

基于最大 Lyapunov 指数不变性的混沌时间序列噪声水平估计*

姚天亮 刘海峰[†] 许建良 李伟锋

(华东理工大学煤气化教育部重点实验室, 上海 200237)

(2011年5月14日收到; 2011年8月9日收到修改稿)

提出了一种基于最大 Lyapunov 指数不变性的计算混沌时间序列噪声水平的新方法. 首先分析了噪声对相空间中两点距离的影响, 然后基于最大 Lyapunov 指数在不同维数的嵌入相空间不变的性质, 建立了估计噪声水平的方法. 仿真计算结果表明, 当噪声水平小于 10% 时, 估计值与真实值符合良好. 该方法对噪声分布类型不敏感, 是一种有效的混沌时间序列噪声估计方法.

关键词: 混沌, 时间序列, 最大 Lyapunov 指数, 噪声

PACS: 05.40.Ca, 05.45.Tp, 05.45.Pq

1 引言

确定性混沌已经在流体力学^[1-3]、生物医学^[4,5]以及金融学^[6]等科学领域得到广泛应用. 混沌时间序列分析是目前非线性科学的重要课题^[7-11], 然而, 各种测量方法获得的实际时间序列数据不可避免地含有各种噪声^[12-14]. 噪声水平估计是时间序列分析的重要内容. Schreiber^[15]通过考察噪声对关联积分的影响, 提出了一种基于误差函数的噪声水平估计方法, 该算法仅对小的 Gauss 测量噪声有效. Yu 等^[16]对由噪声时间序列计算得到的 Gauss 核关联积分进行非线性拟合, 可以同时估计出噪声水平及关联维. 后来, Jayawardena 等^[17]在此基础上发展了一种线性拟合算法, 由于该算法是基于 Gauss 白噪声得到的, 因此对噪声类型比较敏感. Urbanowicz 和 Holyst^[18,19]先后提出了基于粗糙纹理关联熵 (coarse-grained correlation entropy) 的噪声水平估计方法和基于重构相空间中最近邻域距离估计噪声水平的方法, 这两种算法可以估计大的噪声水平, 但结果精度有限. Strumik 等^[20]基于关联熵的标度性质, 提出了一种估计混沌时间序列噪声水平的方法. 最近, Moriya^[21]提出了一种基

于二项式滤波的噪声水平估计方法, 但该方法需要事先知道噪声的概率密度函数.

本文首先考察了噪声对计算最大 Lyapunov 指数的修正, 然后通过假设的噪声水平 σ , 定义了类似计算最大 Lyapunov 指数的二元函数 $f(d, \sigma^2)$. 理论分析发现, 可以根据 $f(d, \sigma^2)$ 随维数的变化趋势, 估计出混沌时间序列的噪声水平. 仿真结果验证了该算法的有效性.

2 理论分析

对于单变量混沌时间序列 $\{x(n) : n = 1, 2, \dots, N_s\}$, N_s 为采样点数目, 采样时间间隔为 Δt . 本文中, 所有时间序列均按下式归一化

$$x(n) \rightarrow \frac{x(n)}{\sigma_x}, \quad (1)$$

这里, σ_x 表示时间序列 $\{x(n)\}$ 的标准差. 含有噪声的测量信号 $y(n)$ 为

$$y(n) = x(n) + e(n), \quad (2)$$

这里, $e(n)$ 为随机噪声. 因此, 如果噪声的标准差 $\sigma_e = 0.05$, 那么噪声水平为 5%.

* 国家重点基础研究发展计划 (批准号: 2010CB227005)、国家自然科学基金 (批准号: 20906020) 和教育部新世纪优秀人才计划 (批准号: NCE-08-0775) 资助的课题.

[†] E-mail: hfliu@ecust.edu.cn

根据 Takens 时间延迟嵌入定理 [22], 将 $x(n)$ 进行相空间重构, 得:

$$\mathbf{x}(n) = [x(n), x(n + \tau), \dots, x(n + (d - 1)\tau)], \quad (3)$$

这里, τ 和 d 分别为时间延迟和嵌入维数, $\{\mathbf{x}(n) : n = 1, 2, \dots, N\}$ 为重构吸引子, 其中 $N = N_s - (d - 1)\tau$. 同理可得 $\mathbf{y}(n) = [y(n), y(n + \tau), \dots, y(n + (d - 1)\tau)]$, $\mathbf{e}(n) = [e(n), e(n + \tau), \dots, e(n + (d - 1)\tau)]$.

