Why Dempster’s rule doesn’t behave as Bayes rule with Informative Priors

Jean Dezert - The French Aerospace Lab, ONERA, Palaiseau, France.
Albena Tchamova - Inst. of I&C Tech., Bulg. Acad. of Sciences, Sofia, Bulgaria.
Deqiang Han - Inst. of Integrated Automation, Xi’an Jiaotong Univ., Xi’an, China.
Jean-Marc Tacnet - IRSTEA, UR ETGR, Saint-Martin d’Hères, France.
Purpose of this paper: We prove why Dempster’s fusion rule is incompatible with Bayes fusion rule in general (i.e. when prior information is informative)

1 - Conditional probabilities and Bayes fusion rule
   Symmetrization of Bayes fusion rule
   Properties of Bayes fusion rule
   Symbolic representation

2 - Belief functions and Dempster’s fusion rule

3 - Analysis of compatibility between Dempster’s and Bayes rules
   Examples

4 - Conclusions & References
Conditional probabilities

Given two random events $X$ and $Z$ taking values in $\Theta(X) \triangleq \{x_i, i = 1, 2, \ldots, N\}$ and $\Theta(Z) \triangleq \{z_j, i = j, 2, \ldots, M\}$ with $P(X)>0$ and $P(Z)>0$, one defines the conditional probabilities by

$$P(X|Z) \triangleq \frac{P(X \cap Z)}{P(Z)} \quad \text{and} \quad P(Z|X) \triangleq \frac{P(X \cap Z)}{P(X)}$$

Bayes Theorem

From above definitions, one gets, $P(X \cap Z) = P(X|Z)P(Z) = P(Z|X)P(X)$. hence

$$P(X|Z) = \frac{P(Z|X)P(X)}{P(Z)} \quad \text{and} \quad P(Z|X) = \frac{P(X|Z)P(Z)}{P(X)}$$

$$P(Z) = \sum_{i=1}^{N} P(Z|X = x_i)P(X = x_i)$$

$$P(X) = \sum_{j=1}^{M} P(X|Z = z_j)P(Z = z_j)$$
1 - Conditional probabilities and Bayes fusion rule

The «Fusion» challenge:

How to compute $P(X|Z_1 \cap Z_2)$ knowing $P(X|Z_1)$ and $P(X|Z_2)$?

The exact solution requires the knowledge of $P(X)$ and $P(X|Z_1 \cup Z_2)$ (usually unknown)

**Approximate (naive) solution:** obtained with statistical conditional independence assumption (A1)

$$(A1): \quad P(Z_1 \cap Z_2|X) = P(Z_1|X)P(Z_2|X)$$

**Bayes (parallel) fusion rule** Using (A1) and Bayes Theorem, one gets the Bayes fusion rule

$$P(X|Z_1 \cap Z_2) = \frac{P(X|Z_1)P(X|Z_2)}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_2)}{P(X=x_i)}} = \frac{1}{K(X, Z_1, Z_2)} \cdot P(X|Z_1) \cdot P(X|Z_2)$$

with the normalization factor

$$K(X, Z_1, Z_2) \triangleq P(X) \cdot \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_2)}{P(X=x_i)}$$
Conditional probabilities and Bayes fusion rule

Symmetrization of Bayes (parallel) fusion rule: 2 sources case

\[
P(X|Z_1 \cap Z_2) = \frac{P(X|Z_1) \cdot P(X|Z_2)}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}}} = \frac{1}{K'(Z_1, Z_2)} \cdot \frac{P(X|Z_1)}{\sqrt{P(X)}} \cdot \frac{P(X|Z_2)}{\sqrt{P(X)}}
\]

with the normalization factor

\[
K'(Z_1, Z_2) \triangleq \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}} = GA_2
\]

The Bayes fusion is in fact the ratio of the agreement factor \(A_2\) over the global agreement factor \(GA_2\)

\[
P(X = x_i|Z_1 \cap Z_2) = \frac{A_2(X = x_i)}{GA_2}
\]

\[
A_2(X = x_i) \triangleq \frac{P(X = x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X = x_i|Z_2)}{\sqrt{P(X=x_i)}}
\]

\[
GA_2 \triangleq \sum_{i_1, i_2=1|i_1 \neq i_2}^{N} \frac{P(X = x_{i_1}|Z_1)}{\sqrt{P(X=x_{i_1})}} \cdot \frac{P(X = x_{i_2}|Z_2)}{\sqrt{P(X=x_{i_2})}}
\]

