Problem definition

- Given a data set $X = \{x_1, x_2, \ldots, x_n\} \in R^m$ and a label set $L = \{1, 2, \ldots, c\}$, some samples x_i are labeled as $y_i \in L$ and some are unlabeled as $y_i = \varnothing$
- The goal is to provide a label to these unlabeled samples
- Define a graph $G = (V,E)$, with $V = \{v_1, v_2, \ldots, v_n\}$
  - Each node v_i corresponds to a sample x_i
- An affinity matrix W defines the weight between the edges in E as follows:

$$
\begin{aligned}
W_{ij} &= \exp -||x_i - x_j||^2/2\sigma^2 \quad \text{if} \quad i \neq j, \\
W_{ii} &= 0,
\end{aligned}
$$

Model definition

- Set of particles $P = (\rho_1, \rho_2, \ldots, \rho_c)$
  - Each particle corresponds to a label in *L*
- Particle variables:
  - Node being visited by the particle $\rho_j^v(t) \in V$
  - Particle potential $\rho_j^\omega \in [\omega_{\min} \quad \omega_{\max}]$
  - Target node by the particle $\rho_j^\tau(t) \in V$
- Node variables:
  - Owner particle $v_i^\rho(t) \in P \quad v_i^\rho(t) = \arg\max_j v_i^{\omega_j}(t).$
  - Level of ownership $v_i^\omega(t) = \{v_i^{\omega_1}(t), v_i^{\omega_2}(t), \ldots, v_i^{\omega_c}(t)\}$