对于无噪相空间 $\{\mathbf{x}(n)\}$ 中的任意一点 $\mathbf{x}(n)$, $\mathbf{x}_{r_i}(n)$ 和 $\mathbf{x}_{s_i}(n)$ 分别为 $\mathbf{x}(n)$ 的第 r_i 和 s_i 个邻域, 其中 $r_i \neq s_i$, 那么 $\mathbf{x}_{s_i}(n)$ 与 $\mathbf{x}_{r_i}(n)$ 也互为邻域, $\mathbf{x}(n)$, $\mathbf{x}_{r_i}(n)$ 和 $\mathbf{x}_{s_i}(n)$ 经时间 $t = T\Delta t$ 分别演化得到 $\mathbf{x}(n, T)$, $\mathbf{x}_{r_i}(n, T)$ 和 $\mathbf{x}_{s_i}(n, T)$. 由于相邻轨迹的距离按指数规律增长 [12,23,24], 则

$$\begin{aligned} & |\mathbf{x}_{r_i}(n, T) - \mathbf{x}_{s_i}(n, T)| \\ &= |\mathbf{x}_{r_i}(n) - \mathbf{x}_{s_i}(n)| e^{\lambda_1 T \Delta t}, \end{aligned} \quad (4)$$

这里, $|\cdot|$ 表示 Euler 距离, λ_1 为最大 Lyapunov 指数.

由于局部相空间中相邻轨迹的距离的演化规律相似 [24,25], 则由 (4) 式可得

$$\ln h(T) = 2\lambda_1 T \Delta t + \ln h(0), \quad (5)$$

这里, $h(T) = \langle |\mathbf{x}_{r_i}(n, T) - \mathbf{x}_{s_i}(n, T)|^2 \rangle$, $h(0) = \langle |\mathbf{x}_{r_i}(n) - \mathbf{x}_{s_i}(n)|^2 \rangle$, 其中, $\langle \dots \rangle$ 表示对不同的 i 值取平均, 即,

$$\begin{aligned} & \langle |\mathbf{x}_{r_i}(n, T) - \mathbf{x}_{s_i}(n, T)|^2 \rangle \\ &= \frac{1}{N_p} \sum_{i=1}^{N_p} |\mathbf{x}_{r_i}(n, T) - \mathbf{x}_{s_i}(n, T)|^2, \\ & \langle |\mathbf{x}_{r_i}(n) - \mathbf{x}_{s_i}(n)|^2 \rangle \\ &= \frac{1}{N_p} \sum_{i=1}^{N_p} |\mathbf{x}_{r_i}(n) - \mathbf{x}_{s_i}(n)|^2, \end{aligned}$$

N_p 为所有邻域点对的数目. 本文中, 对任意的 $i \neq j$, 有 $r_i \neq r_j$, $s_i \neq s_j$ 以及 $r_i \neq s_j$.

对于含有噪声的测量信号, 由于 $\mathbf{y}_{r_i}(n) = \mathbf{x}_{r_i}(n) + \mathbf{e}_{r_i}(n)$, 则

$$\begin{aligned} & \langle |\mathbf{y}_{r_i}(n, T) - \mathbf{y}_{s_i}(n, T)|^2 \rangle \\ &= A + B - C - D + E + F - G + H, \end{aligned} \quad (6)$$

其中,

$$A = \sum_{l=0}^{d-1} \langle (x_{r_i}(n + l\tau, T) - x_{s_i}(n + l\tau, T))^2 \rangle,$$

$$\begin{aligned} B &= \sum_{l=0}^{d-1} \langle 2x_{r_i}(n + l\tau, T)e_{r_i}(n + l\tau, T) \rangle, \\ C &= \sum_{l=0}^{d-1} \langle 2x_{r_i}(n + l\tau, T)e_{s_i}(n + l\tau, T) \rangle, \\ D &= \sum_{l=0}^{d-1} \langle 2x_{s_i}(n + l\tau, T)e_{r_i}(n + l\tau, T) \rangle, \\ E &= \sum_{l=0}^{d-1} \langle 2x_{s_i}(n + l\tau, T)e_{s_i}(n + l\tau, T) \rangle, \\ F &= \sum_{l=0}^{d-1} e_{r_i}^2(n + l\tau, T), \\ G &= \sum_{l=0}^{d-1} \langle 2e_{r_i}(n + l\tau, T)e_{s_i}(n + l\tau, T) \rangle, \\ H &= \sum_{l=0}^{d-1} \langle e_{s_i}^2(n + l\tau, T) \rangle. \end{aligned}$$

