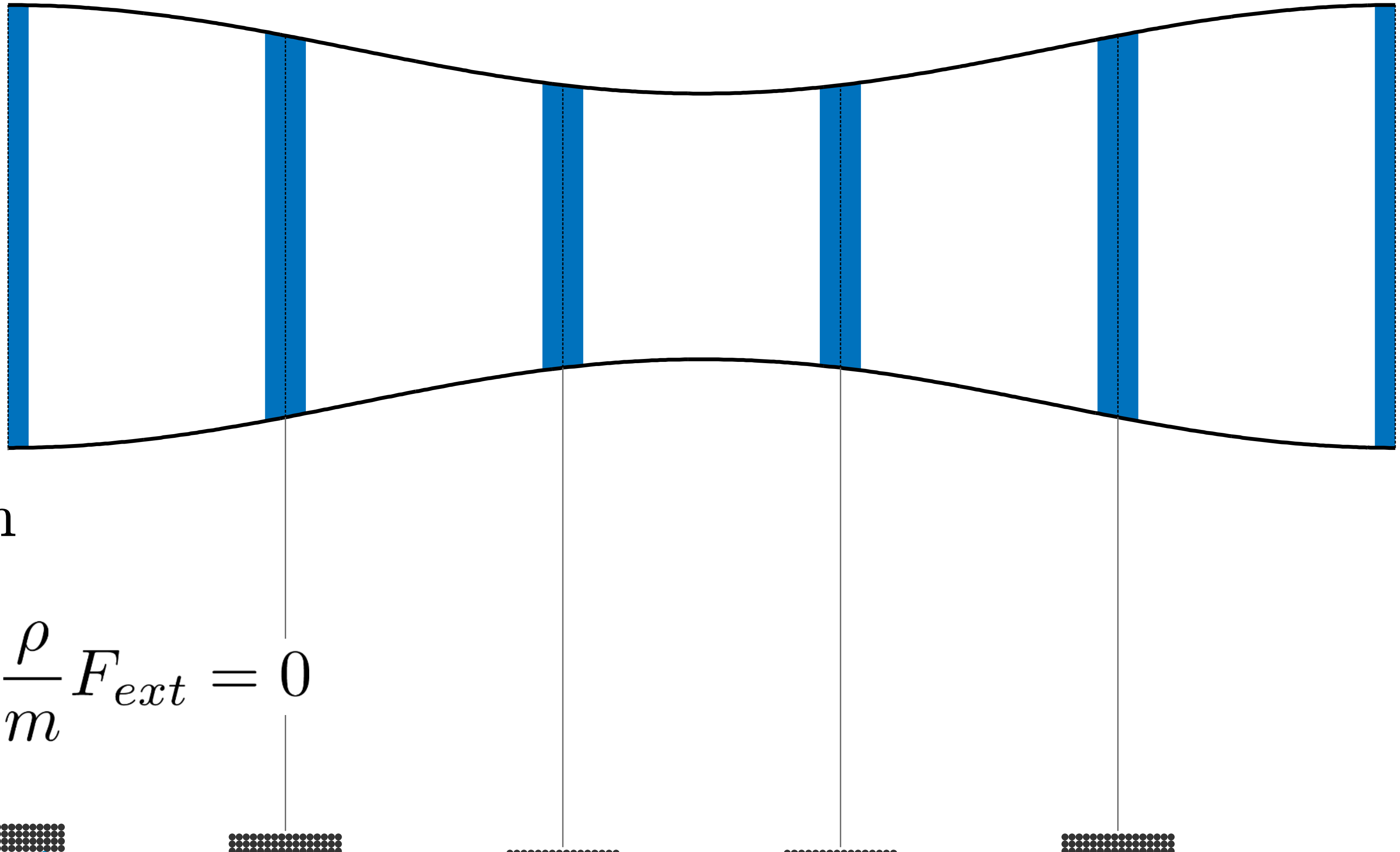
Hybrid method – macro model

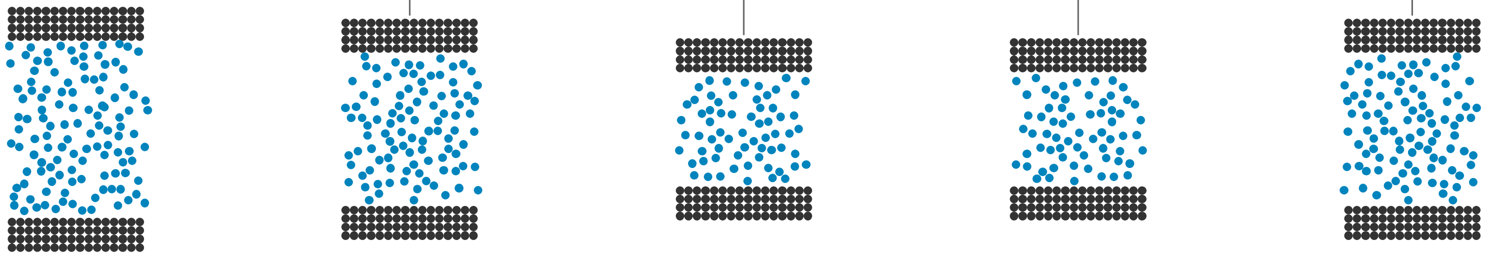
Mass conservation

$$\frac{\partial \rho}{\partial t} + \frac{1}{A}\frac{\partial q}{\partial s} = 0$$

Momentum conservation

$$-\frac{\partial p}{\partial s} + \frac{\partial \tau_{sy}}{\partial y} + \frac{\partial \tau_{sz}}{\partial z} + \frac{\rho}{m}F_{ext} = 0$$



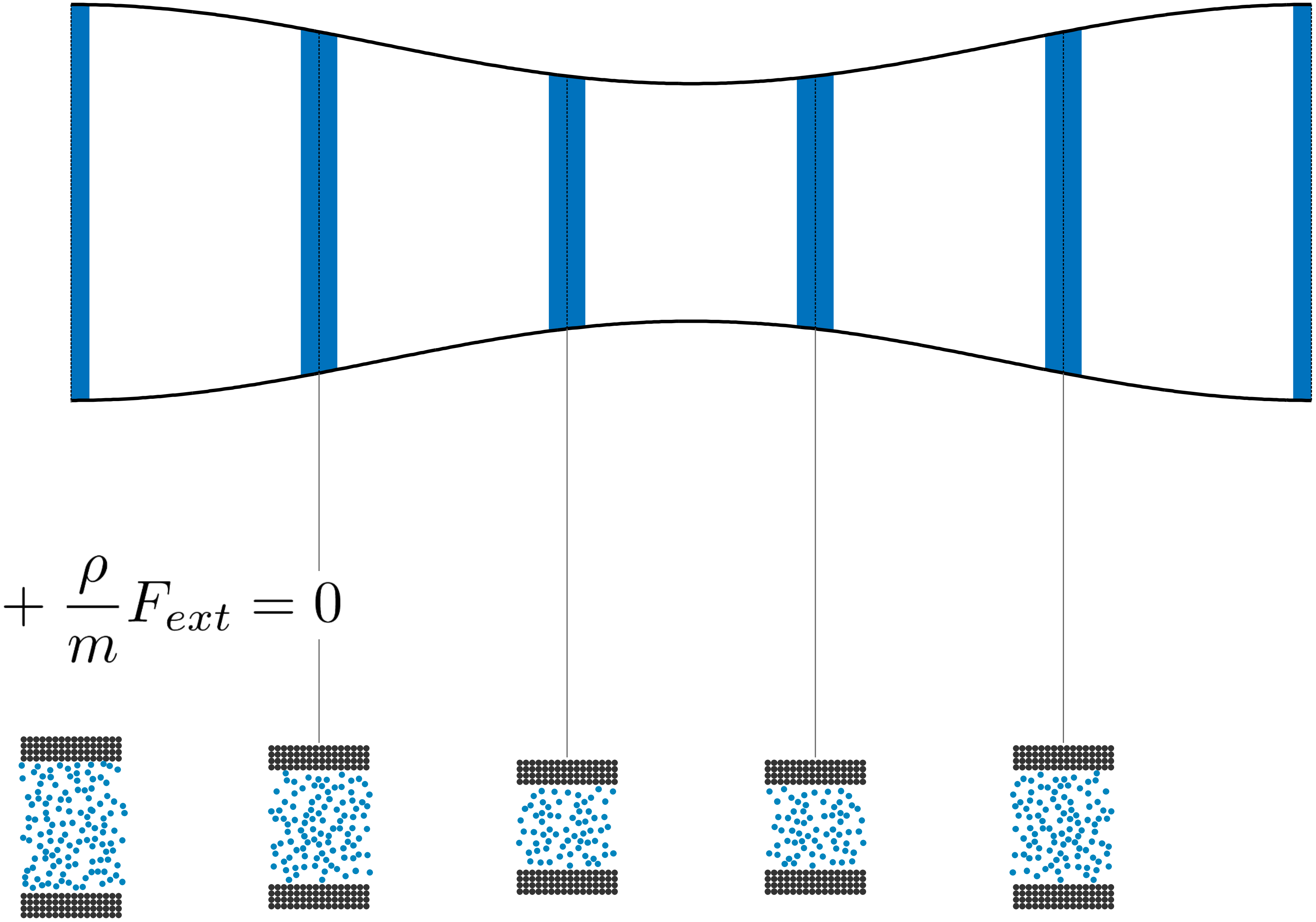


Hybrid method – micro model

MD subdomain simulations conserve mass.

MD cannot support a pressure gradient.

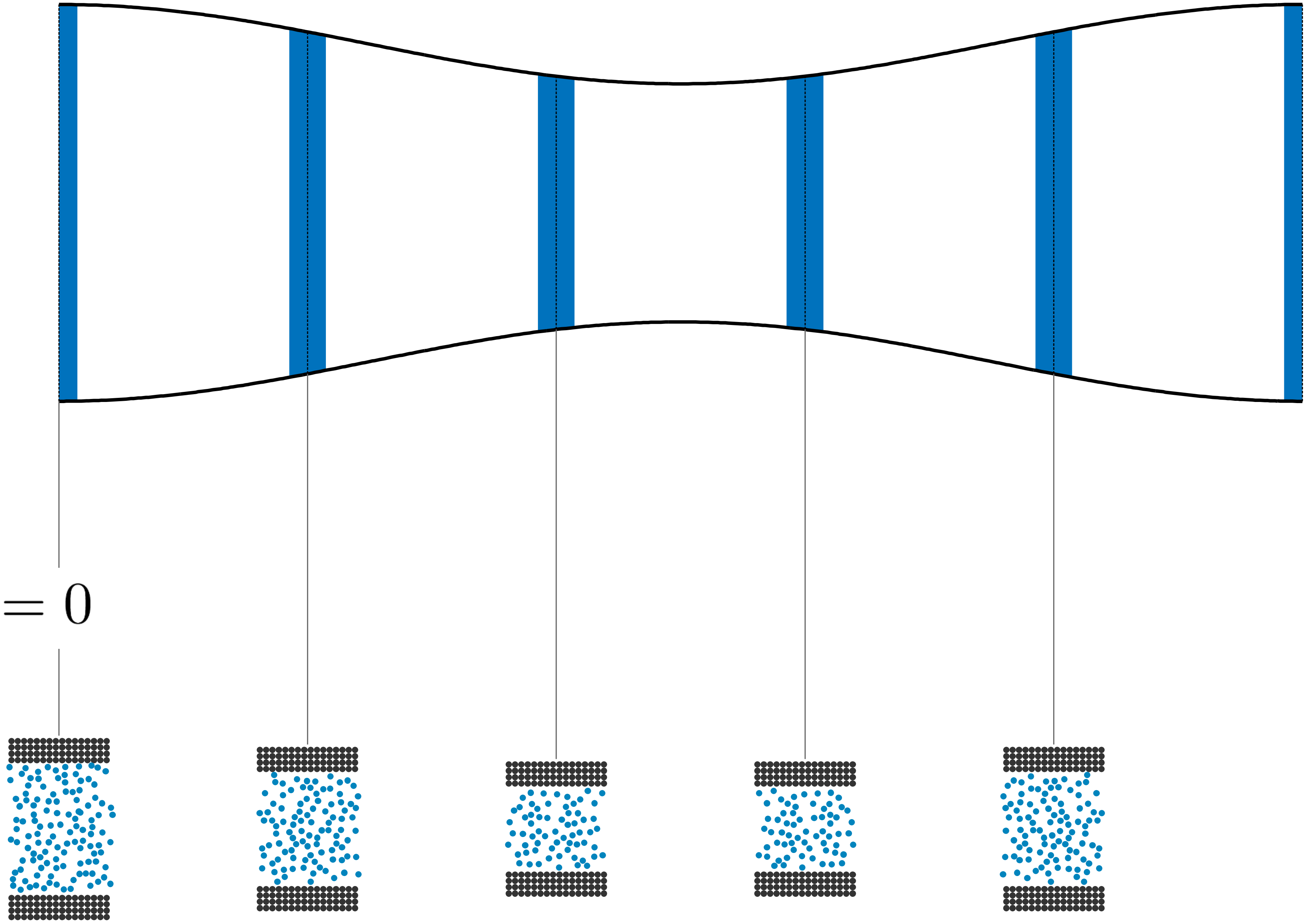
$$-\frac{\partial p}{\partial s} + \frac{\partial \tau_{sy}}{\partial y} + \frac{\partial \tau_{sz}}{\partial z} + \frac{\rho}{m} F_{ext} = 0$$



Hybrid method – micro model

MD subdomain simulations conserve mass.