$$v_i^{\omega_j}(t) \in [\omega_{\min} \quad \omega_{\max}]$$

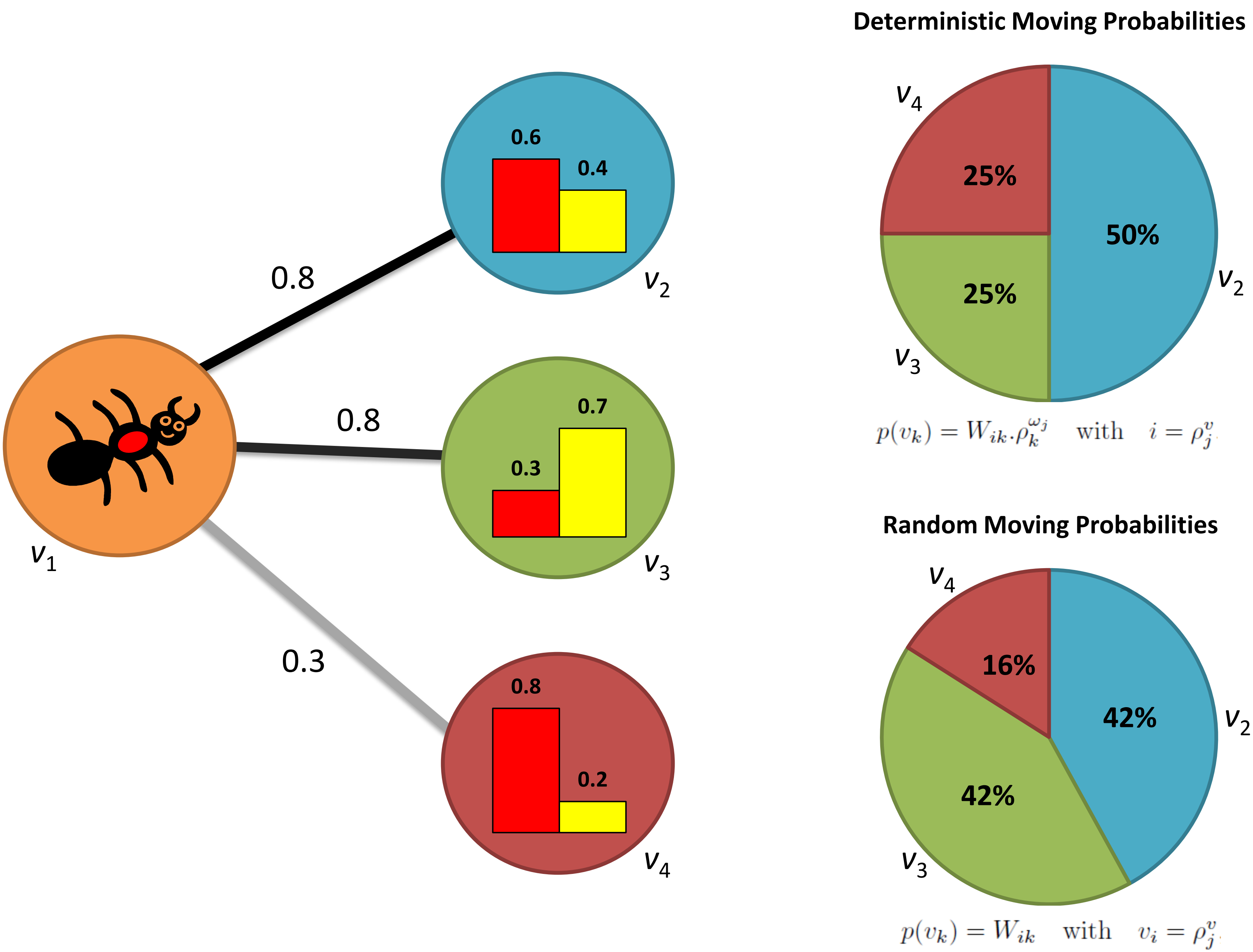
Variables Initialization

$$v_i^{\omega_j}(0) = \begin{cases} \omega_{\max} & \text{if } y_i = j \\ \omega_{\min} & \text{if } y_i \neq j \text{ and } y_i \neq \emptyset \\ \omega_{\min} + (\frac{\omega_{\max} - \omega_{\min}}{c}) & \text{if } y_i = \emptyset \end{cases}$$

$$\sum_{j=1}^{c} v_i^{\omega_j} = \omega_{max} + \omega_{\min}(c-1)$$

$$\rho_j^v(0) = \{v_i | y_i = j\}$$

$$\rho_j^\omega(0) = \omega_{\max}$$



v_1
0.8
0.8
0.3
v_2
0.6
0.4
v_3
0.3
0.7
v_4
0.8
0.2
Deterministic Moving Probabilities
v_4 25%
v_3 25%
v_2 50%
$p(v_k) = W_{ik} . \rho_k^{\omega_j}$ with $i = \rho_j^v$
Random Moving Probabilities
v_4 16%
v_3 42%
v_2 42%
$p(v_k) = W_{ik}$ with $v_i = \rho_j^v$

Particle and Nodes Dynamics

- Node dynamics
  - For each node v_i selected by a particle ρ_j as its target $\rho_j^{\tau}(t)$