对于独立同分布 (IID) 的测量噪声, 当 $N_p \rightarrow \infty$ 时, 有 $A = h(T)$, $B = C = D = E = G = 0$, $F = H = d\sigma_e^2$, 则

$$h_n(T) = h(T) + 2d\sigma_e^2, \quad (7)$$

这里, $h_n(T) = \langle |\mathbf{y}_{r_i}(n, T) - \mathbf{y}_{s_i}(n, T)|^2 \rangle$. 将 (7) 式代入 (5) 式, 当 $T_1 \neq T_2$ 时, 有

$$\lambda_1 = \frac{\ln(h_n(T_2) - 2d\sigma_e^2) - \ln(h_n(T_1) - 2d\sigma_e^2)}{2\Delta t(T_2 - T_1)}. \quad (8)$$

定义二元函数

$$f(d, \sigma^2) = \frac{\ln(h_n(T_2) - 2d\sigma^2) - \ln(h_n(T_1) - 2d\sigma^2)}{2\Delta t(T_2 - T_1)}, \quad (9)$$

显然, 当 $\sigma^2 = \sigma_e^2$ 时, $f(d, \sigma^2) = \lambda_1$ 并且不依赖于嵌入维数 d . (7) 式代入 (9) 式后, 将 $f(d, \sigma^2)$ 对 d 求偏导, 得

$$\begin{aligned} \frac{\partial f(d, \sigma^2)}{\partial(d)} &= \frac{(\sigma^2 - \sigma_e^2)(h(T_2) - h(T_1))}{\Delta t(T_2 - T_1)(h(T_1) + 2d\sigma_e^2 - 2d\sigma^2)} \\ &\quad \times \frac{1}{(h(T_2) + 2d\sigma_e^2 - 2d\sigma^2)}. \end{aligned} \quad (10)$$

对于混沌系统, $\lambda_1 > 0$, 由 (5) 式得 $\frac{h(T_2) - h(T_1)}{T_2 - T_1} > 0$, 则有如下结论:

(i) 当 $\sigma^2 < \sigma_e^2$ 时, $\frac{\partial f(d, \sigma^2)}{\partial(d)} < 0$, 此时 $f(d, \sigma^2)$ 随着维数的增加而减小;

(ii) 当 $\sigma^2 = \sigma_e^2$ 时, $\frac{\partial f(d, \sigma^2)}{\partial(d)} = 0$, 此时 $f(d, \sigma^2)$ 不随维数的变化而变化;

(iii) 当 $\sigma_e^2 < \sigma^2 < \sigma_e^2 + \frac{\min(h(T_1), h(T_2))}{2d}$ 时, $\frac{\partial f(d, \sigma^2)}{\partial(d)} > 0$, 此时 $f(d, \sigma^2)$ 随着维数的增加而增大.

因此, 可以根据 $f(d, \sigma^2)$ 随维数的变化趋势, 估计出噪声水平. 大体上讲, 该方法经过以下几个步骤:

1) 将时间序列重构到 d 维相空间;

2) 对相空间中某一点 $\mathbf{y}(n)$, 搜索 ε -邻域内的所有邻点, 如果邻域点个数大于某个阈值 N_T (一般而言, 使得 $dN_T \approx 200$ 的 N_T 即可满足本文算法的条件), 那么将这些邻点分成两部分, 然后取一系列不同值的 σ , 并计算 $\ln(h_n(T) - 2d\sigma^2)$, $T = 0, 1, 2, \dots$, 否则, 放弃这一点, 接着计算下一点;

3) 取 N_c 个不同的可用的计算点, 并计算

$$S(T) = \frac{1}{N_c} \sum \ln(h_n(T) - 2d\sigma^2); \quad (11)$$

4) 选取 $S(T)$ 的线性较好的区间, 对该区间进行线性最小二乘拟合, 并将其斜率的一半作为 $f(d, \sigma^2)$ 的值;

5) 将时间序列重构到不同维数的相空间, 采用相同的方法计算不同维数 d 下的 $f(d, \sigma^2)$;

6) 对某个确定的 σ , 采用线性最小二乘拟合方法计算 $f(d, \sigma^2)$ 随嵌入维数 d 变化的斜率 $S_f(\sigma)$, 然后采用线性插值的方法, 取第一次使得 $S_f(\sigma) = 0$ 的 σ 作为噪声水平的估计值 σ_{est} .