The degree of conflict can be defined by

\[
GC_2 \triangleq \sum_{i_1, i_2=1|i_1 \neq i_2}^{N} \frac{P(X = x_{i_1}|Z_1)}{\sqrt{P(X=x_{i_1})}} \cdot \frac{P(X = x_{i_2}|Z_2)}{\sqrt{P(X=x_{i_2})}}
\]
Symmetrization of Bayes (parallel) fusion rule: General case (s>2 sources)

The Bayes fusion is in fact the ratio of the agreement factor $A_s$ over the global agreement factor $G_{A_s}$

The degree of conflict can be defined by

$$GC_s = \sum_{i_1, \ldots, i_s=1}^{N} \frac{P(X = x_{i_1} | Z_1)}{\sqrt{P(X = x_{i_1})}} \cdots \frac{P(X = x_{i_s} | Z_s)}{\sqrt{P(X = x_{i_s})}} - G_{A_s}$$
1 - Conditional probabilities and Bayes fusion rule

Symbolic representation of (naive parallel) Bayes fusion rule

For 2 sources and with prior $P(X)$

$$P(X|Z_1 \cap Z_2) = \text{Bayes}(P(X|Z_1), P(X|Z_2); P(X))$$

For $s>2$ sources and with prior $P(X)$

$$P(X|Z_1 \cap \ldots \cap Z_s) = \text{Bayes}(P(X|Z_1), \ldots, P(X|Z_s); P(X))$$

Particular case : $P(X)$ is the uniform prior distribution

If $P(X)$ is the uniform pmf then \( \sqrt{P(X = x_i)} = \sqrt{1/N} \) and \( \sqrt{P(X = x_i)} = \sqrt{1/N} \)

simplify in previous formulas, and one gets

$$P(X|Z_1 \cap Z_2) = \frac{\sum_{i=1}^{N} P(X = x_i|Z_1)P(X = x_i|Z_2)}{P(X = x_i|Z_1)P(X = x_i|Z_2)} = \text{Bayes}(P(X|Z_1), P(X|Z_2))$$

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\prod_{k=1}^{s} P(X|Z_k)}{\sum_{i=1}^{N} \prod_{k=1}^{s} P(X = x_i|Z_k)} = \text{Bayes}(P(X|Z_1), \ldots, P(X|Z_s))$$
1 - Conditional probabilities and Bayes fusion rule

Particular case: \( P(X) \) is the uniform prior distribution

The previous formulas can be rewritten as for \( s \geq 2 \)

\[
P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\prod_{k=1}^{s} P(X|Z_k)}{GA_{s}^{unif}} = \frac{\prod_{k=1}^{s} P(X|Z_k)}{1 - GC_{s}^{unif}}
\]

where

\[
GA_{s}^{unif} = \sum_{i_1, \ldots, i_s=1}^{N} P(X=x_{i_1}|Z_1) \ldots P(X=x_{i_s}|Z_s) \quad \text{(global agreement)}
\]

\[
GC_{s}^{unif} = 1 - GA_{s}^{unif} \quad \text{(global conflict)}
\]

The naive parallel Bayes Fusion rule Eq. (1) is in fact similar to Dempster’s rule of combination using the classical belief alike notations in this particular case, that is

\[
GA_{s}^{unif} = \sum_{x_{i_1}, \ldots, x_{i_s} \in \Theta(X) \atop x_{i_1} \cap \ldots \cap x_{i_s} \neq \emptyset} P(X=x_{i_1}|Z_1) \ldots P(X=x_{i_s}|Z_s)
\]

\[
GC_{s}^{unif} = \sum_{x_{i_1}, \ldots, x_{i_s} \in \Theta(X) \atop x_{i_1} \cap \ldots \cap x_{i_s} = \emptyset} P(X=x_{i_1}|Z_1) \ldots P(X=x_{i_s}|Z_s)
\]
Main Properties of Bayes fusion rule