Instead, we apply a larger body force.



$$\frac{\partial \tau_{sy}}{\partial y} + \frac{\partial \tau_{sz}}{\partial z} + \frac{\rho}{m} F_\mu = 0$$

$$F_\mu = F_{ext} - \frac{m}{\rho}\frac{\partial p}{\partial s}$$

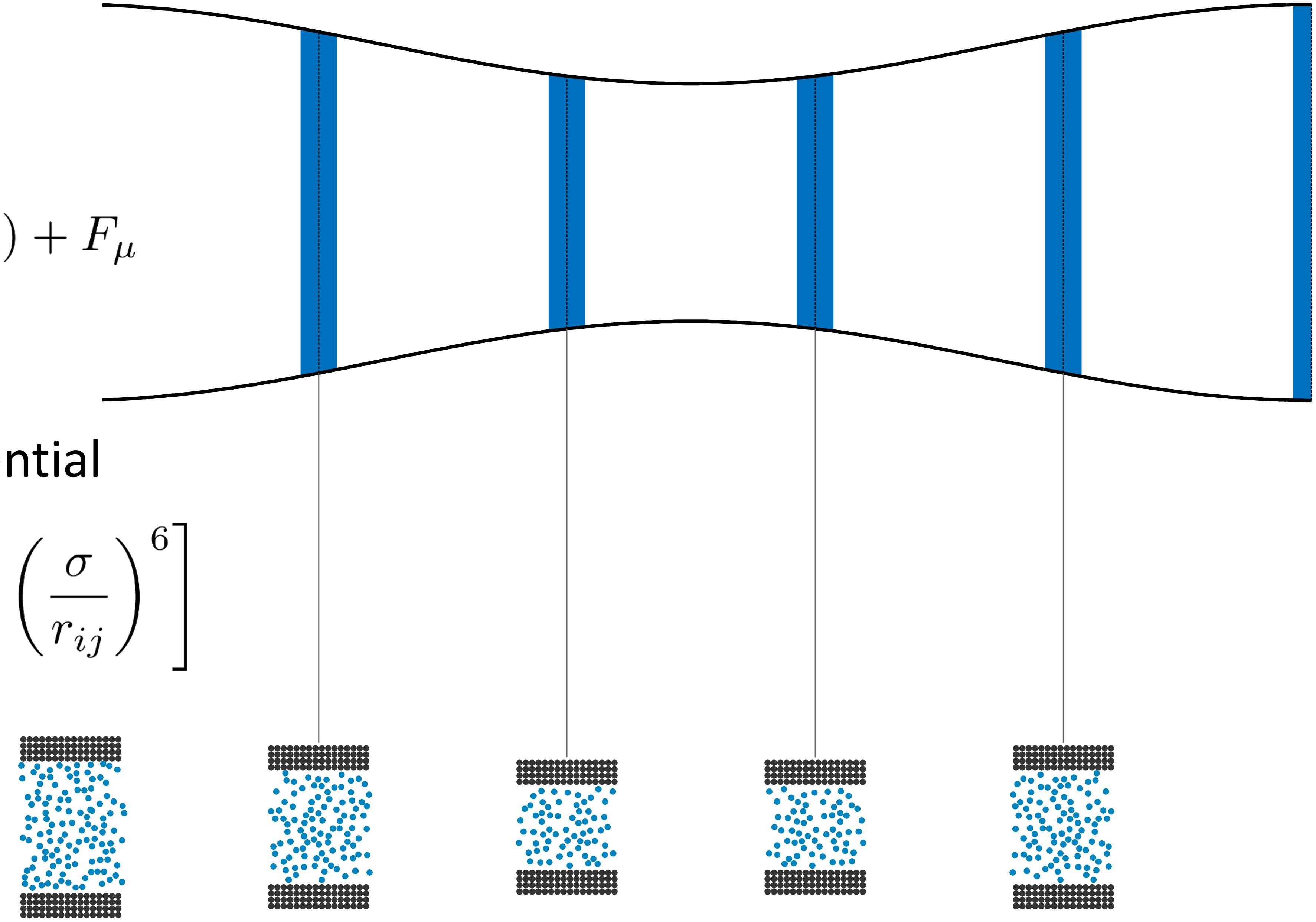
Hybrid method – micro model

Newton's 2nd law

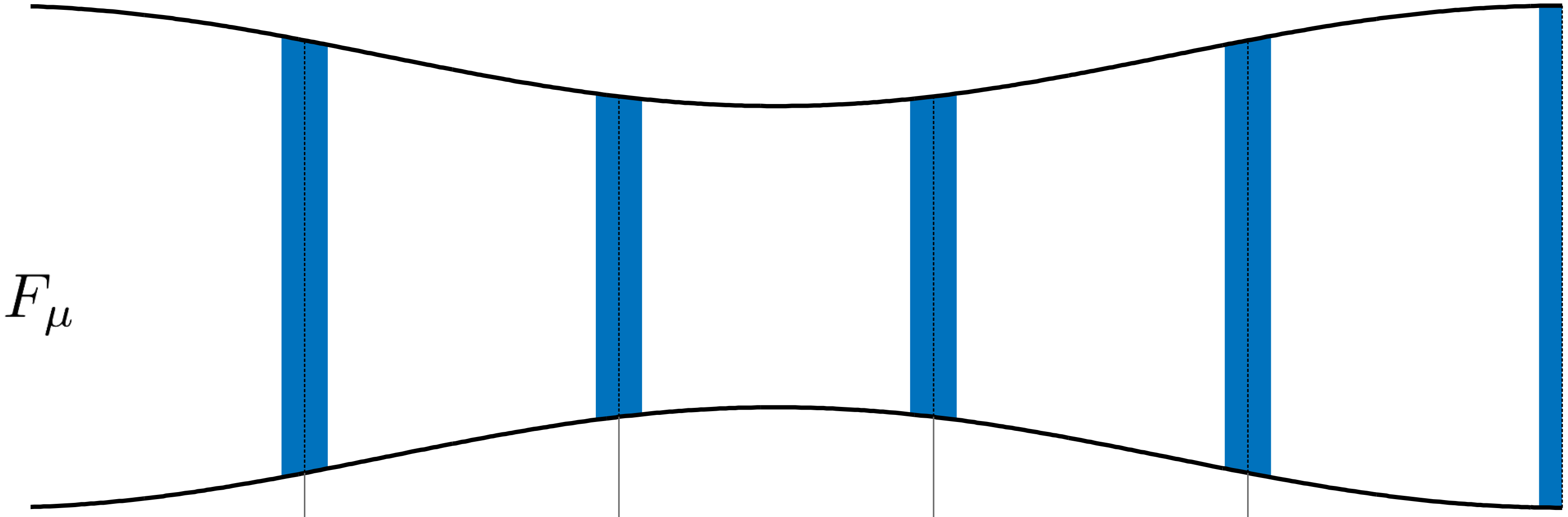
$$m_i \ddot{\mathbf{r}}_i = \sum_{j=1(\neq i)}^{N_m} -\nabla U(r_{ij}) + F_\mu$$

Lennard Jones potential

$$U(r_{ij}) = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]$$



Hybrid method – micro model



Newton's 2nd law

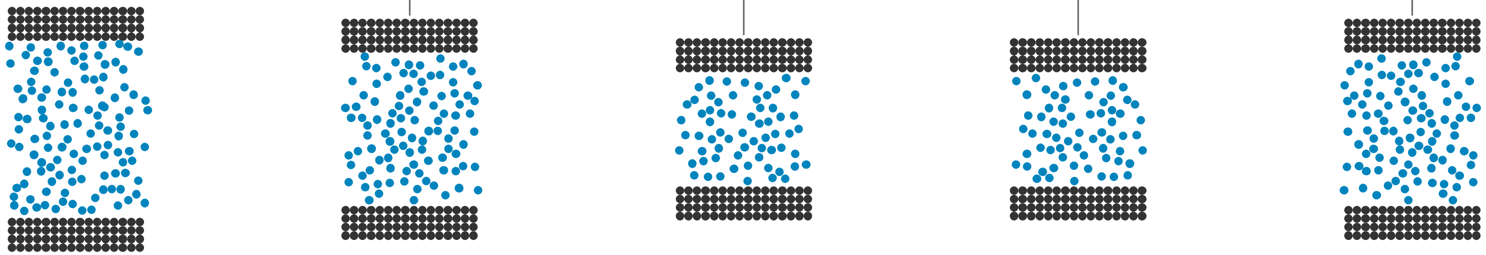
$$m_i \ddot{\mathbf{r}}_i = \sum_{j=1(\neq i)}^{N_m} -\nabla U(r_{ij}) + F_\mu$$

Lennard Jones potential

$$U(r_{ij}) = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]$$

Distance to zero potential

Potential well depth



Hybrid method – micro model

Newton's 2nd law

$$m_i \ddot{\mathbf{r}}_i = \sum_{j=1(\neq i)}^{N_m} -\nabla U(r_{ij}) + F_\mu$$

Lennard Jones potential

$$U(r_{ij}) = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]$$

Potential well depth

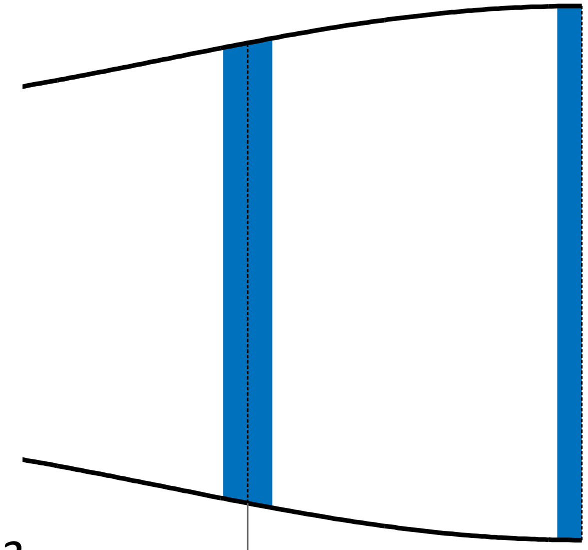
Distance to zero potential

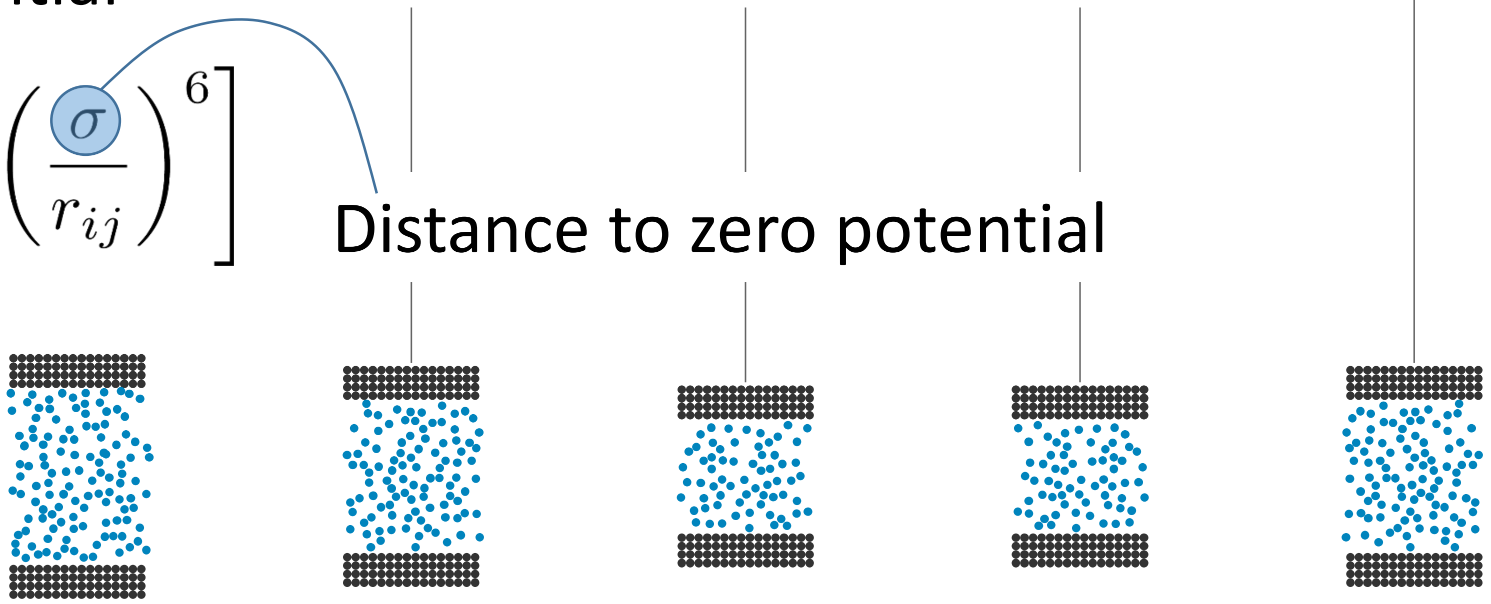
For argon:

$$\sigma_{w-f} = 2.55 \times 10^{-10}\ m$$
$$\epsilon_{w-f} = 0.33 \times 10^{-21}\ J$$

$$\sigma_{f-f} = 3.4 \times 10^{-10}\ m$$
$$\epsilon_{f-f} = 1.65 \times 10^{-21}\ J$$

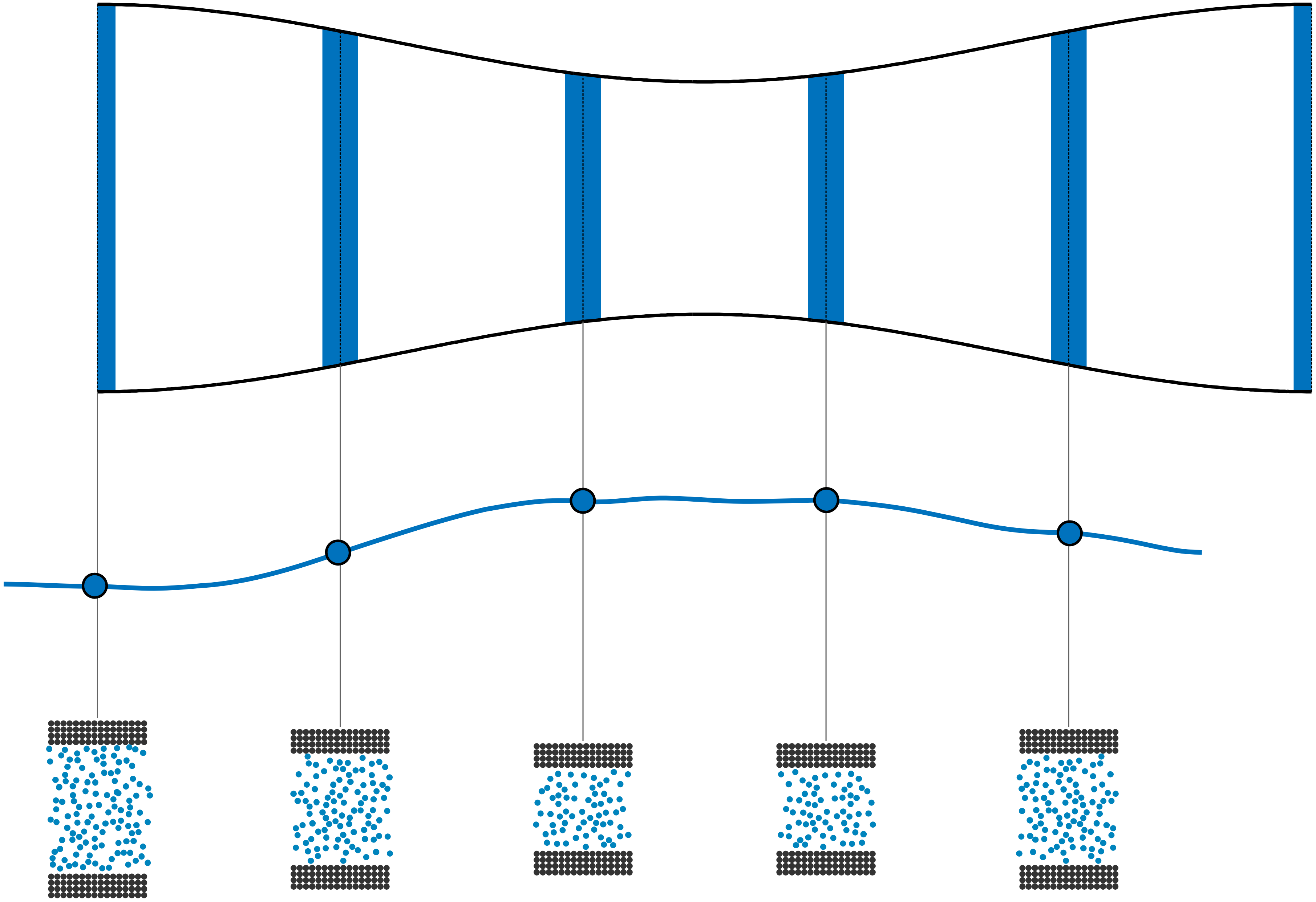
Thompson and Troian (1997), Nature, 389:360:362





Hybrid methods – machine learning

We'd like to use the existing information to make predictions.