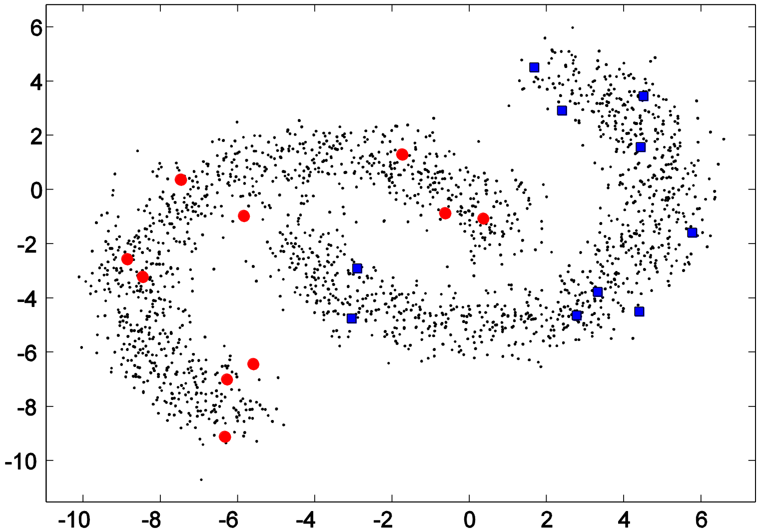
$$v_i^{\omega_k}(t+1) = \begin{cases} v_i^{\omega_k}(t) & \text{if } y_i \neq \emptyset \\ \max\{\omega_{\min}, v_i^{\omega_k}(t) - \frac{\Delta_v \rho_j^{\omega}(t)}{c-1}\} & \text{if } y_i = \emptyset \text{ and } k \neq j \\ v_i^{\omega_k}(t) + \sum_{q \neq k} v_i^{\omega_q}(t) - v_i^{\omega_q}(t+1) & \text{if } y_i = \emptyset \text{ and } k = j \end{cases}$$

- Particle dynamics

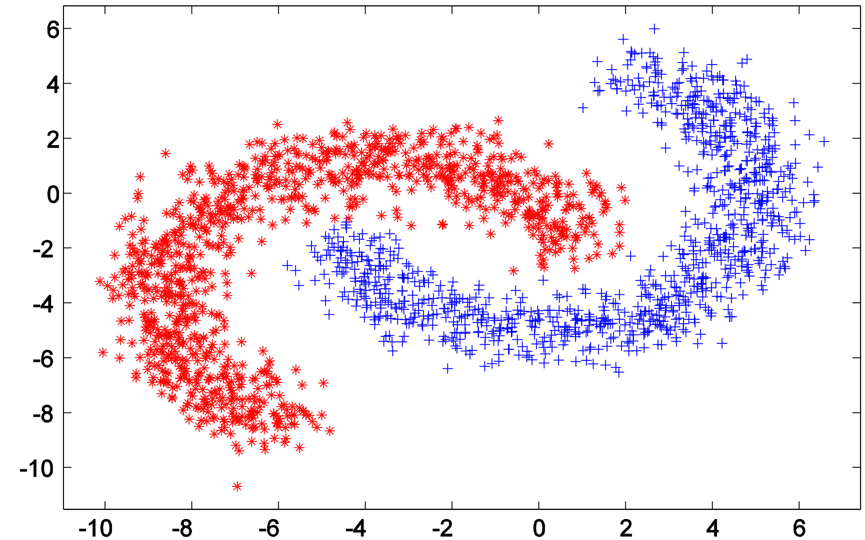
$$\rho_j^{\omega}(t+1) = v_i^{\omega_j}(t+1) \quad \text{with} \quad v_i(t+1) = \rho_j^{\tau}(t+1)$$

$$\rho_j^{v}(t+1) = \begin{cases} \rho_j^{\tau}(t+1) & \text{if } v_i^{\rho}(t+1) = \rho_j \\ \rho_j^{v}(t) & \text{if } v_i^{\rho}(t+1) \neq \rho_j \end{cases}$$

