

# 基于泊松模型的 X 射线脉冲星信号的最大似然 TOA 估计\*

李建勋<sup>†</sup> 柯熙政

(西安理工大学自动化与信息工程学院 西安 710048)

**摘要** 讨论了 X 射线脉冲星光子到达的周期平稳 Poisson 模型和到达时刻 (Time of arrival, TOA) 估计问题. 在此基础上, 得到了脉冲到达时刻的最大似然估计和克拉美-罗限 (Cramer-Rao boundary, CRB), 并推导了低信噪比情况下的似然函数近似表达和克拉美-罗限. 利用解析的脉冲轮廓, 对 PSR B1821-24 的 TOA 估计进行了 Monte Carlo 仿真, 讨论了不同观测时间和信噪比下的估计误差, 给出了相应的信噪比门限. 研究表明: 该分析方法能够有效估计 X 射线脉冲星的 TOA 定时精度, 有利于评估其在各种应用中的性能.

**关键词** 脉冲星: 普通, 脉冲星: 个别, X 射线: 双星, 方法: 数据分析  
**中图分类号:** P 172; **文献标识码:** A

## 1 引言

脉冲星是快速自转的中子星, 辐射信号具有稳定的周期性, 尤其是毫秒脉冲星的长期稳定度可与原子钟相媲美, 有望建立一种新的时间尺度<sup>[1-5]</sup>. X 射线脉冲星主要分为 rotation-powered 和 accretion-powered 两种类型<sup>[6]</sup>, 由于其允许用较小的探测器进行光子探测, 在脉冲星导航方面也具有应用潜力<sup>[7]</sup>.

获得脉冲星脉冲的到达时刻信息是脉冲星各种研究和应用的基础<sup>[8-9]</sup> (射电脉冲星的定时观测如图 1 所示), 通过比较观测的 TOA 和定时模型预测的 TOA, 可以得到丰富的天文信息, 如脉冲星的自转周期、运动轨道、脉冲星质量、内部结构、引力波效应、星际介质特性等<sup>[10-11]</sup>. 经典的方法是把观测脉冲与理想的高信噪比脉冲模板进行相关, 并结合插值的方法进行估计. 但相关法需要假设噪声类型为白噪声, 且在信噪比低的情况下估计精度不高.

对于 X 射线脉冲星, 最基本的观测对象是由脉冲星释放出的单个光子. 首先, 探测器要能够探测源光子的入射信号. 其次, 一旦目标源的入射光子被检测到, 探测器的相应部件将以系统的时钟记录下光子的到达时刻. 第三, 观测时间需要转化为惯性参考系

2009-07-06 收到原稿, 2010-02-01 收到修改稿

\* 国家自然科学基金 (60977054)、陕西省自然科学基金 (2010JQ8002)、西安市科技计划项目 (CXY1012) 及校特色研究项目 (105-210911) 资助

<sup>†</sup> Ljx088@xaut.edu.cn

的等效时间,即在惯性系时间轴上进行重排.探测的光子流并不能直接代表脉冲轮廓,设检测的  $N$  个光子在时间跨度上包含了很多脉冲周期,每个光子的到达时刻对应其所属脉冲的特定相位值,必须把所有的光子按其所在单一脉冲中的相位值进行对齐后累加,才能生成脉冲轮廓.我们称这个整理过程为历元折叠,即按期望的脉冲周期(由定时模型预测)同步累加所有的到达光子数目(如图 2 所示).具体步骤是,把观测脉冲星的整个相位等分为  $M$  段(可称为时间仓),每一个光子事件按折叠周期依次放入相应的仓中,可想而知,一些收集仓将累加到更多的光子事件,最后在一个周期长度上生成代表脉冲轮廓的直方图,本质上相当于多个脉冲波形的相位平均.观测轮廓一旦创建,就可以获得脉冲的各种特征,包括有效的脉冲幅度、脉峰个数、形状和噪声特征等,其 TOA 确定一般与源的辐射强度、背景辐射以及脉冲形状有关 [12].

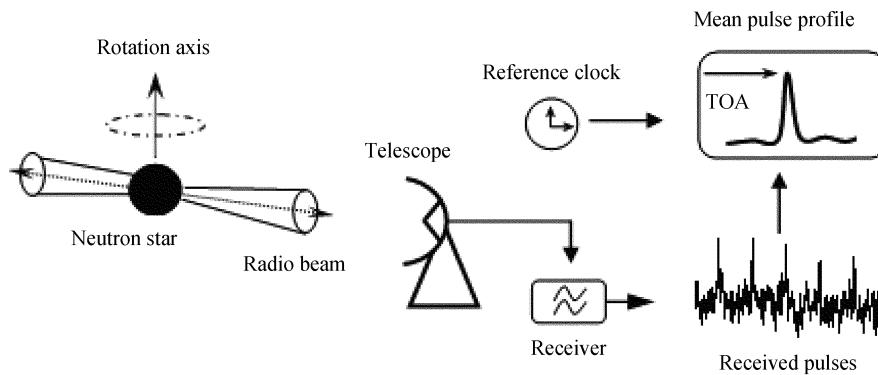


图 1 射电脉冲星定时观测的基本概念图

Fig. 1 Depiction of timing observation for radio pulsar

本文基于非齐次 Poisson 过程,直接从光子流到达规律来研究 X 射线脉冲星 TOA 的最大似然估计方法和估计性能,由于脉冲信号的周期平稳性, TOA 估计问题也等价于脉冲相位的估计问题.