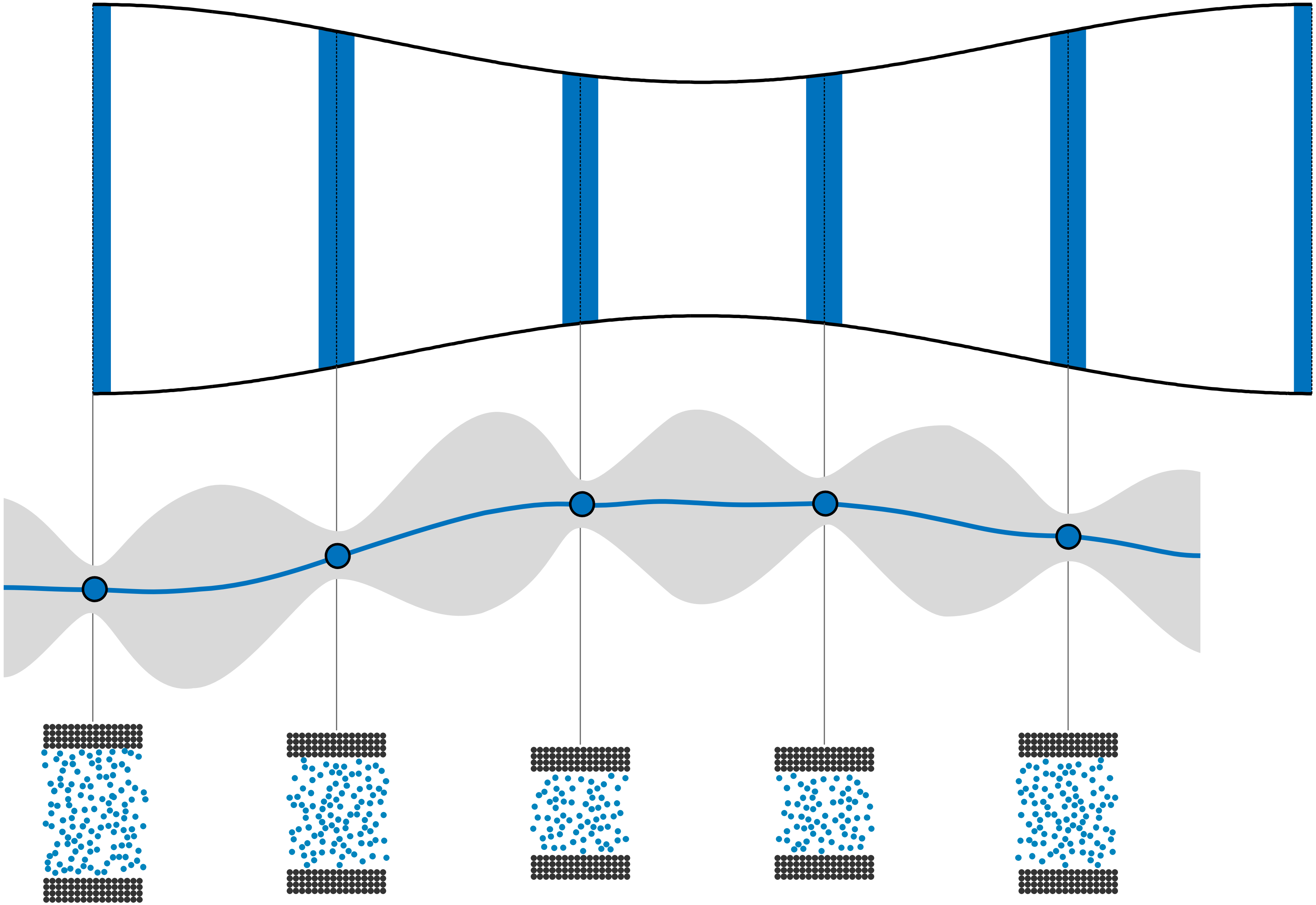


Hybrid methods – machine learning

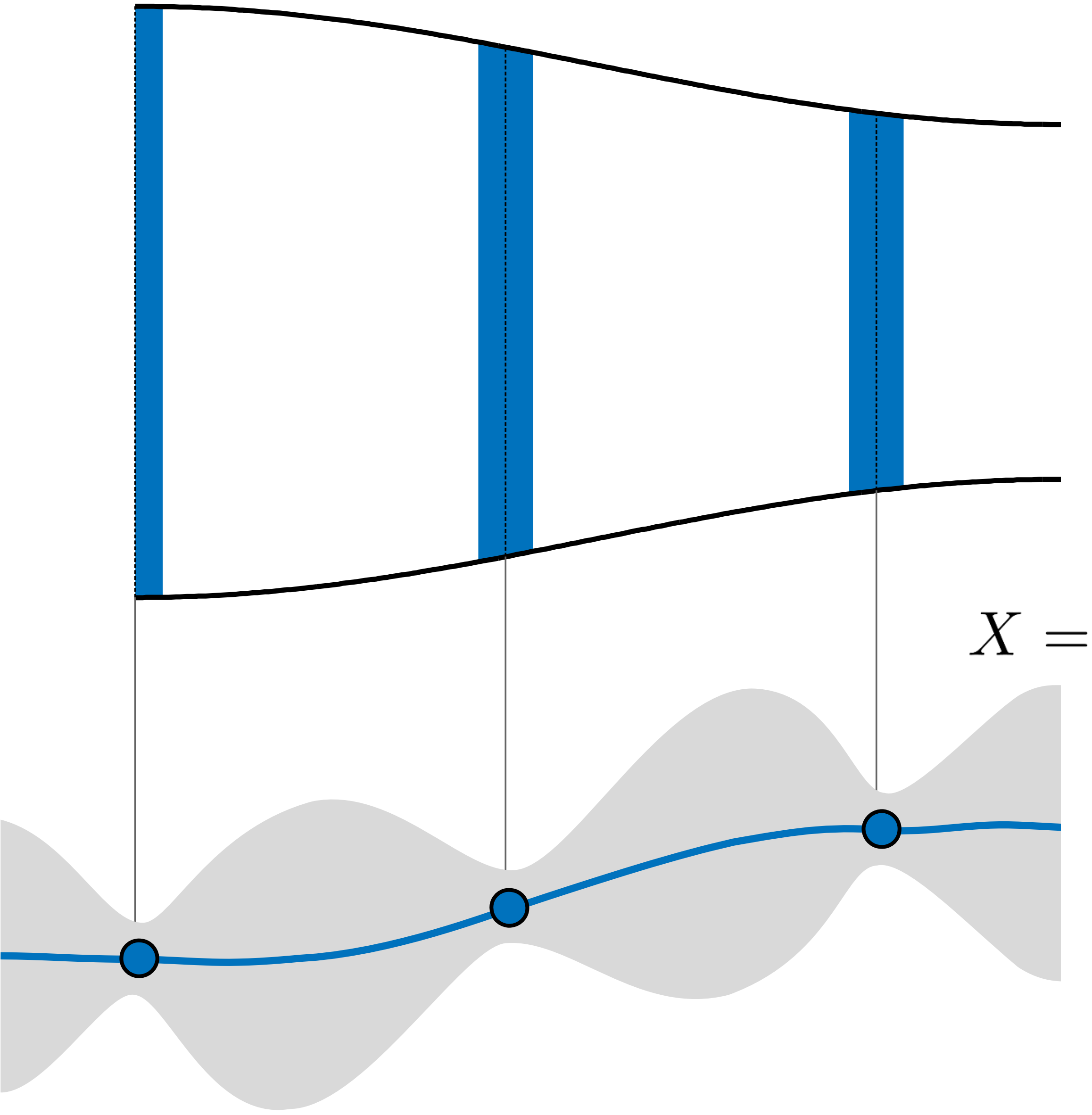
We'd like to use the existing information to make predictions.

We want to judge on-the-fly if a new simulation is required.

For this, we use a Gaussian process (GP).



GP regression for our multiscale system



input

$$\boldsymbol{x}_i = \{h_i,\ \rho_i,\ F_i\}$$

output

$$y_i = q_i \text{ or } p_i$$

Gaussian process

$$f(\mathbf{x}_i) \sim \mathcal{GP}\big(\mu(\mathbf{x}_i), K(\mathbf{x}_i, \mathbf{x}_j)\big)$$

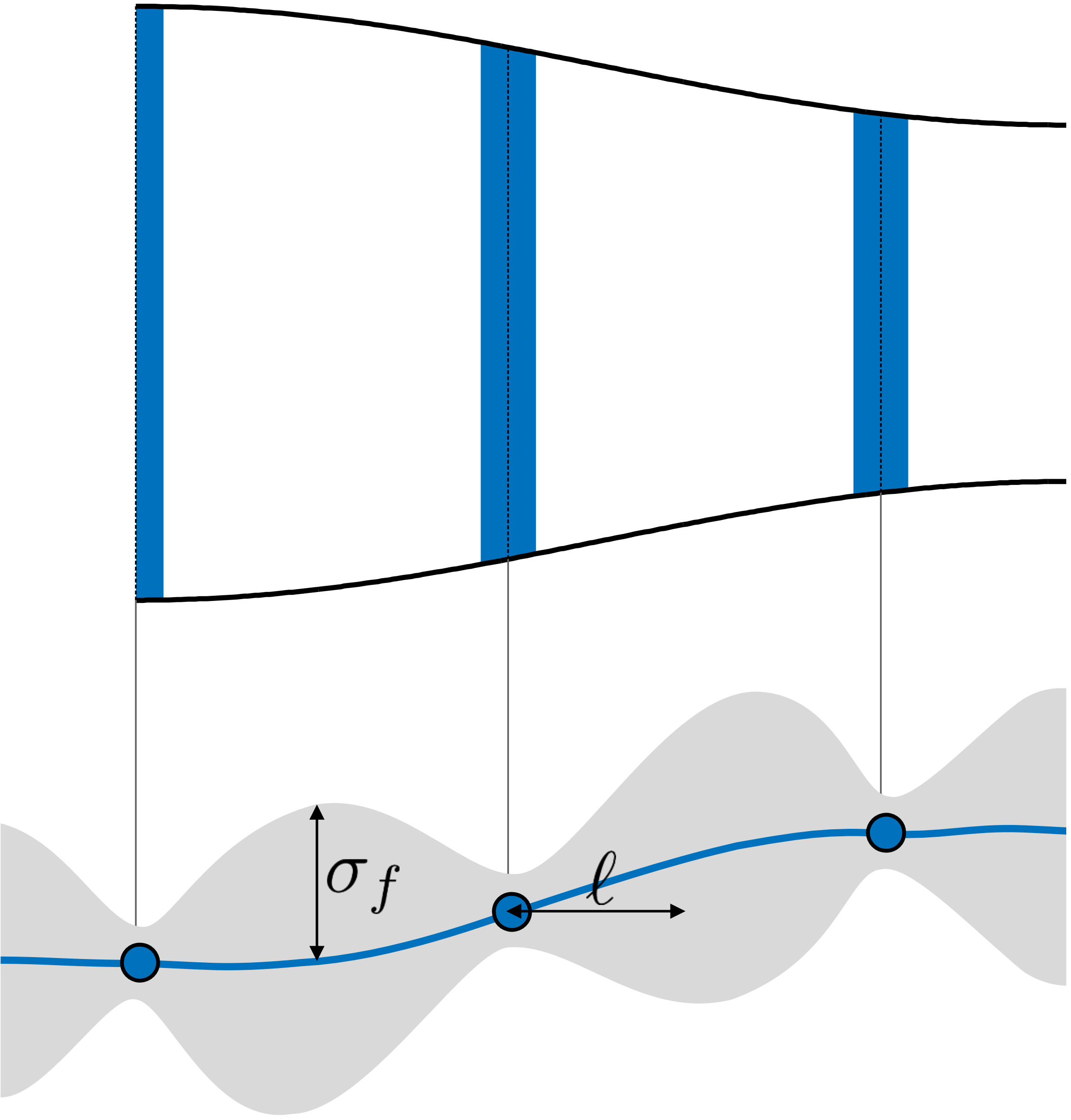
training points

$$X = \begin{bmatrix} h_1 & h_2 & \dots & h_M \\ \rho_1 & \rho_2 & \dots & \rho_M \\ F_1 & F_2 & \dots & F_M \end{bmatrix} \quad \langle \boldsymbol{y} \rangle = [\langle y_1 \rangle\ \langle y_2 \rangle\ \dots\ \langle y_M \rangle]$$

posterior distribution – Bayes' theorem

$$P(f|\langle \mathbf{y} \rangle, X) = \frac{P(\langle \mathbf{y} \rangle | X, f) P(f)}{P(\langle \mathbf{y} \rangle | X)}$$

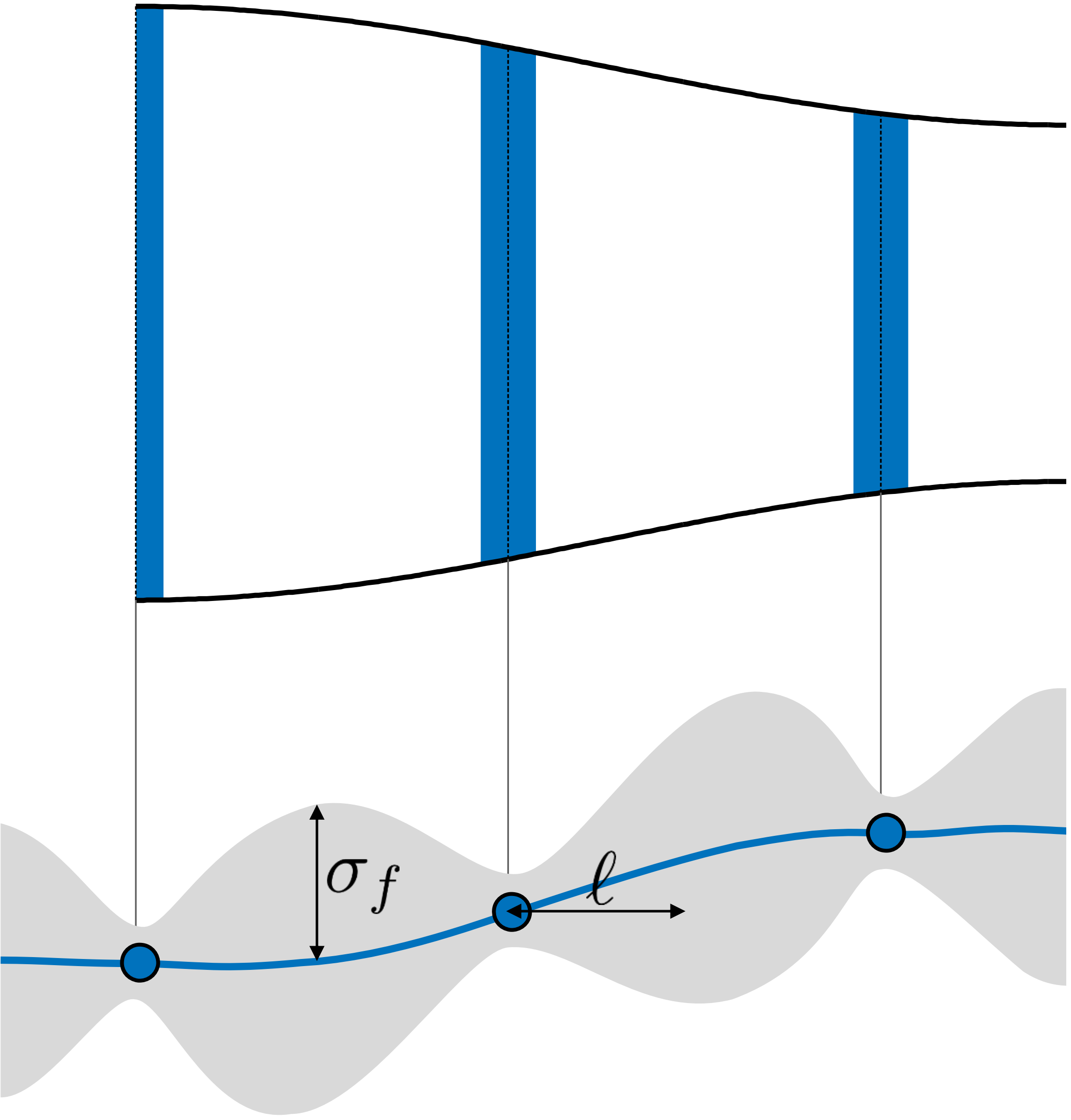
GP regression for our multiscale system



function covariance –
squared exponential kernel

$$K(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left(-\frac{d_{ij}^2}{2\ell^2}\right)$$