Simulation Results: Banana-Shaped data set



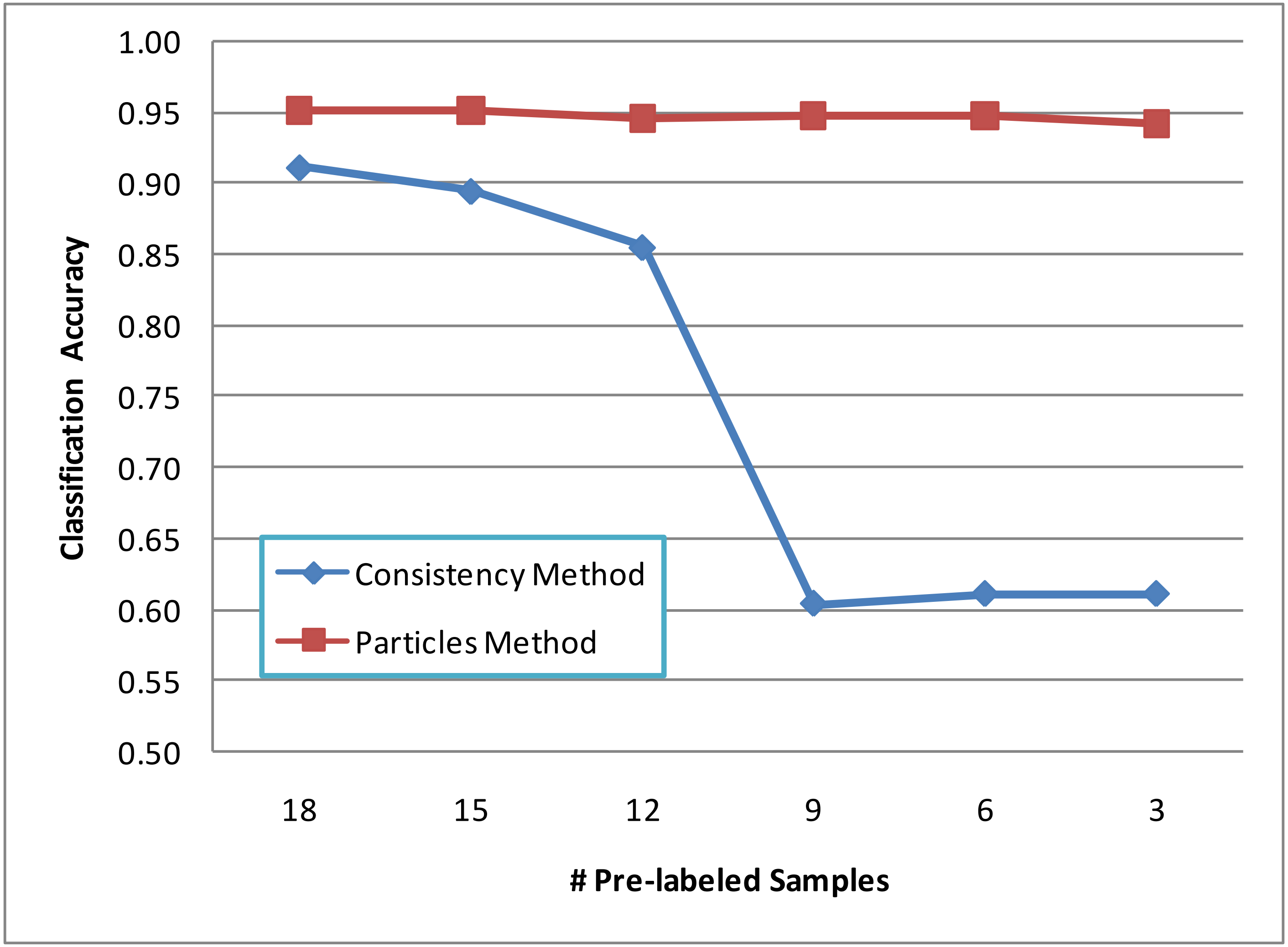
(a) Toy Data (Banana-Shaped)