3 仿真计算

3.1 计算条件

以 Hénon 映射^[26] 和广义 Hénon 映射^[27](超混沌) 为例, 时间序列长度分别为 $N_s = 50000$ 和 $N_s = 150000$, 加入的噪声水平从 0% 到 10%, 采用本文算法进行仿真计算. 表 1 列出了这两种混

沌系统的有关参数. 相空间重构时时间延迟均取 1, 若无特别说明, 加入的噪声均为 IID Gauss 噪声. 两混沌系统均以计算得到的 x 值作为仿真计算的时间序列.

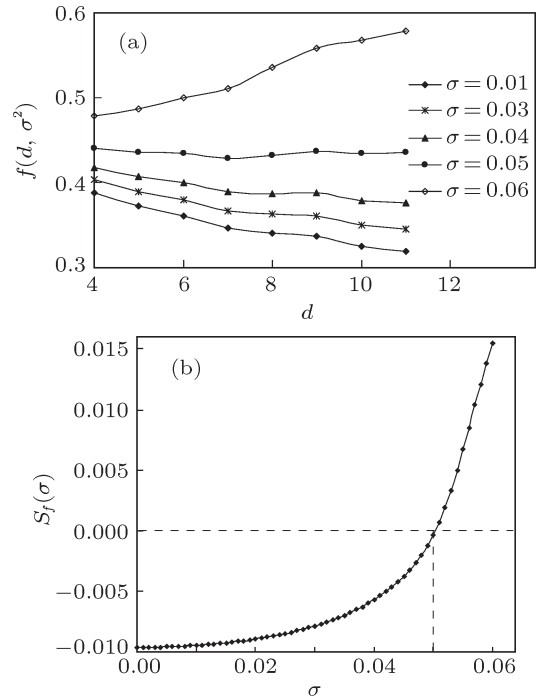


图 1 Hénon 映射时间序列计算结果 (a) $f(d, \sigma^2)$ 随 d 的变化曲线; (b) $S_f(\sigma)$ 随 σ 的变化曲线

3.2 从混沌时间序列估计噪声水平

对于 Hénon 映射时间序列, 加入 5% 的噪声, 采用本文算法计算得到的 $f(d, \sigma^2)$ 及 $S_f(\sigma)$ 如图 1 所示, 取第一次使得 $S_f(\sigma) = 0$ 的 σ , 可得 $\sigma_{est} = 0.05037$. 图 2 为广义 Hénon 映射时间序列加入 7% 的噪声的计算结果, 同理可得 $\sigma_{est} = 0.06962$. 更加详细的计算结果如图 3 所示. 可以看出, 当噪声水平不超过 7% 时, 两混沌系统的估计值与真实值符合得较好, 而当噪声水平达到 10% 时, 均出现约 15% 的负偏差.

表 1 混沌系统及其相关参数

混沌系统	方程	参数	λ_1
Hénon	$x_{i+1} = 1 - ax_i^2 + y_i$ $y_{i+1} = bx_i$	$a = 1.4, b = 0.3$	$0.4169^{[12]}$
广义 Hénon	$x_{i+1} = a - y_i^2 - bz_i$ $y_{i+1} = x_i$ $z_{i+1} = y_i$	$a = 1.76, b = 0.1$	$\lambda_1 = 0.225^{[24]}$, $\lambda_2 = 0.187^{[24]}$