GP regression for our multiscale system



function covariance –
squared exponential kernel

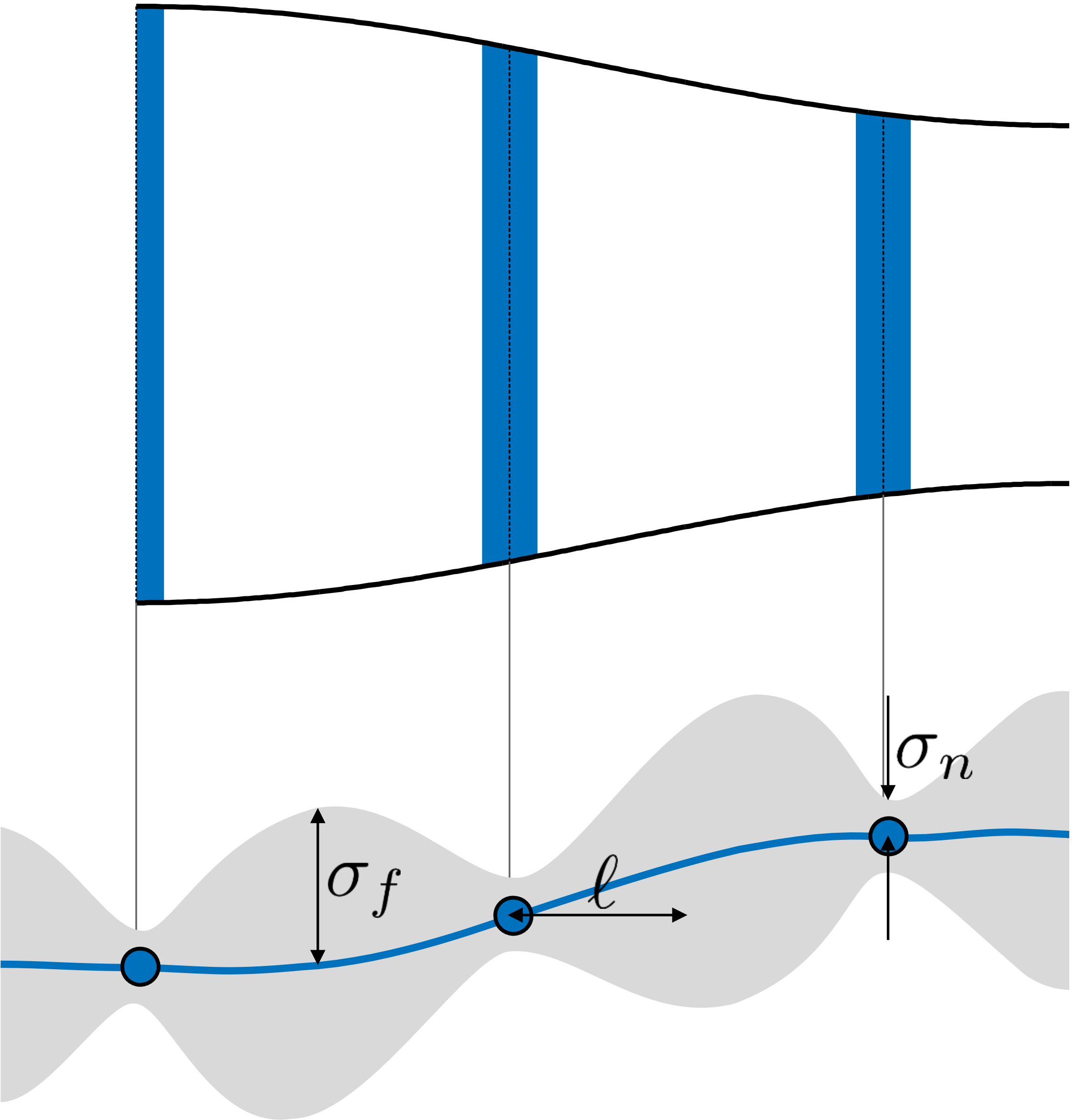
$$K(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left(-\frac{d_{ij}^2}{2\ell^2}\right)$$

input difference

$$d_{ij}^2 = \left(\frac{h_i - h_j}{\overline{\Delta h}}\right)^2 + \left(\frac{\rho_i - \rho_j}{\overline{\Delta \rho}}\right)^2 + \left(\frac{F_i - F_j}{\overline{\Delta F}}\right)^2$$

$$\overline{\Delta h} = \mathbb{E}\begin{bmatrix} \Delta h_{11} & \Delta h_{12} & \dots & \Delta h_{1M} \\ \Delta h_{21} & \Delta h_{22} & \dots & \Delta h_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta h_{M1} & \Delta h_{M2} & \dots & \Delta h_{MM} \end{bmatrix}$$

GP regression for our multiscale system



function covariance

$$K(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left(-\frac{d_{ij}^2}{2\ell^2}\right)$$

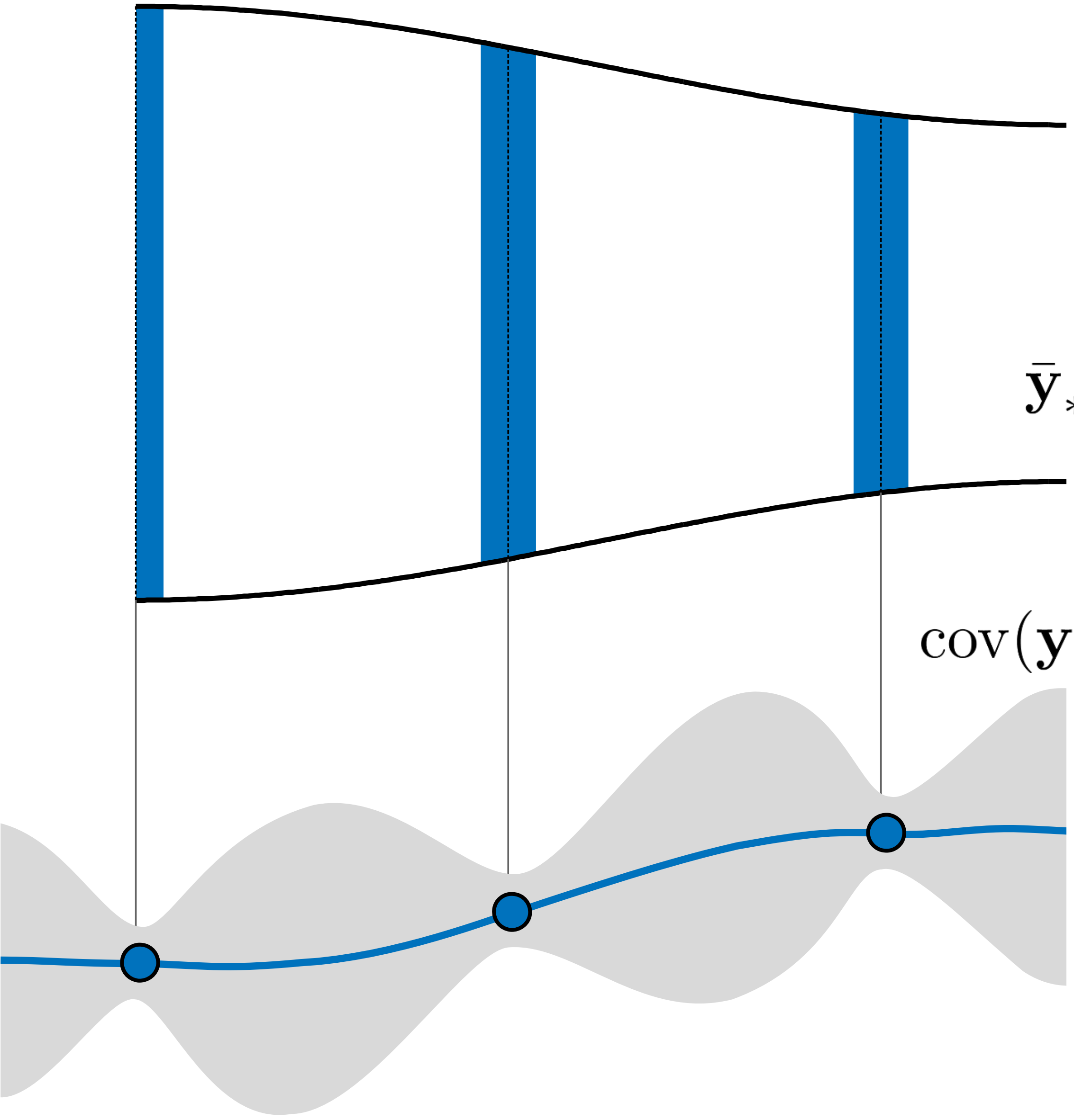
output covariance

$$C(X, X) = K(X, X) + \sigma_n^2 I$$

hyperparameters

	q (ng/s)	p (MPa)
σ_n	0.05	0.003
σ_f	1	1
ℓ	1	1

GP regression for our multiscale system



posterior distribution at test points

$$\mathbf{y}_*|X, \mathbf{y}, X_* \sim \mathcal{N}\left(\bar{\mathbf{y}}_*, \mathrm{cov}(\mathbf{y}_*)\right)$$

posterior mean

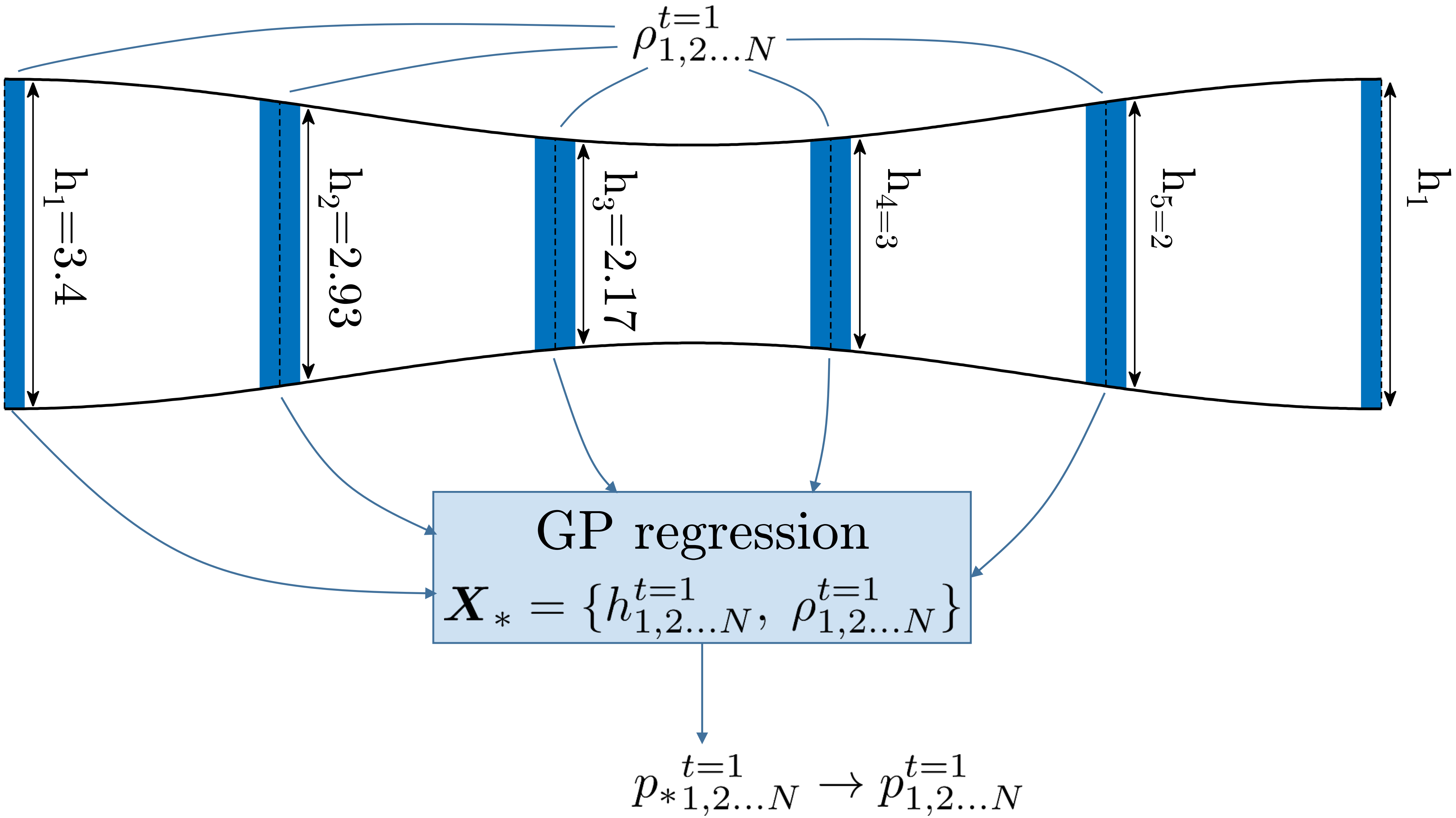
$$\bar{\mathbf{y}}_* = \mu(X_*) + K(X_*, X)C(X, X)^{-1}\left(\langle\mathbf{y}\rangle - \mu(X)\right)$$