- **(P1)**: The pmf $P(X)$ is a neutral element of Bayes fusion rule when combining only two sources.

$$\text{Bayes}(P(X|Z_1), P(X); P(X)) = P(X|Z_1) \quad \text{and} \quad \text{Bayes}(P(X), P(X|Z_2); P(X)) = P(X|Z_2)$$

$P(X)$ is not a neutral element of Bayes fusion when combining 3 sources (or more)

$$\text{Bayes}(P(X|Z_1), P(X|Z_2), P(X); P(X)) \neq \text{Bayes}(P(X|Z_1), P(X|Z_2); P(X))$$

- **(P2)**: Bayes fusion rule is in general not idempotent.

$$\text{Bayes}(P(X|Z_1), P(X|Z_1); P(X)) \neq P(X|Z_1)$$
Main Properties of Bayes fusion rule

- **(P3): Bayes fusion rule is in general not associative.**
  
  In the paper, we give an example with $P(X)$ non uniform showing that

  \[
  \begin{align*}
  P(X|(Z_1 \cap Z_2) \cap Z_3) &\neq P(X|Z_1 \cap Z_2 \cap Z_3) \\
  P(X|Z_1 \cap (Z_2 \cap Z_3)) &\neq P(X|Z_1 \cap Z_2 \cap Z_3) \\
  P(X|Z_2 \cap (Z_1 \cap Z_3)) &\neq P(X|Z_1 \cap Z_2 \cap Z_3)
  \end{align*}
  \]

- **(P4): Bayes fusion rule is associative if and only if $P(X)$ is uniform.**

  From Bayes fusion formula with $P(X)$ uniform, the associativity is satisfied because

  \[
  
  \text{Bayes}(P(X|Z_1), \ldots, P(X|Z_s)) = \text{Bayes}(\text{Bayes}(P(X|Z_1), \ldots, P(X|Z_{s-1})), P(X|Z_s))
  \]

  and independently of the choice of the decomposition (grouping) of the sources.
Main Properties of Bayes fusion rule

• (P5): The levels of global agreement and global conflict between the sources do not matter in Bayes fusion rule.

Example: \( P(X = x_1) = 0.2 \) and \( P(X = x_2) = 0.8 \)

Case 1:
\[
\begin{align*}
P(X = x_1 | Z_1) &\approx 0.0607 \text{ and } P(X = x_2 | Z_1) \approx 0.9393 \\
P(X = x_1 | Z_2) &\approx 0.6593 \text{ and } P(X = x_2 | Z_2) \approx 0.3407
\end{align*}
\]

Bayes fusion gives
\[
\begin{align*}
P(X = x_1 | Z_1 \cap Z_2) &= \frac{0.0607 \cdot 0.6593 / 0.2}{(0.0607 \cdot 0.6593 / 0.2) + (0.9393 \cdot 0.3407 / 0.8)} = 1/3 \\
P(X = x_2 | Z_1 \cap Z_2) &= \frac{0.9393 \cdot 0.3407 / 0.8}{(0.0607 \cdot 0.6593 / 0.2) + (0.9393 \cdot 0.3407 / 0.8)} = 2/3
\end{align*}
\]

Case 2:
\[
\begin{align*}
P'(X = x_1 | Z_1) &\approx 0.8360 \text{ and } P'(X = x_2 | Z_1) \approx 0.1640 \\
P'(X = x_1 | Z_2) &\approx 0.0240 \text{ and } P'(X = x_2 | Z_2) \approx 0.9760
\end{align*}
\]

Therefore \( \text{Bayes}(P(X | Z_1), P(X | Z_2); P(X)) = \text{Bayes}(P'(X | Z_1), P'(X | Z_2); P(X)) \)

even if \( \begin{cases} (GA_2 = 0.60) \neq (GA'_2 = 0.30) \\ (GC_2 = 1.60) \neq (GC'_2 = 2.05) \end{cases} \)

What really matters is the distribution of all relative (ratios) agreement factors.
2 - Belief functions and Dempster’s fusion rule

Belief functions [Shafer 1976]

