ADAPTIVE SEQUENTIAL MONTE CARLO APPROACH FOR REAL-TIME APPLICATIONS

Imperial College London

Thomas C.P. Chau[1], Wayne Luk[1], Peter Y.K. Cheung[2], Alison Eele[3], Jan Maciejowski[3]

[1]Department of Computing, [2]Department of Electrical and Electronic Engineering
Imperial College London, United Kingdom
[3]Department of Engineering, University of Cambridge, United Kingdom

Introduction

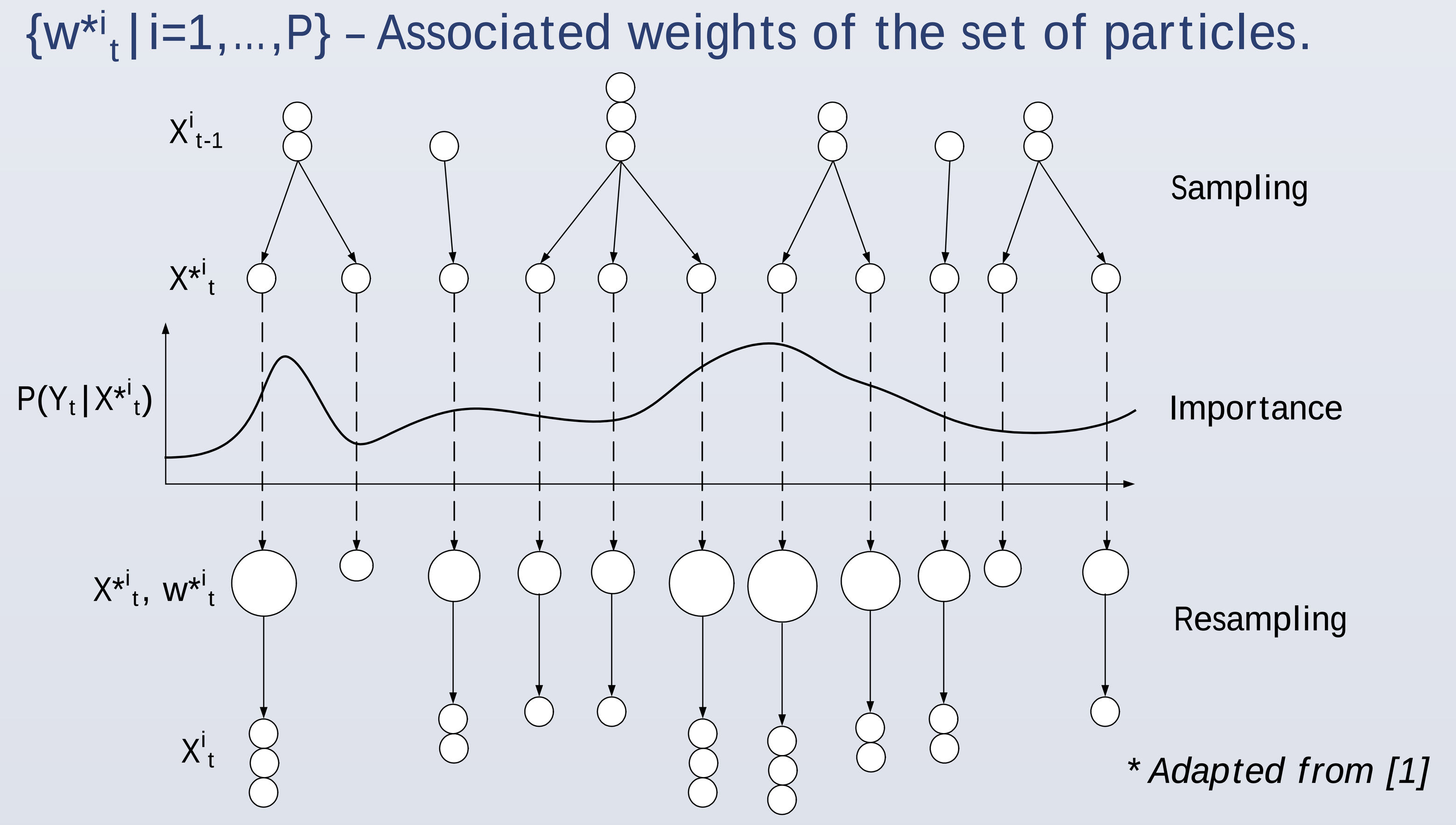
- Sequential Monte Carlo (SMC) methods: estimate states of dynamic systems using weighted particles.
- Adaptive real-time approach: analyse run-time workload for dynamic task allocation to processors on FPGA to improve energy efficiency.
- Example: dynamic energy consumption of robot localisation is reduced by up to 70% without affecting solution quality.