posterior covariance

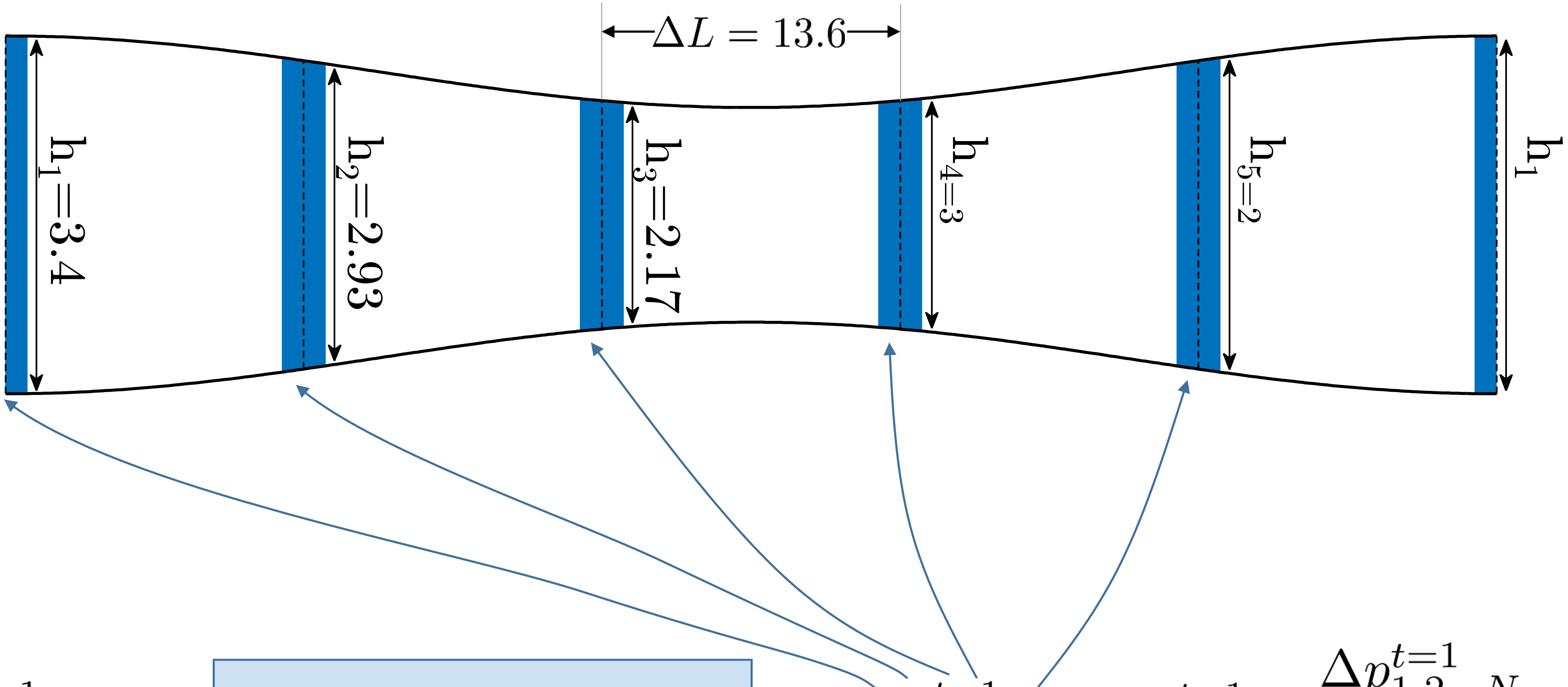
$$\mathrm{cov}(\mathbf{y}_*) = K(X_*, X_*) - K(X_*, X)C(X, X)^{-1}K(X, X_*)$$

$$\text{if } \mathrm{cov}(\mathbf{y}_*) < \sigma_t \text{ then } = y = y_*$$

Implementation procedure

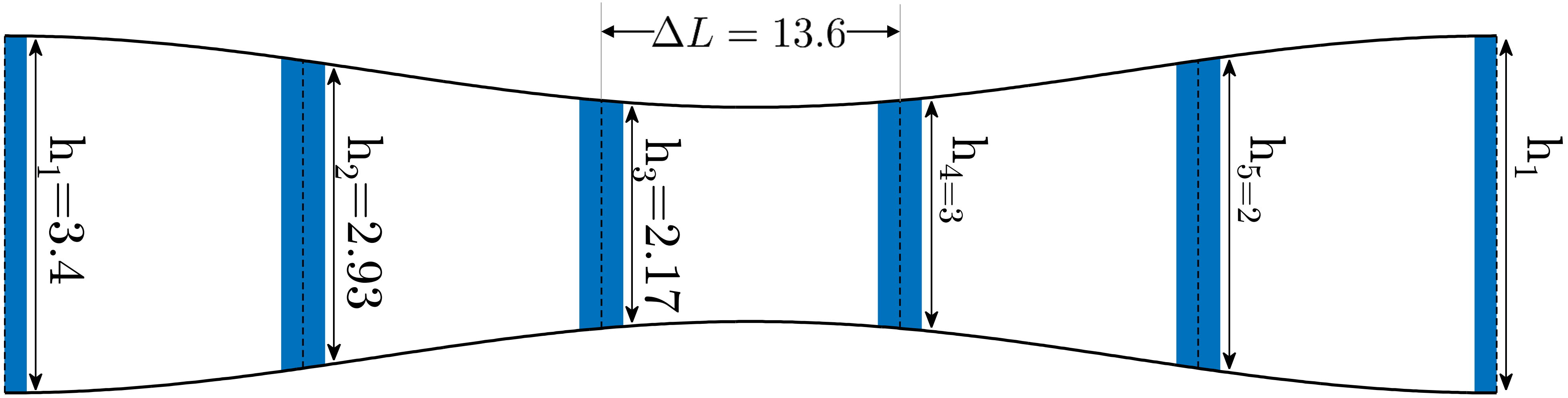


Implementation procedure



$p^{t=1}_{1,2\ldots N}$ → momentum balance → $F_{\mu}{}^{t=1}_{1,2\ldots N} = F^{t=1}_{ext} - \frac{\Delta p^{t=1}_{1,2\ldots N}}{2\Delta L}$

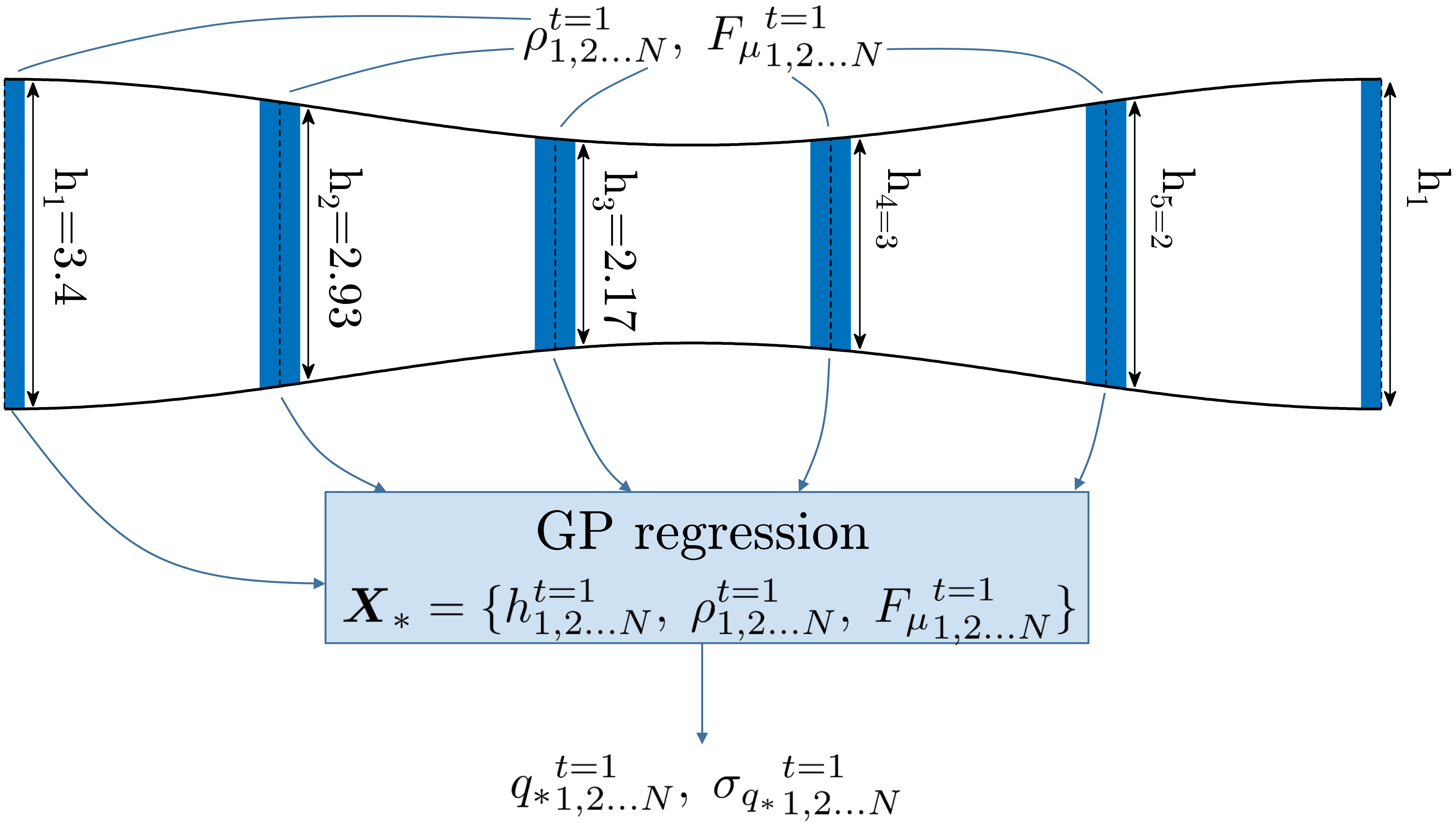
Implementation procedure



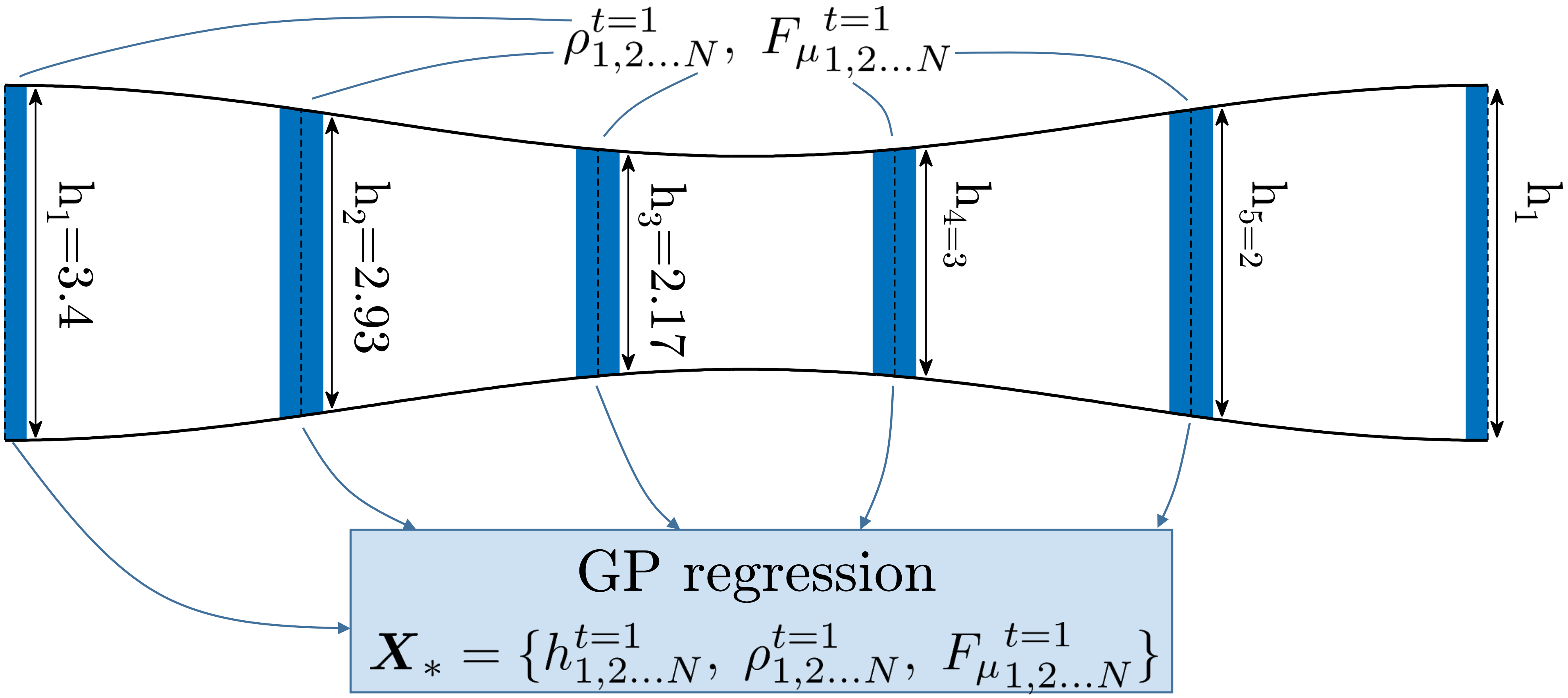
$$F_{ext} = 0.487 \sin\left(\frac{2\pi t}{10.8}\right) \ \text{pN}$$

$p^{t=1}_{1,2\ldots N}$ → **momentum balance** → $F_{\mu}{}^{t=1}_{1,2\ldots N} = F^{t=1}_{ext} - \dfrac{\Delta p^{t=1}_{1,2\ldots N}}{2\Delta L}$