(b) Classication results

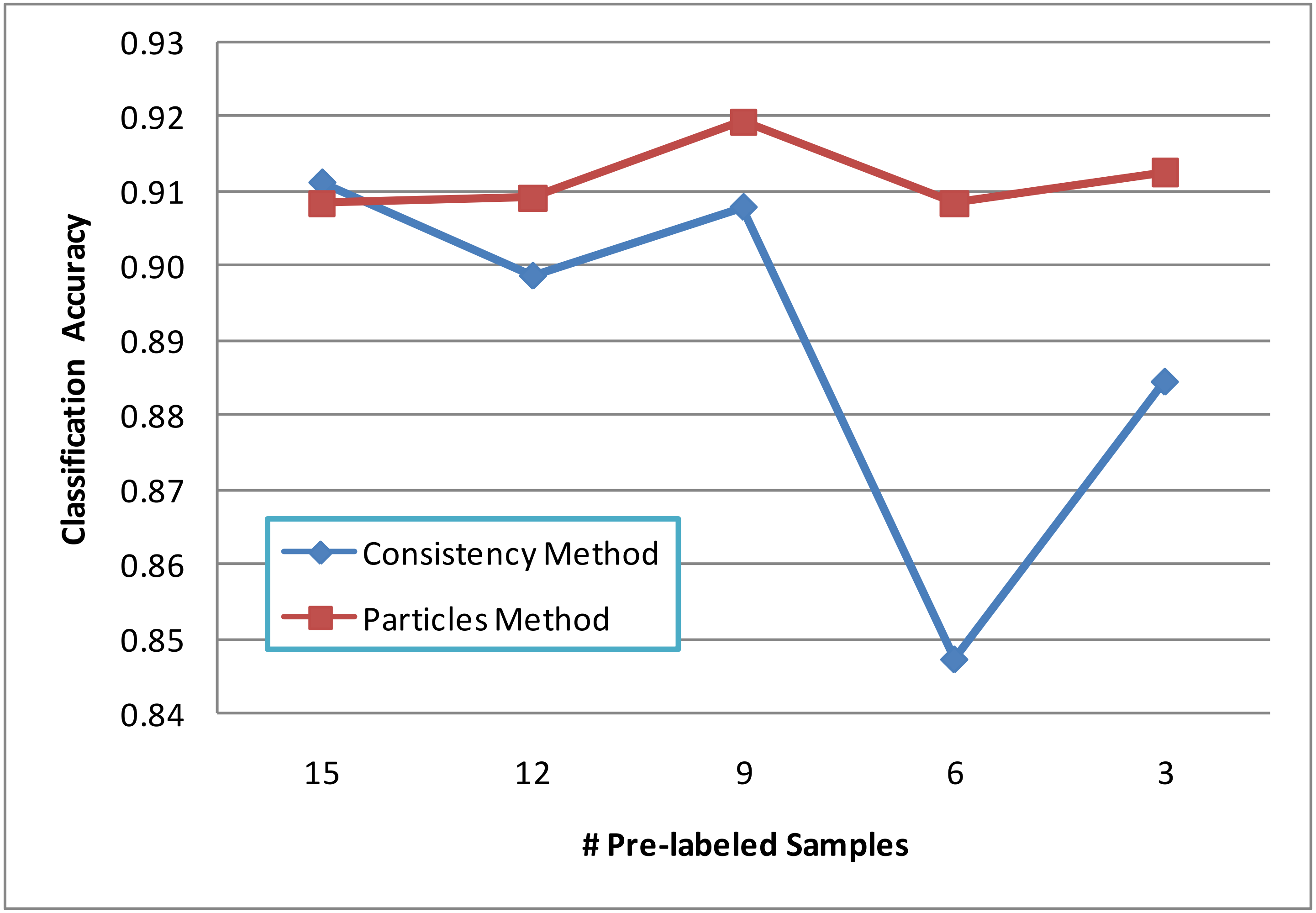
Classification of the banana-shaped patterns. (a) toy data set with 2000 samples divided in two classes, 20 samples are pre-labeled (red circles and blue squares) (b) classification achieved by the proposed method

Simulation Results: Iris data set (UCI)