Motivation and Objective

- Real-time systems are often energy constrained.
- Events often occur at regular intervals for real-time applications.
- Objective: reduce energy by exploiting slack time when system is idle.

Sequential Monte Carlo Methods

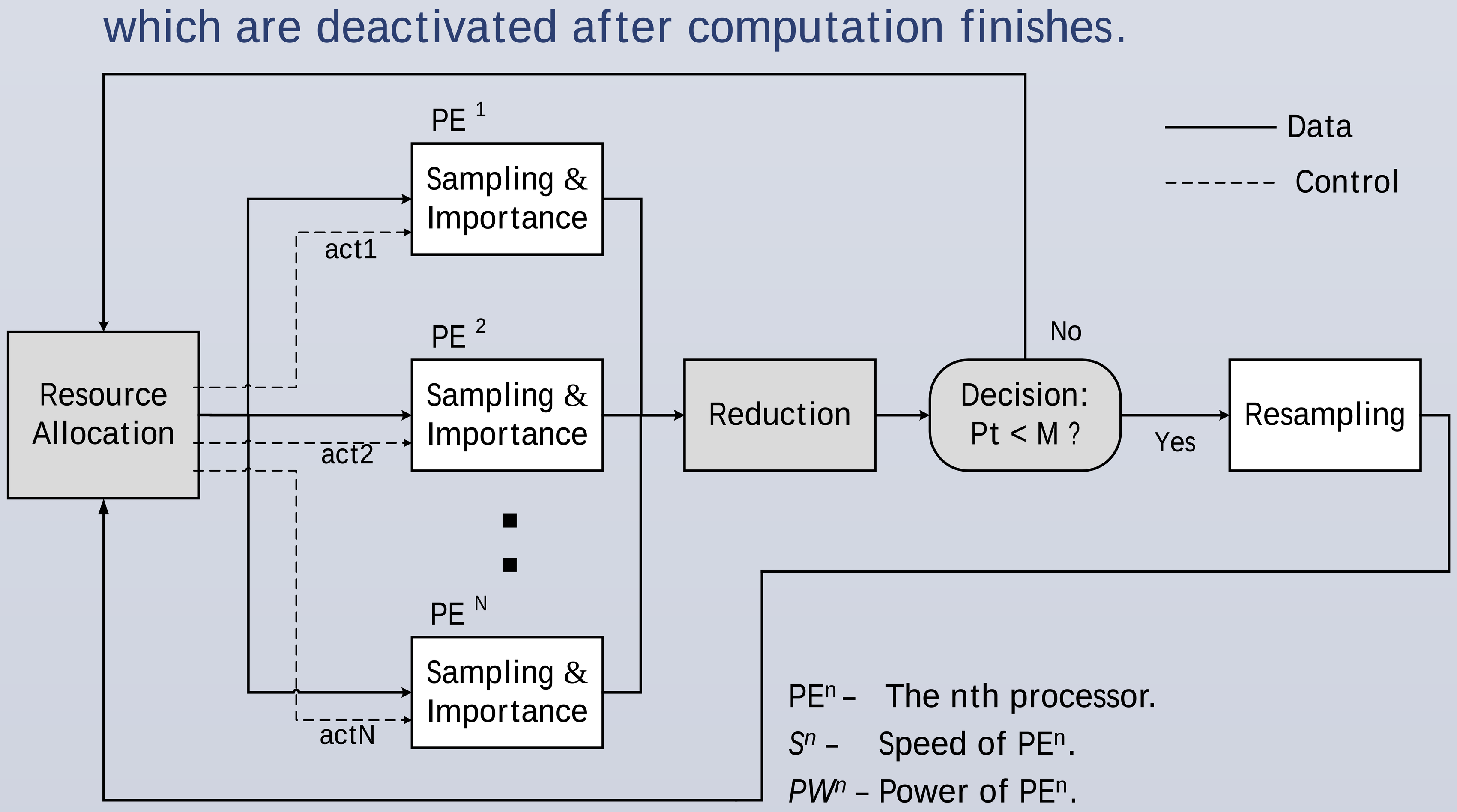
- X_t – State of the system at time t.
- Y_t – Observation of the system at time t.
- $\{x^i_t | i=1,\ldots,P\}$ – A set of P particles representing X_t.
- $\{w^{*i}_t | i=1,\ldots,P\}$ – Associated weights of the set of particles.