Frame of discernment: $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$ (N exhaustive and exclusive elements)

Basic belief assignment (bba): $m(.) : 2^\Theta \rightarrow [0, 1]$ \hspace{1cm} $m(\emptyset) = 0$ and $\sum_{X \in 2^\Theta} m(X) = 1$

Focal element $X$: iff $m(X) > 0$

Belief $Bel(.)$ and Plausibility $Pl(.)$ functions: (in one-to-one correspondence with $m(.)$)

$$\begin{cases} 
Bel(X) = \sum_{Y \in 2^\Theta | Y \subseteq X} m(Y) \\
Pl(X) = \sum_{Y \in 2^\Theta | X \cap Y \neq \emptyset} m(Y)
\end{cases}$$

$Bel(.)$ is subadditive since $\sum_{\theta_i \in \Theta} Bel(\theta_i) \leq 1$.

$Pl(.)$ is superadditive since $\sum_{\theta_i \in \Theta} Pl(\theta_i) \geq 1$.

$\forall X \in 2^\Theta, Bel(X) \leq P(X) \leq Pl(X)$

Vacuous bba: for modeling a full ignorant source of evidence

$m_v(X) = 0$ if $X \neq \Theta$, and $m_v(\theta_1 \cup \theta_2 \cup \ldots \cup \theta_N) = 1$

Bayesian bba: iff focal elements of $m(.)$ are singletons. In such case $Bel(X) = P(X) = Pl(X)$
Dempster’s rule of combination of $s \geq 2$ independent (not totally conflicting) sources

Dempster’s rule is often called Dempster-Shafer (DS) rule in literature

\[ m_{DS}(\emptyset) \triangleq 0 \]
\[ m_{DS}(X) \triangleq \frac{m_{12\ldots s}(X)}{1 - K_{12\ldots s}} \quad \forall X \neq \emptyset \in 2^\Theta \]

where
\[ m_{12\ldots s}(X) \triangleq \sum_{\substack{X_1, X_2, \ldots, X_s \in 2^\Theta \\ X_1 \cap X_2 \cap \ldots \cap X_s = X}} m_1(X_1)m_2(X_2) \ldots m_s(X_s) \quad \text{(conj. agreement)} \]
\[ K_{12\ldots s} \triangleq \sum_{\substack{X_1, X_2, \ldots, X_s \in 2^\Theta \\ X_1 \cap X_2 \cap \ldots \cap X_s = \emptyset}} m_1(X_1)m_2(X_2) \ldots m_s(X_s) \quad \text{(global conflict)} \]

Properties:
- DS rule is commutative and associative
- The vacuous bba is the neutral element of DS rule

DS rule is often claimed a generalization of Bayes rule because when conditioning $m(.)$ by $m_Z(Z)=1$ with DS rule, one gets $PI(X|Z)=PI(X\land Z)/PI(Z)$. 
To make a fair analysis and comparison, we need to work with Bayesian bba’s because Bayes rule works with probabilities only. We consider the following Bayesian bba’s

\[
\begin{align*}
    m_1(.) &\triangleq \{ m_1(\theta_i) = P(X = x_i|Z_1), i = 1, 2, \ldots, N \} \\
    \vdots & \quad \vdots \\
    m_s(.) &\triangleq \{ m_s(\theta_i) = P(X = x_i|Z_s), i = 1, 2, \ldots, N \}
\end{align*}
\]

**Basic idea:**

If DS rule is a true (consistent) generalization of Bayes fusion rule, it must provide same results as Bayes rule when combining Bayesian bba's.

Otherwise DS rule cannot be claimed to be a generalization of Bayes fusion rule.

**What we prove:**

DS rule is compatible with Bayes rule ONLY when the prior is NOT informative (uniform or vacuous).