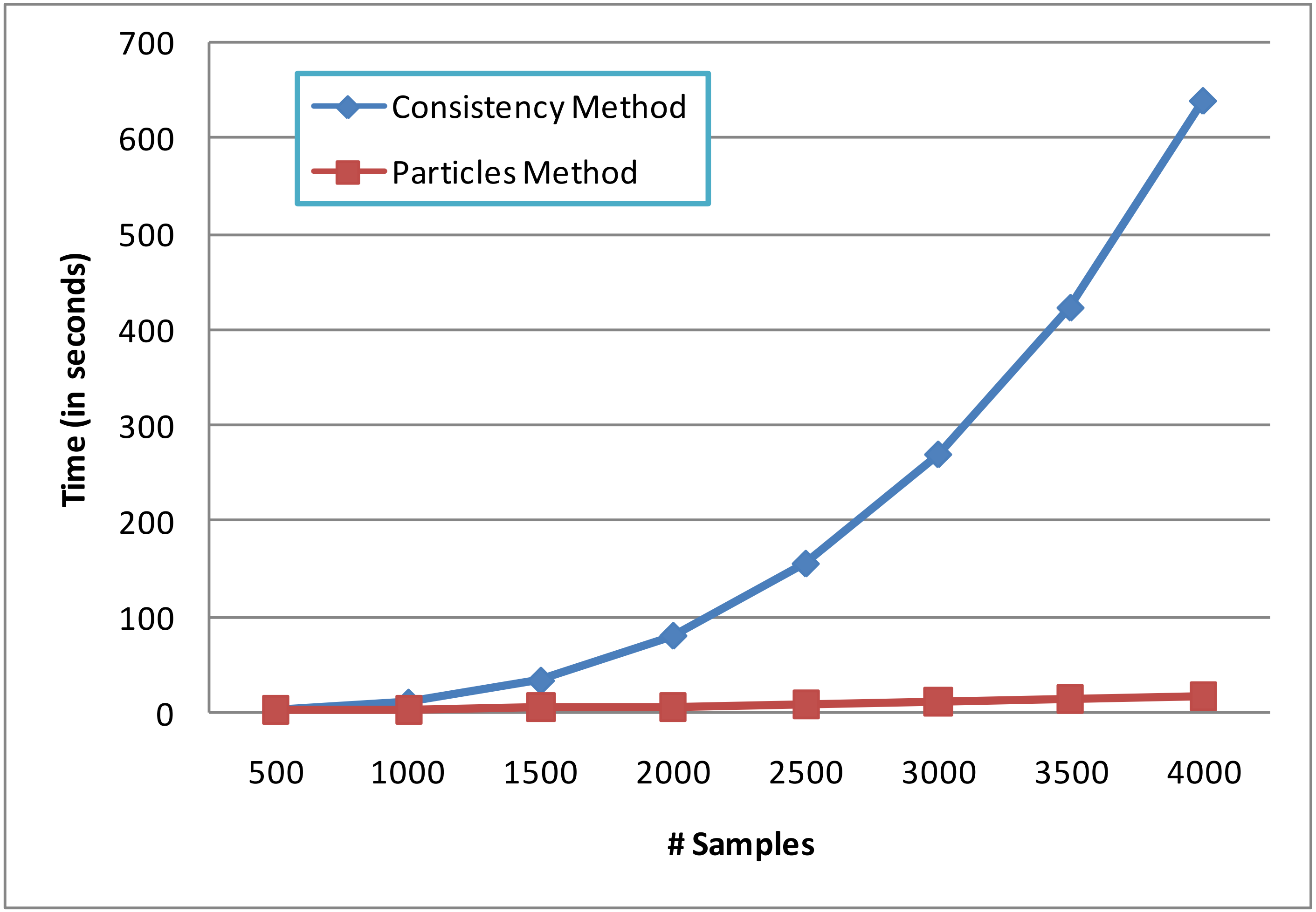
Classification Accuracy in the Iris data set with different number of pre-labeled samples

Simulation Results: Wine data set (UCI)



Classification Accuracy in the Wine data set with different number of pre-labeled samples

Simulation Results: execution time



Time elapsed to reach 95% correct classification with data sets of different sizes.

* Banana-shaped classes, 10% pre-labeled samples