** Adapted from [1]*

Adaptive SMC Approach

- Sampling and Importance: particles are distributed to *N* processors for parallel processing.
- Reduction: particles with low weight are eliminated, i.e. workload is reduced.
- Resampling: the set of particles are restored if the particle count (P_t) drops below a threshold (M).
- Resource allocation: particles are distributed to processors which are deactivated after computation finishes.



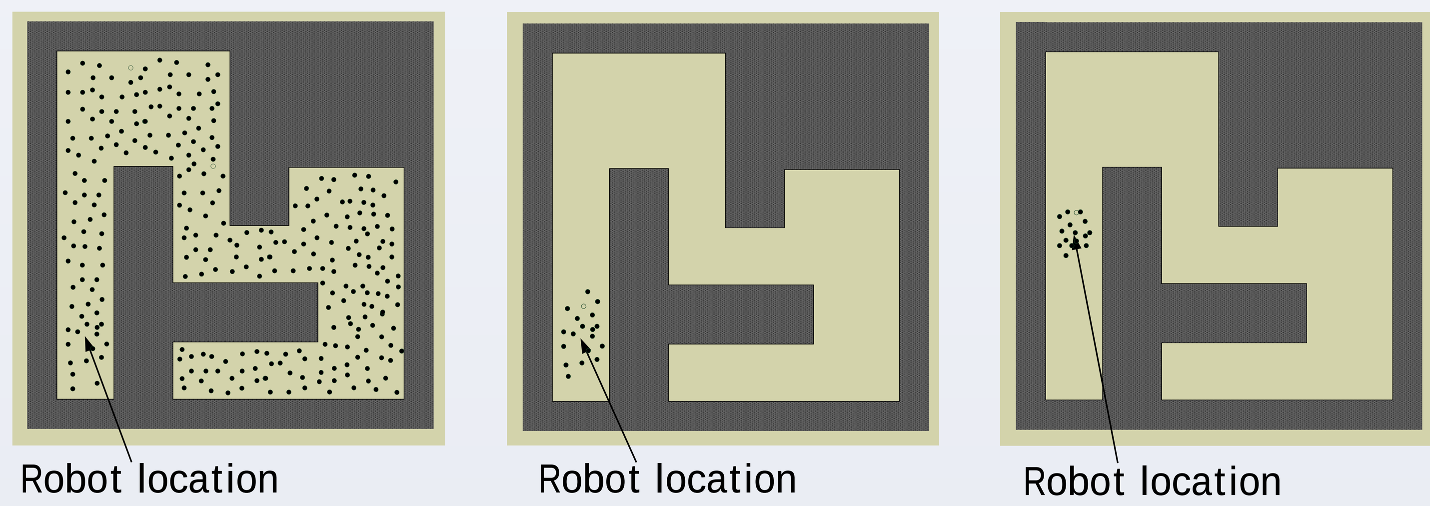
PE^n – The nth processor.
S^n – Speed of PE^n.
PW^n – Power of PE^n.
ρ^n – Number of particles distributed to PE^n.
P – Total number of particles.
T_{RT} – Real-time constraint.