DS rule is incompatible with Bayes in general case (when prior is informative)

**Symbolic notation:**

\[
m_{DS}(.) = DS(m_1(.), \ldots, m_s(.); m_0(.))
\]

\[m_0(.)\text{ being the bba modeling the a priori belief}\]
3 - Analysis of compatibility of DS rule with Bayes rule

From Bayes and DS formulas, one sees that when \( m_i(.) \) are Bayesian bba's and they coincide with \( P(X|Z_i) \), one has

\[
DS(m_1(.), \ldots, m_s(.); m_0(.)) \equiv Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X))
\]

If and only if:
- \( P(X) \) is uniform
- \( m_0(.) \) is either the uniform Bayesian bba, or \( m_0(.) \) is the vacuous bba

Thus, DS rule is compatible with Bayes rule ONLY when the prior is NOT informative

In the more general cases when \( m_0(.)=P(X) \) is really informative (not uniform, nor vacuous), one has

\[
DS(m_1(.), \ldots, m_s(.); m_0(.)) \neq Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X))
\]

because Bayes and DS rules deal differently with informative priors, that is

- In DS rule, the prior bba \( m_0(.) \) is combined in a pure conjunctive (multiplicative) manner
- In Bayes rule, one divides each posterior bba \( m_i(x_j), i=1,2,\ldots,s \) by \( \sqrt{m_0(x_j)} \)
3 - Analysis of compatibility of DS rule with Bayes rule

Example with uniform & vacuous prior

\[ \Theta(X) = \{x_1, x_2, x_3\} \]

\[
\begin{aligned}
m_1(x_1) &= P(X = x_1 | Z_1) = 0.2 \\
m_1(x_2) &= P(X = x_2 | Z_1) = 0.3 \\
m_1(x_3) &= P(X = x_3 | Z_1) = 0.5
\end{aligned}
\]

\[
\begin{aligned}
m_2(x_1) &= P(X = x_1 | Z_2) = 0.5 \\
m_2(x_2) &= P(X = x_2 | Z_2) = 0.1 \\
m_2(x_3) &= P(X = x_3 | Z_2) = 0.4
\end{aligned}
\]

- If \( m_0(.) \) the vacuous bba

\[
\begin{aligned}
m_{DS}(x_1) &= \frac{1}{1-K_{12}^{vacuous}} m_1(x_1)m_2(x_1)m_0(x_1 \cup x_2 \cup x_3) \\
&= \frac{1}{1-0.67} \cdot 0.2 \cdot 0.5 \cdot 1 = \frac{0.10}{0.33} \approx 0.3030 \\
m_{DS}(x_2) &= \frac{1}{1-K_{12}^{vacuous}} m_1(x_2)m_2(x_2)m_0(x_1 \cup x_2 \cup x_3) \\
&= \frac{1}{1-0.67} \cdot 0.3 \cdot 0.1 \cdot 1 = \frac{0.03}{0.33} \approx 0.0909 \\
m_{DS}(x_3) &= \frac{1}{1-K_{12}^{vacuous}} m_1(x_3)m_2(x_3)m_0(x_1 \cup x_2 \cup x_3) \\
&= \frac{1}{1-0.67} \cdot 0.5 \cdot 0.4 \cdot 1 = \frac{0.20}{0.33} \approx 0.6061
\end{aligned}
\]

\[ K_{12}^{vacuous} = 0.67 \]

- If \( m_0(x_1) = m_0(x_2) = m_0(x_3) = 1/3 \) (uniform bba)

\[
\begin{aligned}
m_{DS}(x_1) &= \frac{1}{1-K_{12}^{uniform}} m_1(x_1)m_2(x_1)m_0(x_1) \\
&= \frac{1}{1-0.89} \cdot 0.2 \cdot 0.5 \cdot 1/3 = \frac{0.10/3}{0.11} \approx 0.3030 \\
m_{DS}(x_2) &= \frac{1}{1-K_{12}^{uniform}} m_1(x_2)m_2(x_2)m_0(x_2) \\
&= \frac{1}{1-0.89} \cdot 0.3 \cdot 0.1 \cdot 1/3 = \frac{0.03/3}{0.11} \approx 0.0909 \\
m_{DS}(x_3) &= \frac{1}{1-K_{12}^{uniform}} m_1(x_3)m_2(x_3)m_0(x_3) \\
&= \frac{1}{1-0.89} \cdot 0.5 \cdot 0.4 \cdot 1/3 = \frac{0.20/3}{0.11} \approx 0.6061
\end{aligned}
\]

\[ K_{12}^{uniform} = 0.89 \]