Implementation procedure

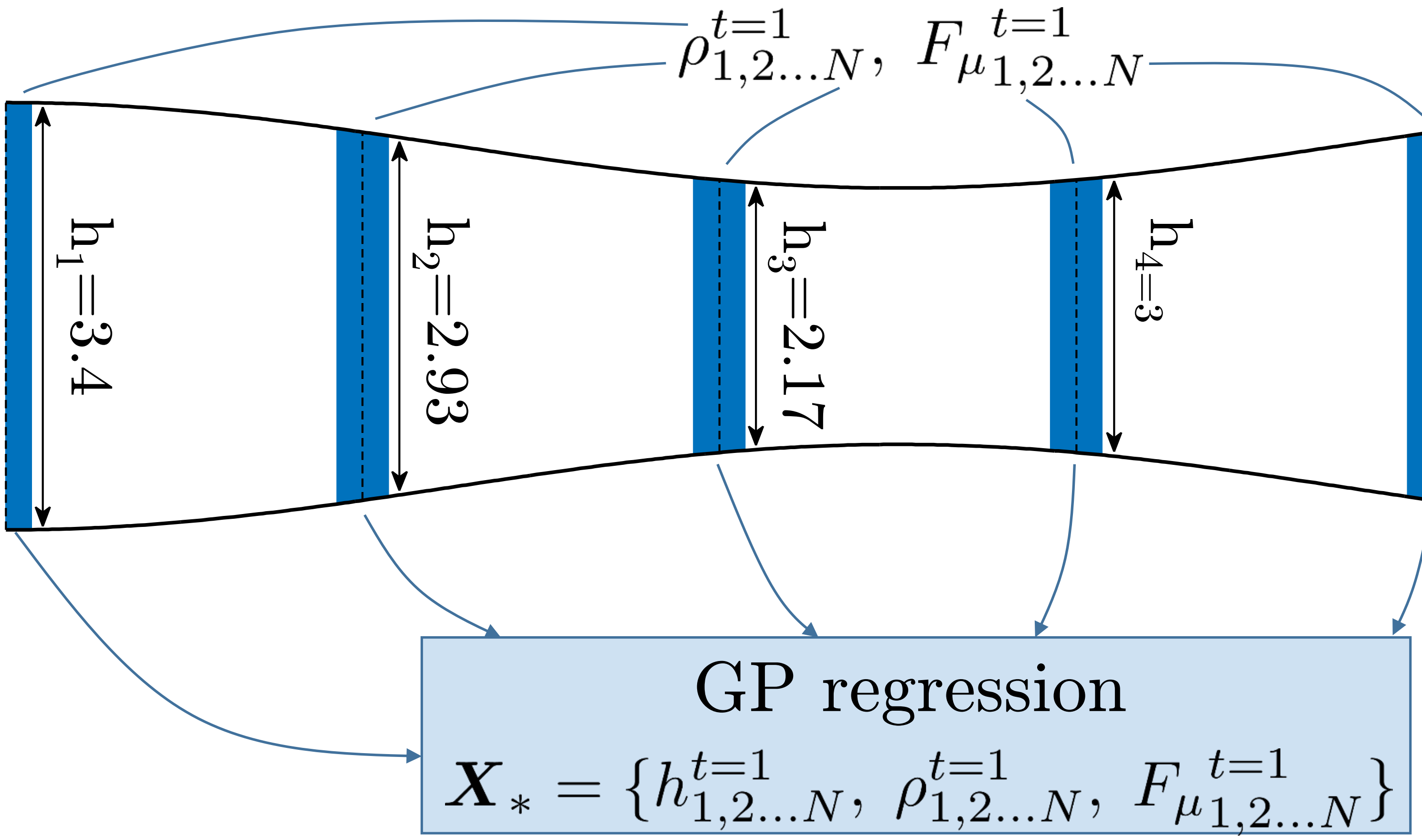


Implementation procedure



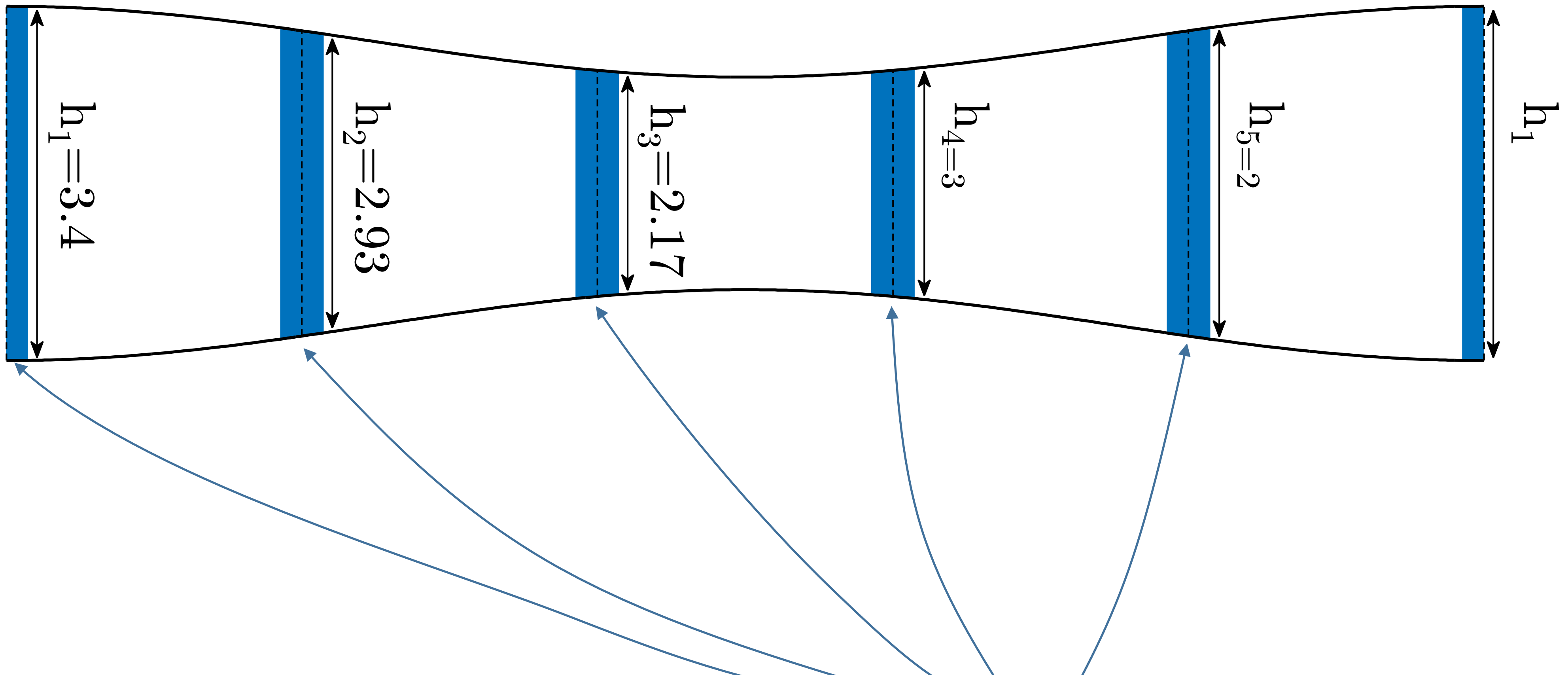
$$\text{if } \sigma_{q_*}{}^{t=1}_{i} < \sigma_t; \quad \text{then } q_*{}^{t=1}_{i} \rightarrow q^{t=1}_{i}$$

Implementation procedure



if $\sigma_{q_*\,i}^{\ t=1} > \sigma_t$ → MD simulation → $q_i^{t=1} = \langle q_i^{t=1} \rangle$ → update database

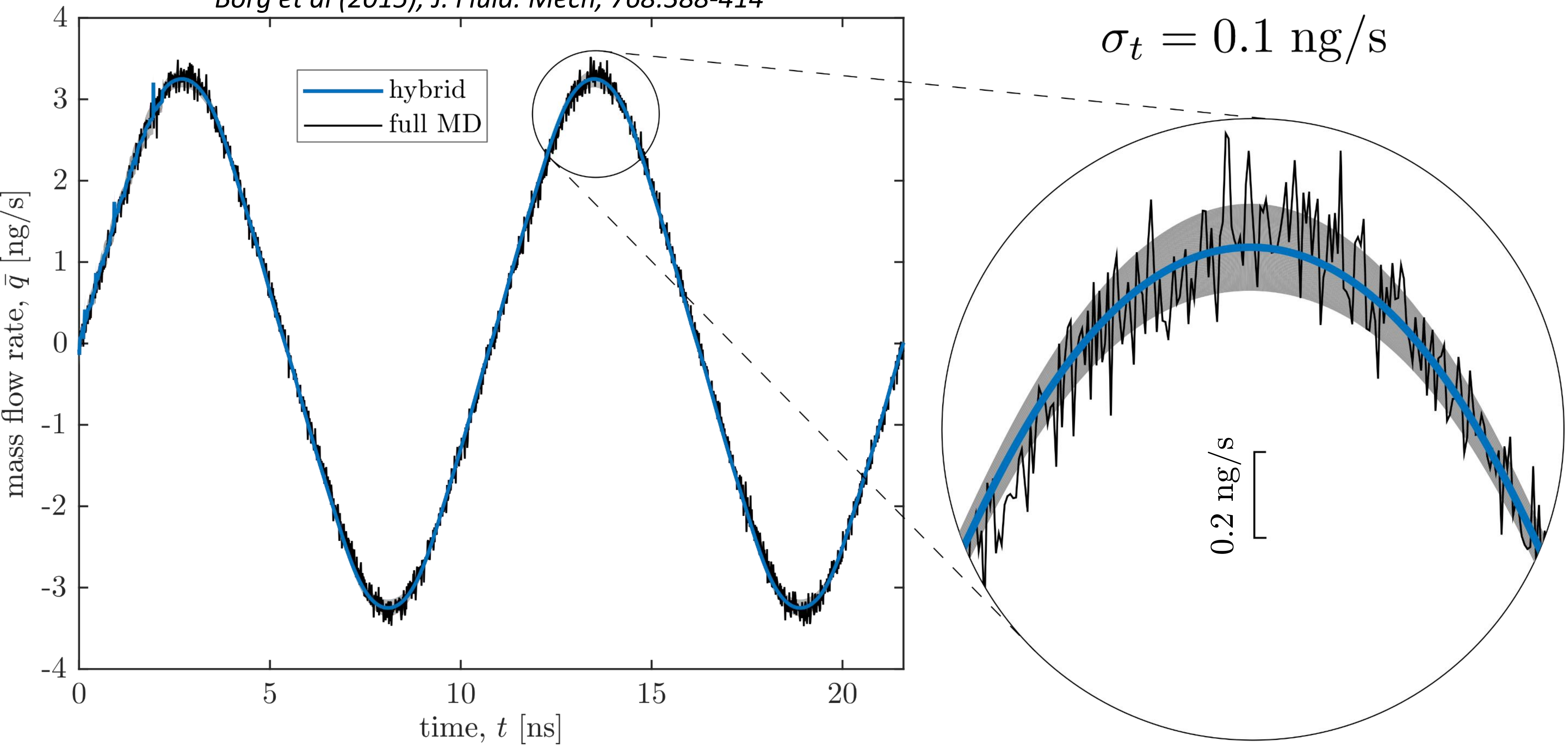
Implementation procedure



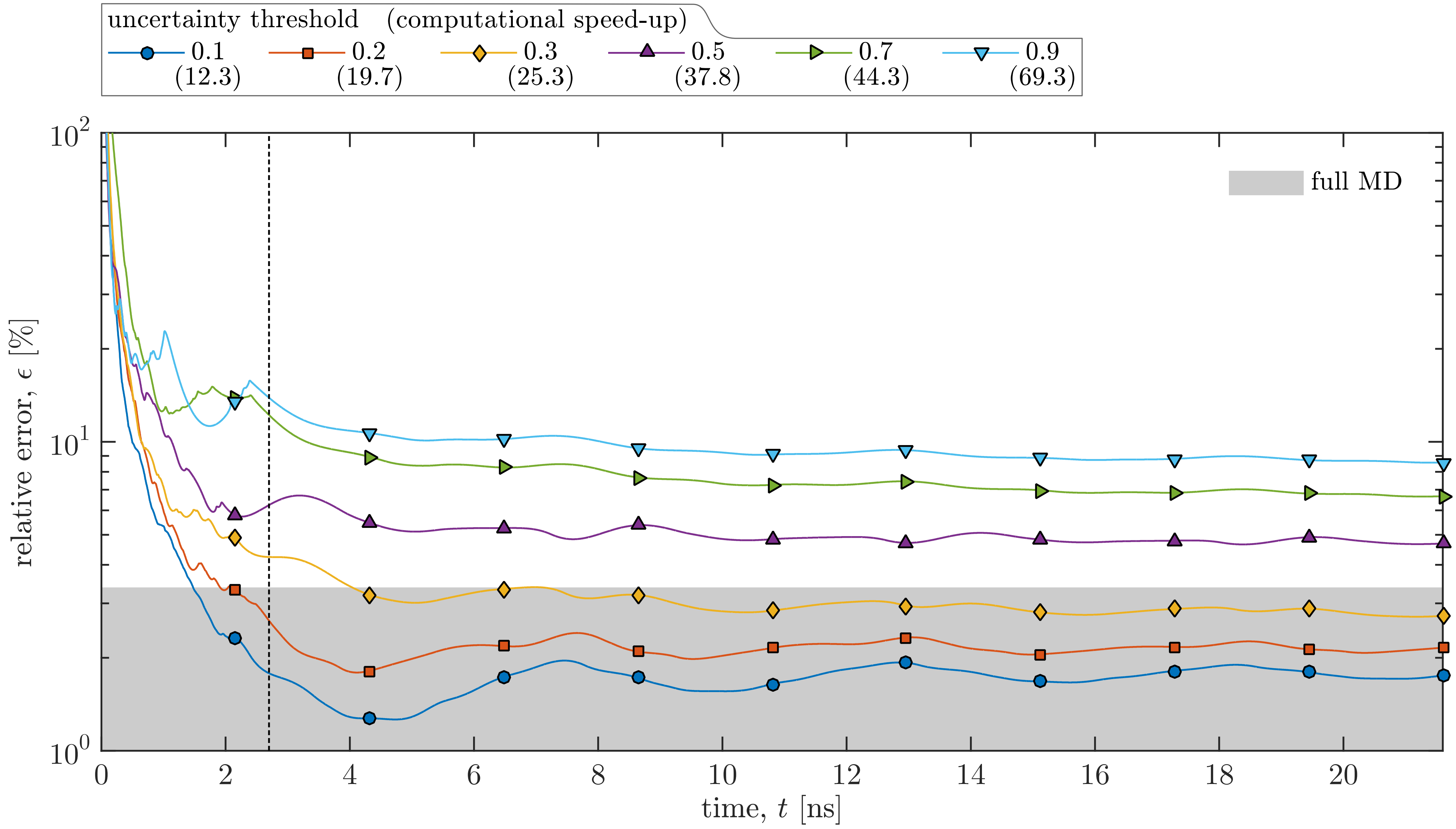
$q^{t=1}_{1,2\ldots N}$ → mass conservation → $\rho^{t=2}_{1,2\ldots N} = -\frac{\Delta t}{A}\frac{\Delta q^{t=1}_{1,2\ldots N}}{2\Delta L}$