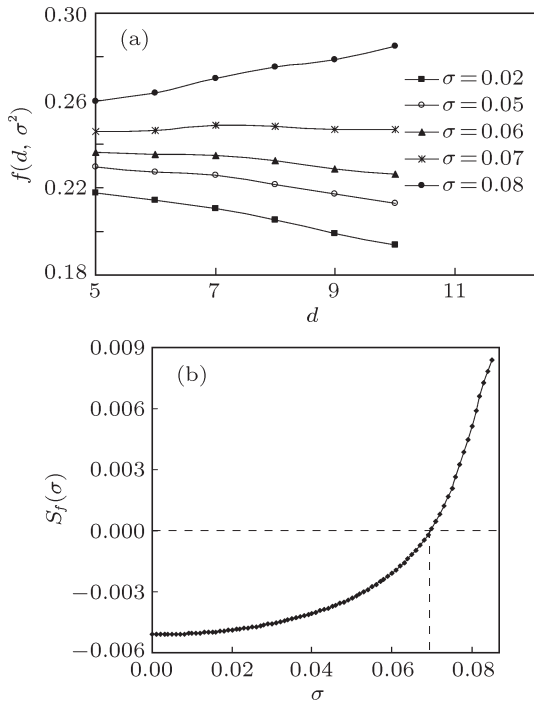


图2 广义 Hénon 映射时间序列计算结果 (a) $f(d, \sigma^2)$ 随 d 的变化曲线; (b) $S_f(\sigma)$ 随 σ 的变化曲线

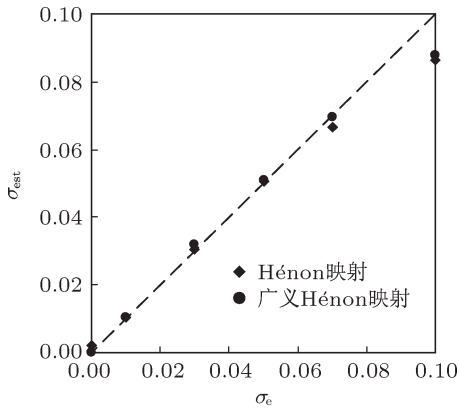


图3 不同混沌系统噪声水平估计结果

3.3 不同的噪声类型

前面的算例里, 加入的噪声均为 IID Gauss 噪

声, 但这一假设对实际的实验数据并不一定都成立. 第二部分的推导过程中, 并没有严格要求噪声为 IID Gauss 噪声.

分别采用三种不同类型的 IID 噪声考察本文的算法: Gauss 分布, 均匀分布, 以及二者各占 50% 的组合分布. 以 Hénon 映射为例, 加入的噪声水平从 0% 到 10%, 数值结果如图 4 所示. 可以看出, 不同噪声类型的计算结果一致. 这一性质对实际计算很重要, 因为实验时间序列数据的噪声类型事先往往并不知道.

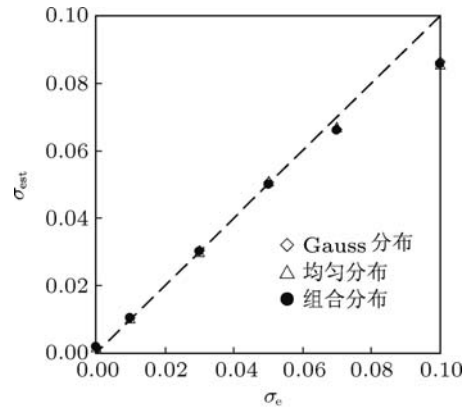


图4 Hénon 映射加入不同类型的噪声的计算结果

4 结论

实际测量的时间序列数据不可避免地含有各种噪声, 正确估计其中含有的噪声水平十分重要. 本文从理论上推导出了一种基于最大 Lyapunov 指数不变性的计算混沌时间序列噪声水平的新方法. 对 Hénon 映射和广义 Hénon 映射分别加入不同强度的噪声进行仿真计算, 结果表明, 当噪声水平不超过 7% 时, 估计值与真实值符合得很好; 当噪声水平达到 10% 时, 估计值出现约 15% 的负偏差. 本文算法对噪声分布类型不敏感, 因此是一种可以应用于实际时间序列数据的有效方法.