These results coincide with Bayes fusion even if the conflict levels differ because the uniform or vacuous prior bba’s do not bring helpful information to discriminate the states.
### Example with informative prior

\[ \Theta(X) = \{ x_1, x_2, x_3 \} \]

\[
\begin{align*}
m_1(x_1) &= P(X = x_1 | Z_1) = 0.2 \\
m_1(x_2) &= P(X = x_2 | Z_1) = 0.3 \\
m_1(x_3) &= P(X = x_3 | Z_1) = 0.5
\end{align*}
\]

\[
\begin{align*}
m_2(x_1) &= P(X = x_1 | Z_2) = 0.5 \\
m_2(x_2) &= P(X = x_2 | Z_2) = 0.1 \\
m_2(x_3) &= P(X = x_3 | Z_2) = 0.4
\end{align*}
\]

with informative prior bba/pmf:

\[
\begin{align*}
m_0(x_1) &= P(X = x_1) = 0.6 \\
m_0(x_2) &= P(X = x_2) = 0.3 \\
m_0(x_3) &= P(X = x_3) = 0.1
\end{align*}
\]

### Applying Bayes rule

\[
\begin{align*}
P(x_1 | Z_1 \cap Z_2) &= \frac{A_2(x_1)}{GA_2} = \frac{0.2 \cdot 0.5 / 0.6}{2.2667} = 0.1667 \\ &= \frac{2.2667}{2.2667} \approx 0.0735
\end{align*}
\]

\[
\begin{align*}
P(x_2 | Z_1 \cap Z_2) &= \frac{A_2(x_2)}{GA_2} = \frac{0.3 \cdot 0.1 / 0.3}{2.2667} = 0.1000 \\ &= \frac{2.2667}{2.2667} \approx 0.0441
\end{align*}
\]

\[
\begin{align*}
P(x_3 | Z_1 \cap Z_2) &= \frac{A_2(x_3)}{GA_2} = \frac{0.5 \cdot 0.4 / 0.1}{2.2667} = 0.0000 \\ &= \frac{2.2667}{2.2667} \approx 0.8824
\end{align*}
\]

### Applying DS rule

\[
\begin{align*}
m_{DS}(x_1) &= \frac{1}{1 - 0.9110} \cdot 0.2 \cdot 0.5 \cdot 0.6 = \frac{0.060}{0.089} \approx 0.6742 \\
m_{DS}(x_2) &= \frac{1}{1 - 0.9110} \cdot 0.3 \cdot 0.1 \cdot 0.3 = \frac{0.009}{0.089} \approx 0.1011 \\
m_{DS}(x_3) &= \frac{1}{1 - 0.9110} \cdot 0.5 \cdot 0.4 \cdot 0.1 = \frac{0.020}{0.089} \approx 0.2247
\end{align*}
\]

Therefore

\[ DS(m_1(.), \ldots, m_s(.); m_0(.)) \neq Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X)) \]
The following points have been shown in this work:

1 - How naive Bayesian parallel fusion rule can be expressed in a symmetrical form.

2 - Properties of Bayes fusion rule (non idempotency, non associativity, insensitivity to the level of conflict).

3 - Dempster’s rule is compatible with Bayes rule only if the bba’s of sources are Bayesian and when the prior is uniform, or vacuous (i.e. non informative).

4 - In General (with informative priors), Dempster’s rule is incompatible/inconsistent with Bayes fusion rule.

5 - Our analysis supports previous Mahler’s conclusions by providing a deeper analysis, proves, and examples.

Dempster’s rule is NOT a generalization of Bayesian fusion rule.
Main references


- J. Dezert, A. Tchamova, D. Han, J. Tacnet, Why Dempsters fusion rule is not a generalization of Bayes fusion rule, Proc. of Fusion 2013 Int. Conf., Istanbul, 2013 (accepted for publication).


This study was partly supported by the project AComIn, grant 316087, funded by the FP7 Capacity Programme and co-supported by Grant for State Key Program for Basic Research of China (973) (No. 2013CB329405), National Natural Science Foundation of China (No. 61104214, No. 61203222)