Results – empty database, low uncertainty threshold

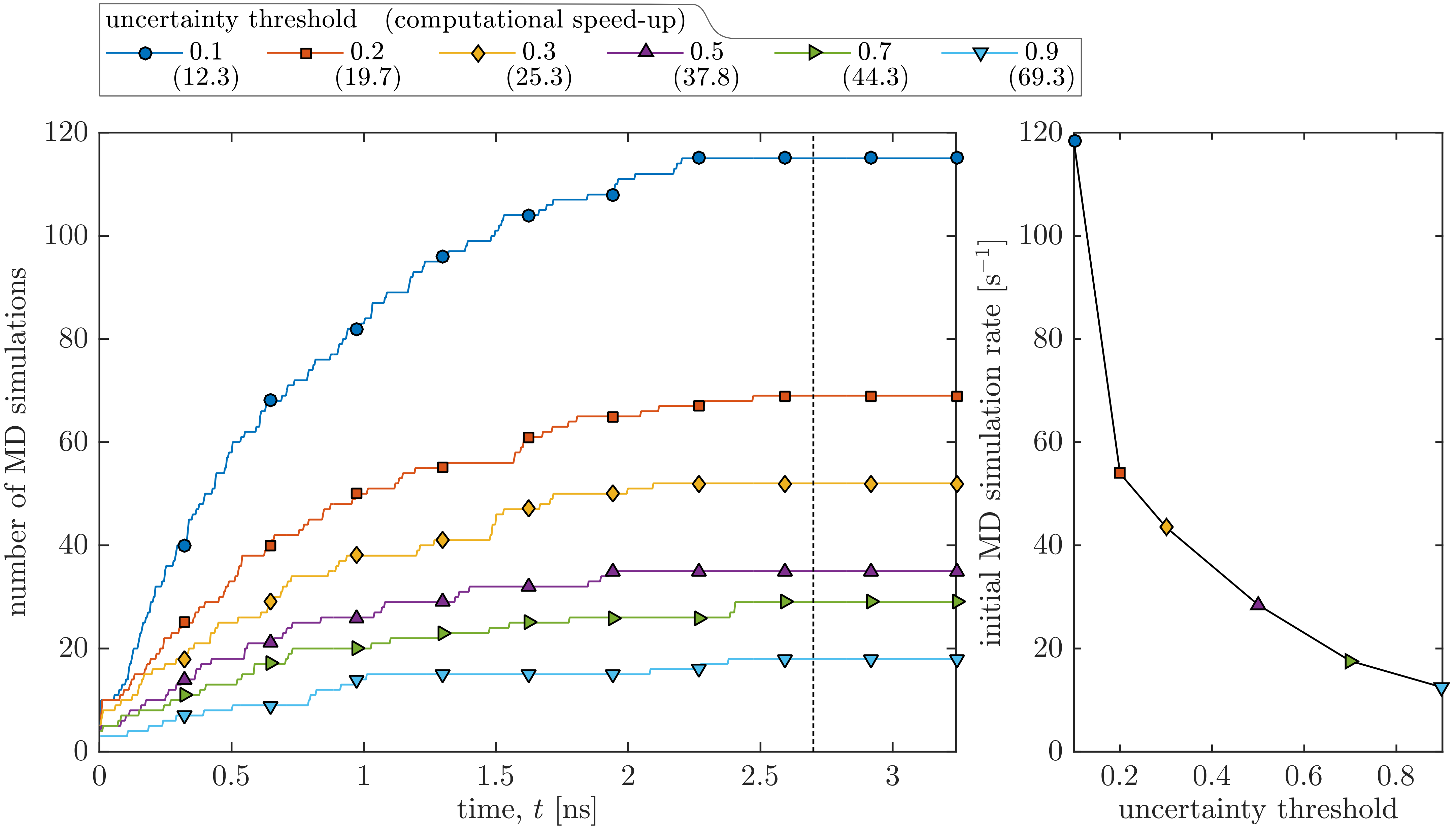
Borg et al (2015), J. Fluid. Mech, 768:388-414



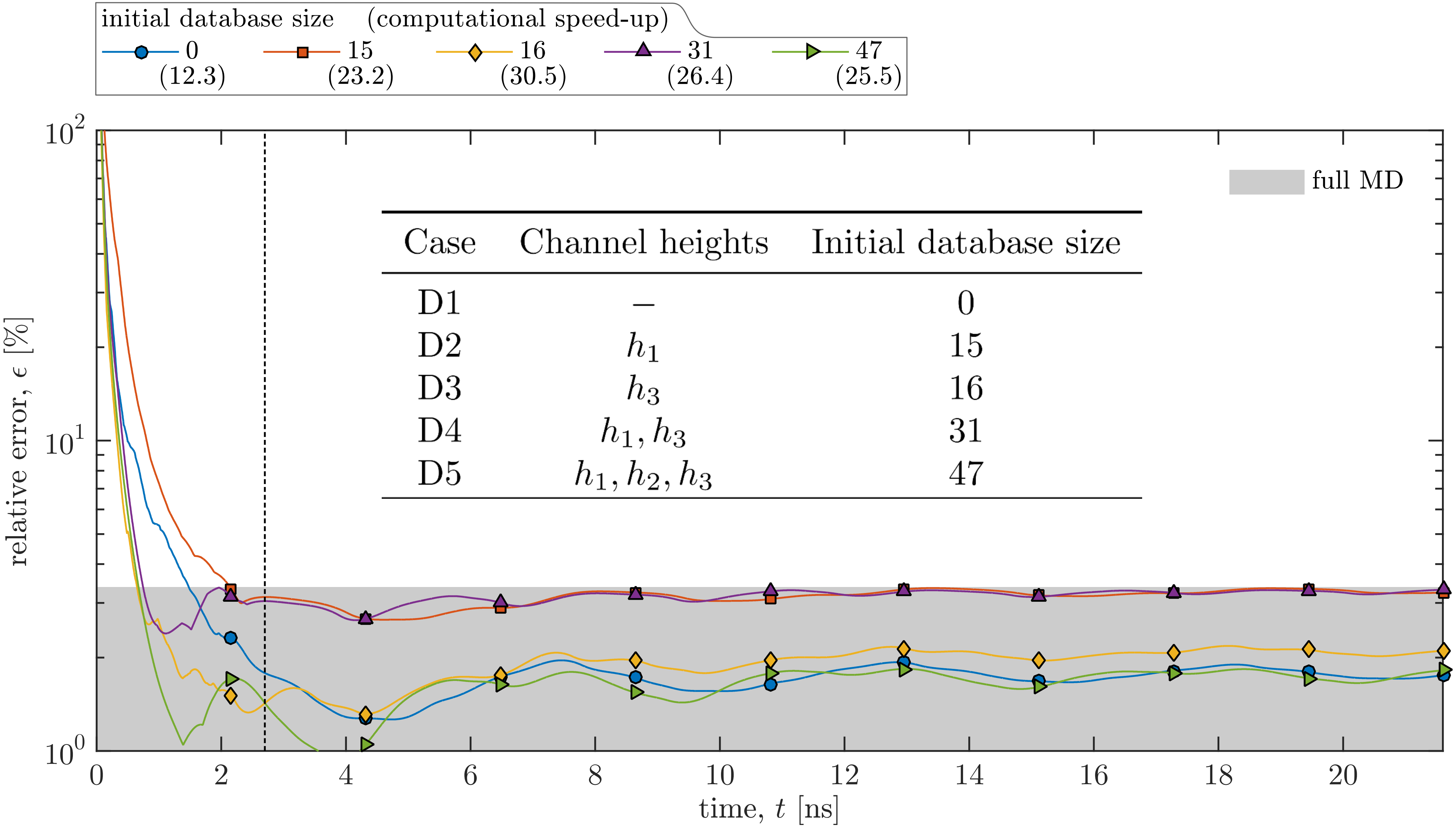
Results – empty database, varying uncertainty threshold



Results – empty database, varying uncertainty threshold

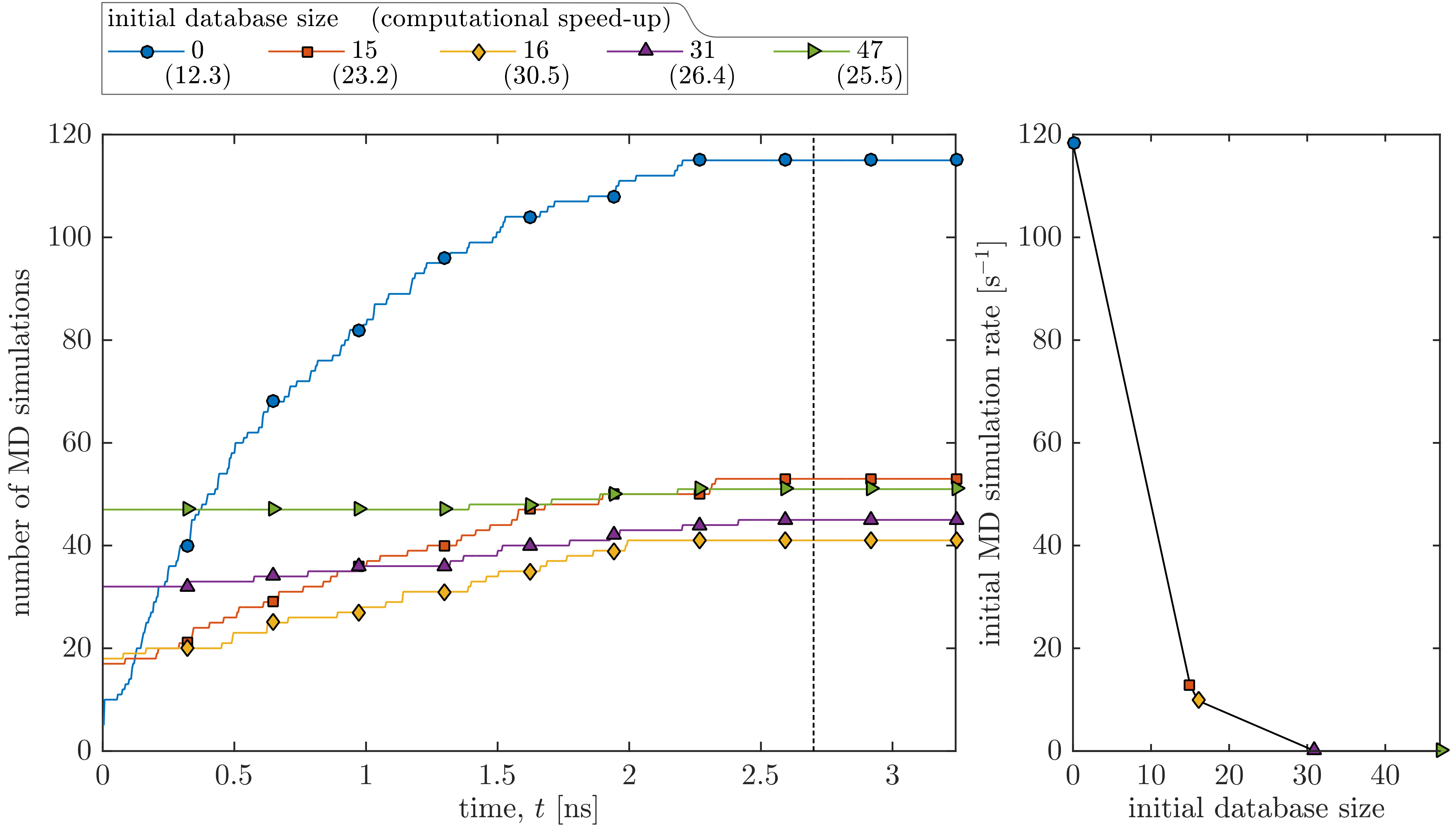


Results – varying database, low uncertainty threshold

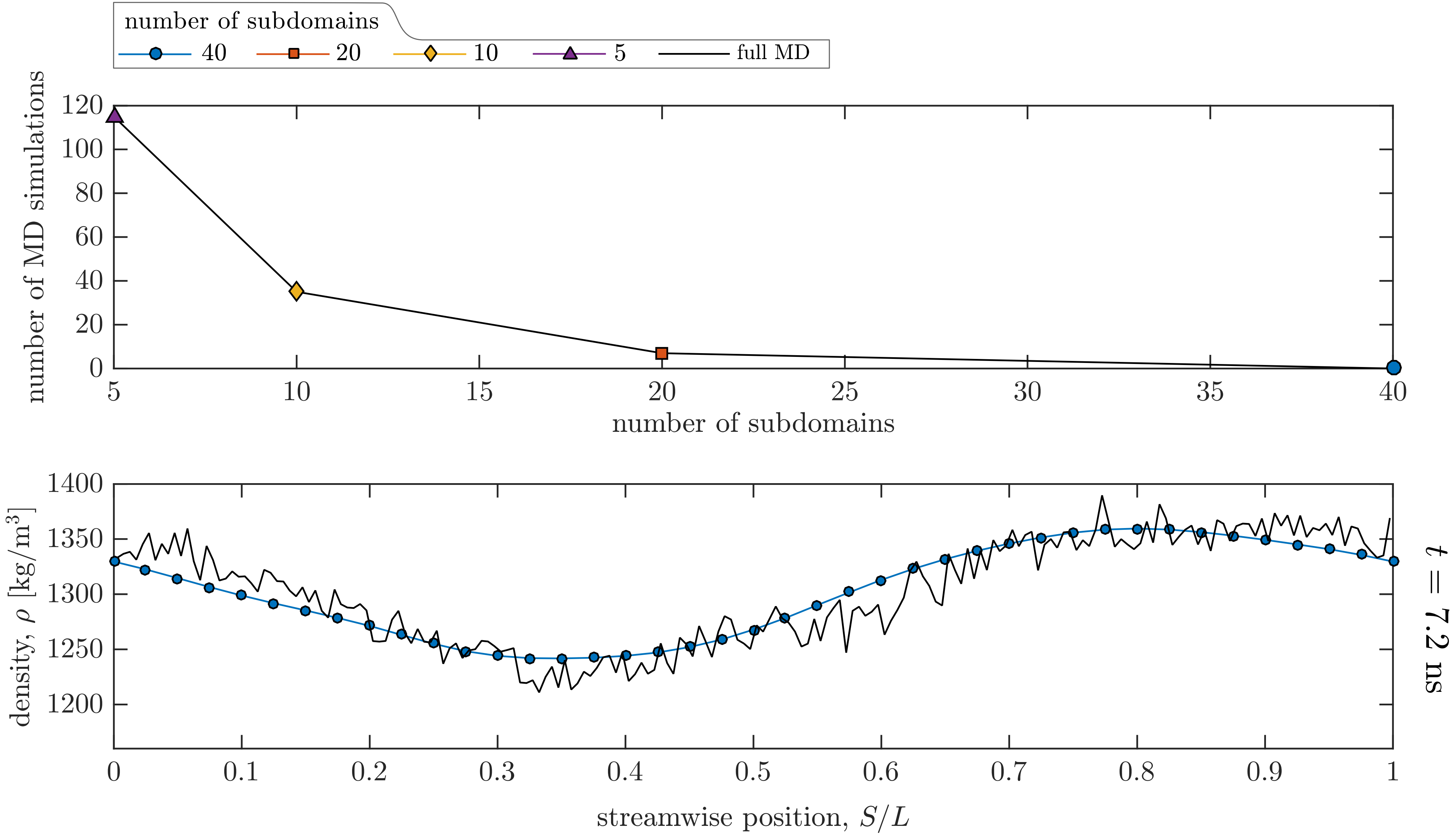


Case	Channel heights	Initial database size
D1	–	0
D2	h_1	15
D3	h_3	16
D4	h_1, h_3	31
D5	h_1, h_2, h_3	47

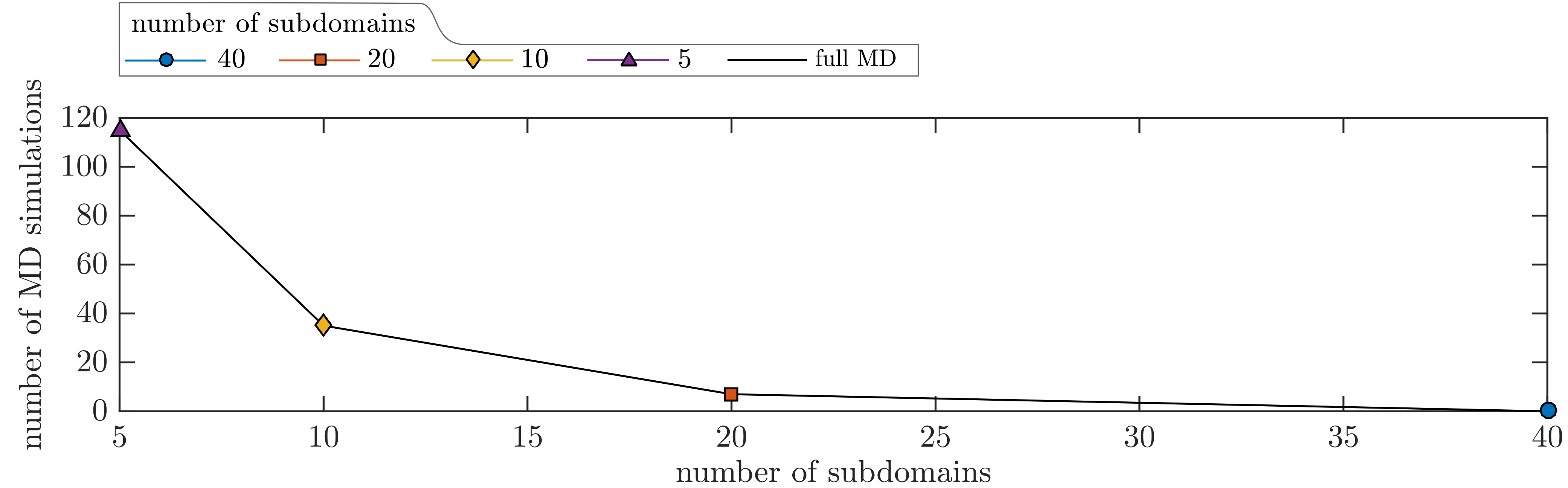
Results – varying database, low uncertainty threshold