Determine $\{\rho^n | n=1,\ldots,N\} = P \cdot \frac{\frac{S^n}{PW^n}}{\sum_{k=1}^{N} \frac{S^k}{PW^k}}$

subject to $\sum_{n=1}^{N} \rho^n = P$ and $T(P) \leq T_{RT}$

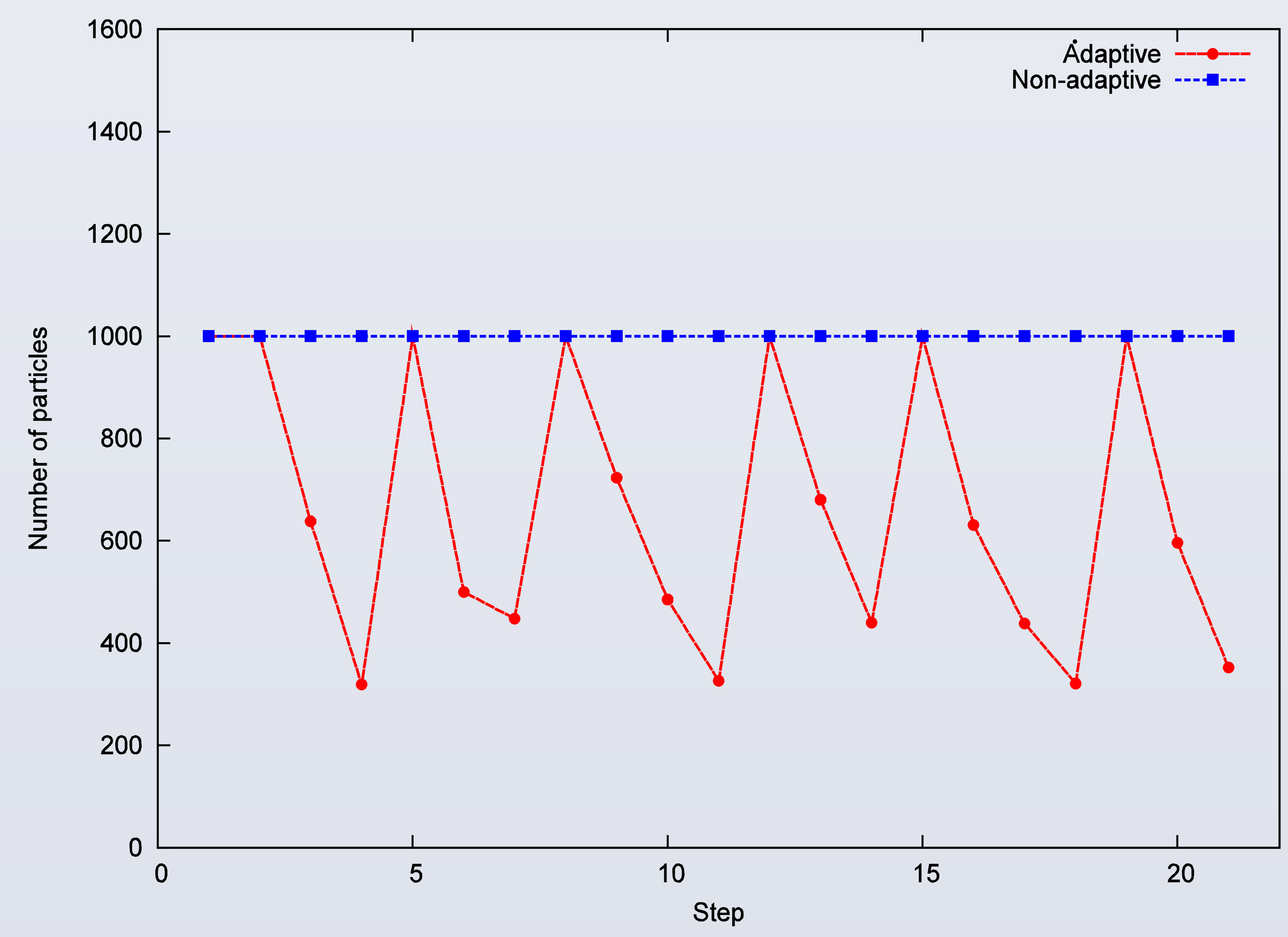
Results

- Platform: Altera DE4 development board with a Stratix IV EP4SGX530 FPGA running at 100MHz.
- Example application: localisation of mobile robots.
- More particles are need for global positioning than local tracking.