[1] Tulu S, Yilmaz O 2010 *Chaos* **20** 043103
 [2] Skufca J D, Yorke J A, Eckhardt B 2006 *Phys. Rev. Lett.* **96** 174101
 [3] Krasny R, Nitsche M 2002 *J. Fluid Mech.* **454** 47
 [4] Brackley C A, Ebenhoh O, Grebogi C, Kurths J, Moura A D, Romano M C, Thiel M 2010 *Chaos* **20** 045101
 [5] Ghosh-Dastidar S, Adeli H, Dadmehr N 2007 *IEEE Trans. Biomed. Eng.* **54** 1545
 [6] Chen W C 2008 *Chaos, Solitons and Fractals* **36** 1305
 [7] Schreiber T 1999 *Phys. Rep.* **308** 1
 [8] Wang H C, Chen G R, Lü J H 2004 *Phys. Lett. A* **333** 246
 [9] Liu H F, Zhao Y Y, Dai Z H, Gong X, Yu Z H 2001 *Acta Phys. Sin.* **50** 2311 (in Chinese) [刘海峰, 赵艳艳, 代正华, 龚欣, 于遵宏 2001 物理学报 **50** 2311]
 [10] Zhou Y D, Ma H, Lü W Y, Wang H Q 2007 *Acta Phys. Sin.* **56** 6809 (in Chinese) [周永道, 马洪, 吕王勇, 王会琦 2007 物理学

- 报 56 6809]
- [11] Zhang J F, Hu S S 2008 *Acta Phys. Sin.* **57** 2708 (in Chinese) [张军峰, 胡寿松 2008 物理学报 **57** 2708]
- [12] Liu H F, Dai Z H, Li W F, Gong X, Yu Z H 2005 *Phys. Lett. A* **341** 119
- [13] Gong Z Q, Feng G L, Dong W J, Li J P 2006 *Acta Phys. Sin.* **55** 3180 (in Chinese) [龚志强, 封国林, 董文杰, 李建平 2006 物理学报 **55** 3180]
- [14] Wu Y D, Xie H B 2007 *Acta Phys. Sin.* **56** 6294 (in Chinese) [吴延东, 谢洪波 2007 物理学报 **56** 6294]
- [15] Schreiber T 1993 *Phys. Rev. E* **48** 13
- [16] Yu D, Small M, Harrison R G, Diks C 2000 *Phys. Rev. E* **61** 3750
- [17] Jayawardena A W, Xu P, Li W K 2008 *Chaos* **18** 023115
- [18] Urbanowicz K, Holyst J A 2003 *Phys. Rev. E* **67** 046218
- [19] Urbanowicz K, Holyst J A 2006 *Int. J. Bifurcat. Chaos* **16** 1865
- [20] Strumik M, Macek W M, Redaelli S 2005 *Phys. Rev. E* **72** 036219
- [21] Moriya N 2010 *Nucl. Instrum. Methods Phys. Res. A* **618** 306
- [22] Takens F 1981 *Dynamical System and Turbulence, Lecture Notes in Mathematics* (Berlin: Springer-Verlag) p366
- [23] Rosenstein M T, Collins J J, De Luca C J 1993 *Physica D* **65** 117
- [24] Kantz H 1994 *Phys. Lett. A* **185** 77
- [25] Guegan D, Leroux J 2009 *Chaos Solitons and Fractals* **41** 2401
- [26] Hénon M 1976 *Commun. Math. Phys.* **50** 69
- [27] Baier G, Klein M 1990 *Phys. Lett. A* **151** 281

Noise-level estimation of noisy chaotic time series based on the invariant of the largest Lyapunov exponent*

Yao Tian-Liang Liu Hai-Feng[†] Xu Jian-Liang Li Wei-Feng

(Key Laboratory of Coal Gasification, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China)

(Received 14 May 2011; revised manuscript received 9 August 2011)

Abstract

A novel method of estimating the noise level from a noisy chaotic time series based on the invariant of the largest Lyapunov exponent is presented in this paper. The influence of noise on the distance between two points in an embedding phase space is considered, and then based on the invariant of the largest Lyapunov exponent in a different dimensional embedding phase space, the algorithm is proposed to estimate the noise level. Simulation results show that the estimated values of noise level agree well with the true values when the noise level is less than 10%. And this method is not sensitive to the distribution of noise. Therefore, the method is useful for estimating the noise level of noisy chaotic time series.

Keywords: chaos, time series, the largest Lyapunov exponent, noise

PACS: 05.40.Ca, 05.45.Tp, 05.45.Pq

* Project supported by the National Basic Research Program of China (Grant No. 2010CB227005), the National Natural Science Foundation of China (Grant No. 20906020), and the Program for New Century Excellent Talents in University of Ministry of Education of China (Grant No. NCET-08-0775).

[†] E-mail: hfiu@ecust.edu.cn