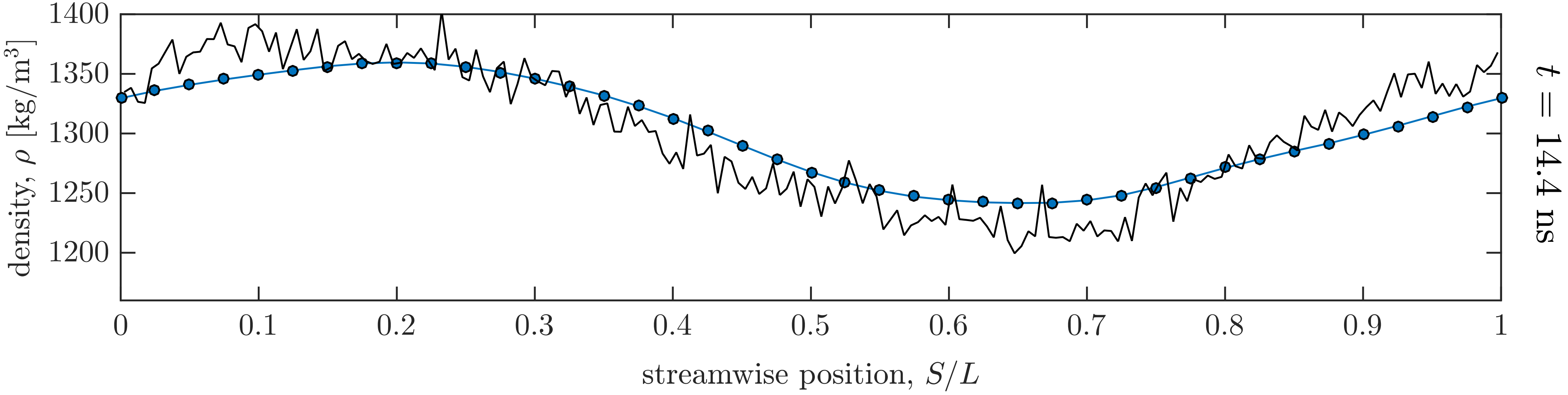


Results – expanding a database: subdomains

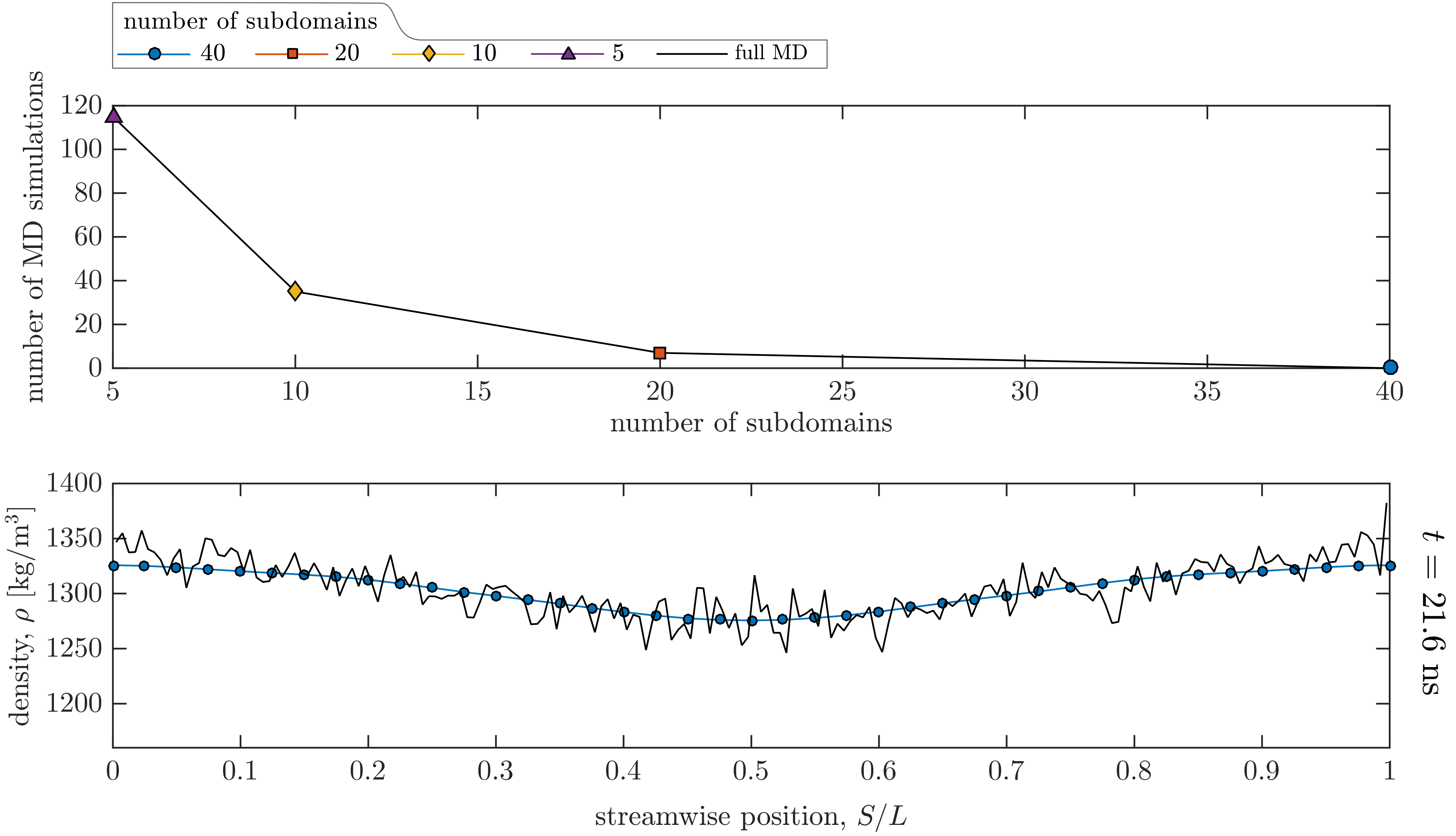


Results – expanding a database: subdomains

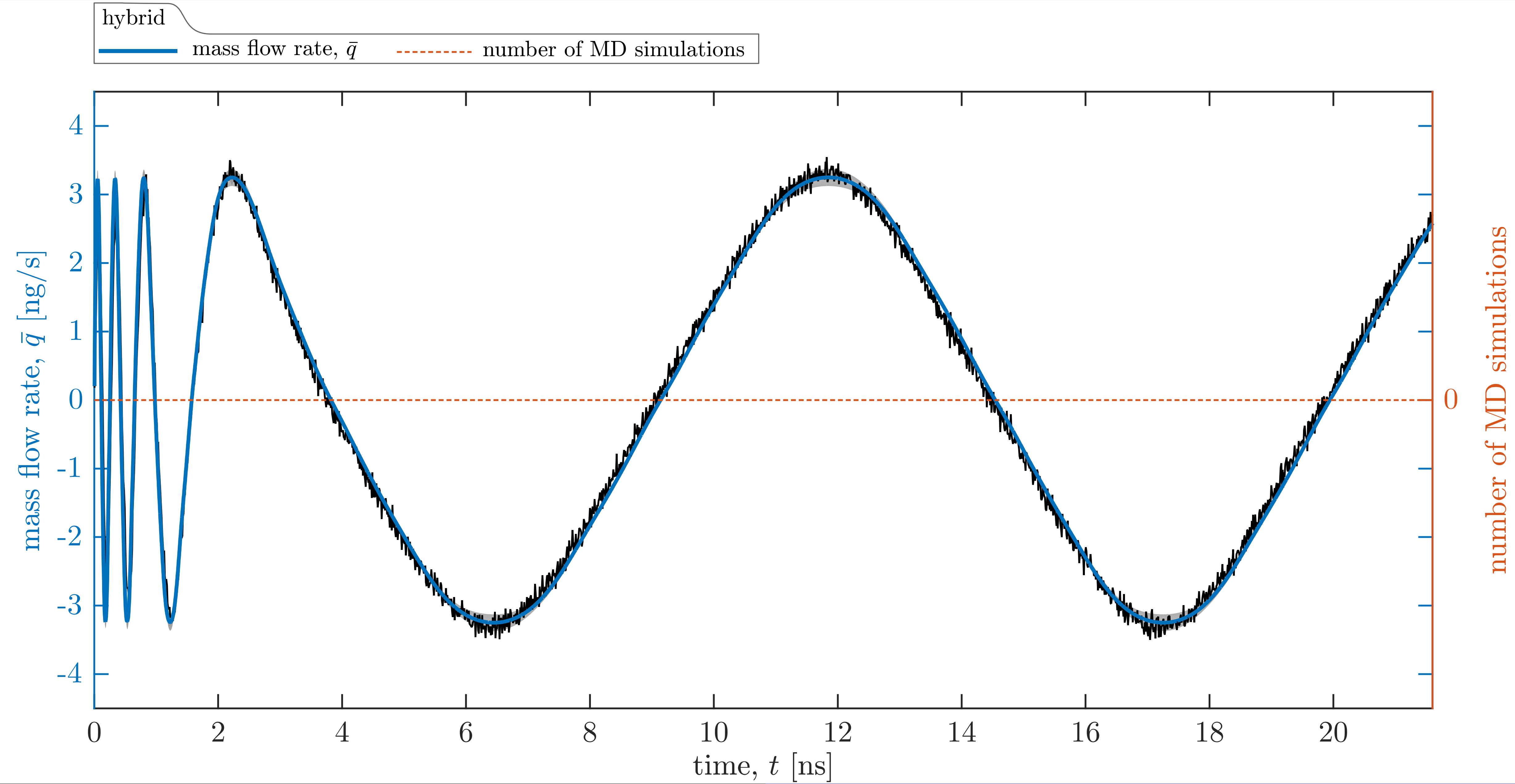




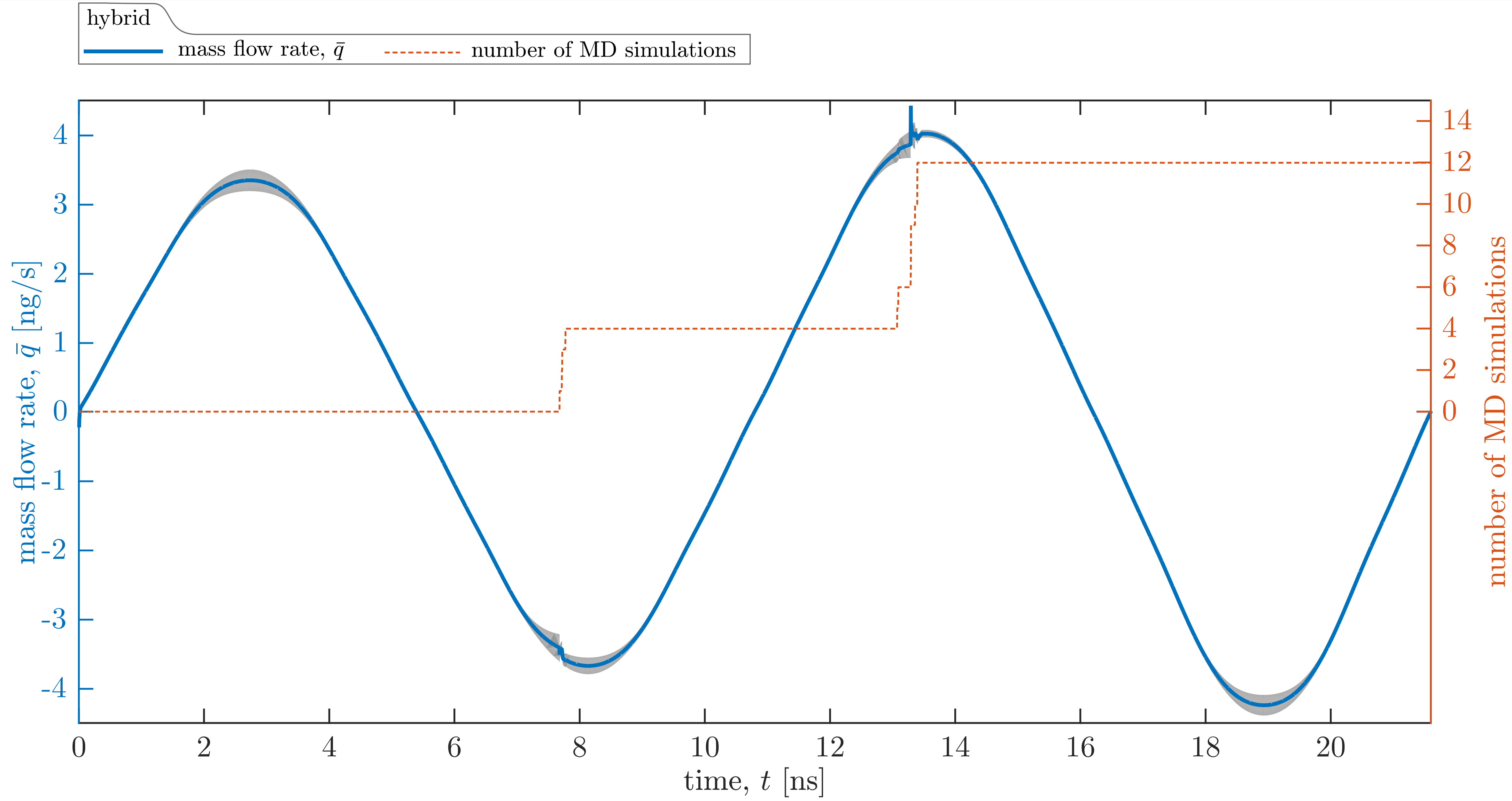
Results – expanding a database: subdomains



Results – expanding a database: forcing functions



Results – expanding a database: forcing functions



Conclusions

- Near-optimal information efficiency.
- Potential for uncertainty quantification.
- Strong agreement with full MD solutions.
- Dramatically enhanced computational speed (when database is extensive).
- Uncertainty threshold is a trade-off between accuracy and efficiency.
- Constructing an initial database is likely beneficial.
- Future work includes making the subdomain selection “smarter” and applying our algorithm to more complex engineering problems.

Thanks for listening