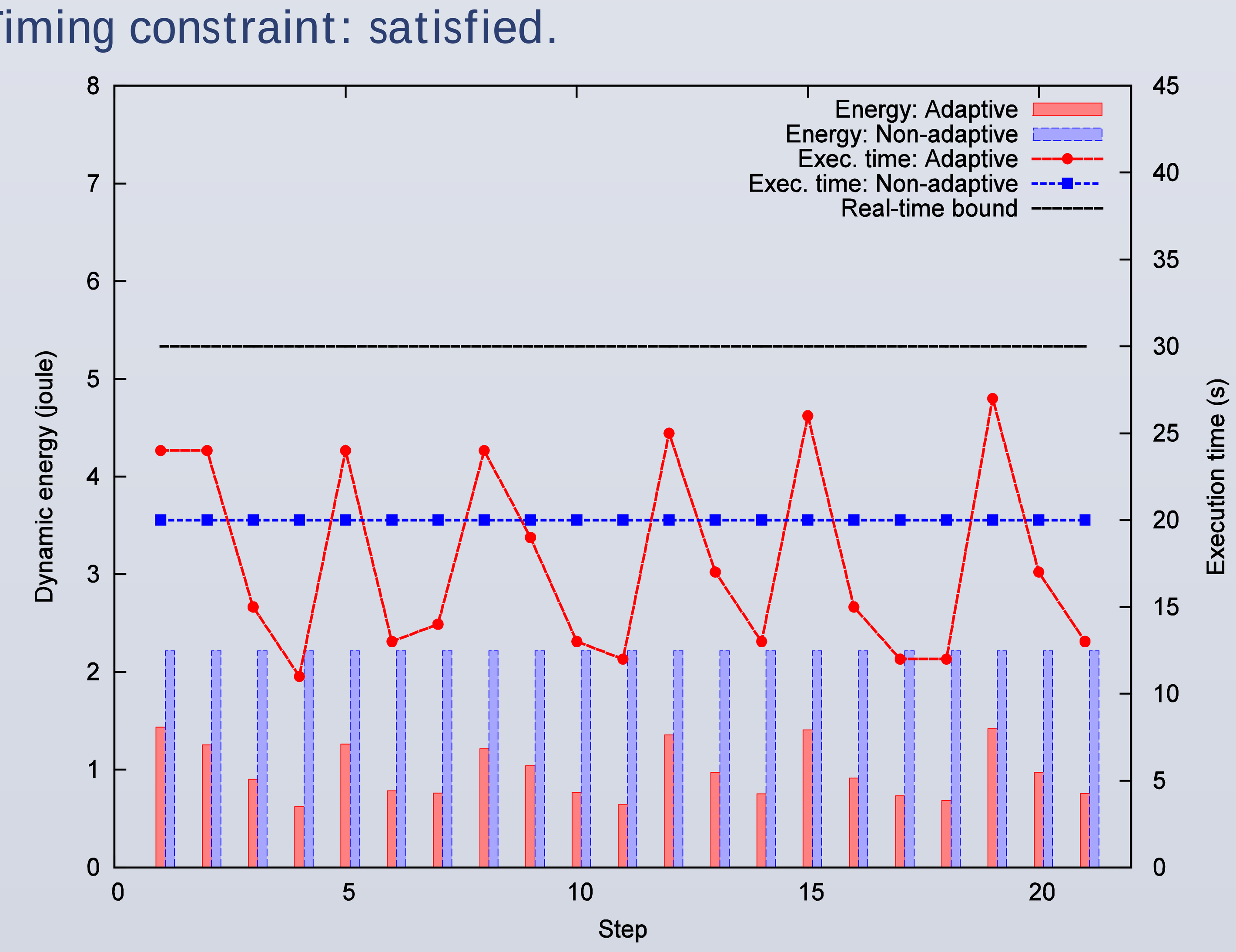


Start | Global positioning | Local tracking

- Computational complexity: reduced as the size of particle set changes dynamically.



- Dynamic energy consumption: reduced by 35-70%.
- Timing constraint: satisfied.



Summary

- Estimate states of dynamic systems: weighted particles.
- Dynamic workload: run-time task allocation to processors.
- Reduced energy: timing and quality of results maintained.

[1] M. Happe, E. Lubbers, and M. Platzner, "A Multithreaded Framework for Sequential Monte Carlo Methods on CPU/FPGA Platforms," International Workshop on Applied Reconfigurable Computing, 2009.

This work is supported in part by the European Union Seventh Framework Programme under grant agreement number 248976, 257906 and 287804, by UK EPSRC, by the Croucher Foundation, and by Altera.