## 2 信号模型

若不考虑脉冲星周期变率的影响, X 射线脉冲星的辐射特征可以用周期平稳过程的概率密度来表征.探测器接收信号的特征可用 Poisson 点过程描述.因此,可以用 Poisson 过程的循环平稳特性来建模 X 射线光子流的到达规律,进行脉冲到达时刻(或相位)的估计.

平稳的 Poisson 计数过程由强度  $\lambda$  刻画,在时间间隔  $T_b$  内,  $k$  个事件发生的概率为

$$Pr(k) = \frac{(\lambda T_b)^k}{k!} e^{-\lambda T_b}. \quad (1)$$

设探测器的时间分辨率为  $T_b$ ,指能检测到光子到达的最小时间间隔(时间仓).在该间隔内,不管有多少个光子到达,都被检测为单光子事件.则由(1)式,在任一个给定的时

间仓  $T_b$  内, 探测器没有检测到光子到达的概率为  $P_0 = e^{-\lambda T_b}$ , 检测到光子到达的概率为  $P_1 = 1 - e^{-\lambda T_b}$ .

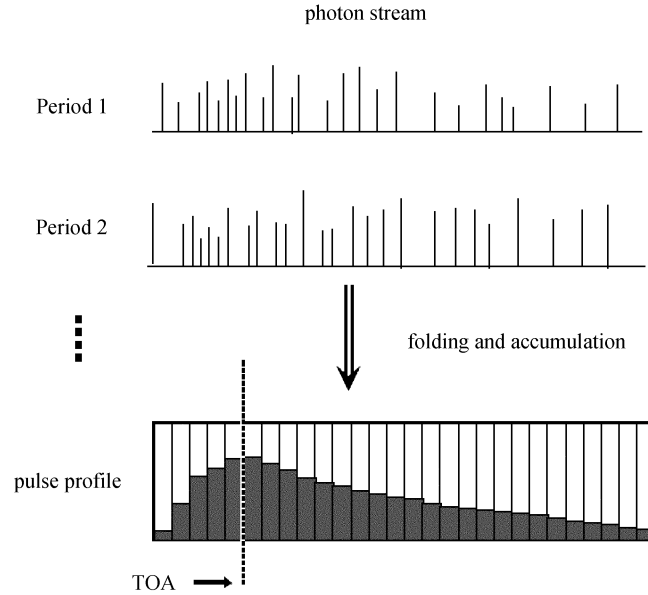


图 2 X 射线脉冲星的光子流周期累加

Fig. 2 Period accumulation of photon stream from X-ray pulsar

对于 X 射线脉冲星的光子流来说, 其到达密度具有非齐次性和周期平稳性<sup>[13]</sup>(统计特征周期变化), 可以记为

$$\lambda(t) = \lambda_p(t) + \lambda_n = \lambda(t + P). \quad (2)$$

其中,  $P$  是脉冲星的循环周期,  $\lambda_p(t)$  为脉冲星单光子到达密度, 具有周期性,  $\lambda_n$  是平稳背景噪声的到达密度.  $\lambda(t)$  的带宽在  $1/(2T_b)$  内. 接收脉冲的信噪比可定义为  $R_{SN} = \lambda_p/\lambda_n$ ,  $\lambda_p$  为脉冲星信号在一个周期中的平均到达密度.

设探测器一个周期的时间仓个数为整数  $N_b = P/T_b$ , 则任意一个仓 (第  $n_1$  个周期的第  $n_2$  个时间间隔) 的记录时刻可用下式表示:

$$t_{n_1, n_2} = n_1 P + n_2 T_b = (n_1 N_b + n_2) T_b. \quad (3)$$

在该时间仓上定义一个指示函数  $I(n_1, n_2)$  表示是否有光子到达

$$I(n_1, n_2) = \begin{cases} 1 & \text{detected} \\ 0 & \text{undetected} \end{cases}, \quad (4)$$

则有

$$P_0(t_{n_1, n_2}) = Pr\{I(n_1, n_2) = 0\} = \exp(-T_b \lambda(t_{n_1, n_2})), \quad (5)$$

$$P_1(t_{n_1, n_2}) = 1 - P_0(t_{n_1, n_2}). \quad (6)$$

### 3 TOA 的最大似然估计

TOA 参数估计的最大似然准则为, 对于一组样本数据, 参数的估计值应使下面的联合概率达到最大

$$Pr(\tau_{\text{bin}}; I) = \prod_{n_1=0}^{N_p-1} \prod_{n_2=0}^{N_b-1} P_1(t_{n_1, n_2} - \tau_{\text{bin}})^{I(n_1, n_2)} P_0(t_{n_1, n_2} - \tau_{\text{bin}})^{(1-I(n_1, n_2))}. \quad (7)$$

其中,  $N_p$  为脉冲星观测的周期数.  $\tau_{\text{bin}}$  指以时间仓为单位的时间延迟. 与 (7) 式等价的对数形式为

$$\lg Pr(\tau_{\text{bin}}; I) = \sum_{n_1=0}^{N_p-1} \sum_{n_2=0}^{N_b-1} I(n_1, n_2) \lg P_1(t_{n_1, n_2} - \tau_{\text{bin}}) + (1 - I(n_1, n_2)) \lg P_0(t_{n_1, n_2} - \tau_{\text{bin}}). \quad (8)$$

由于光子到达密度具有周期平稳性, 上式按脉冲周期累加后得到

$$\lg Pr(\tau_{\text{bin}}; I) = \sum_{n_2=0}^{N_b-1} S_1(n_2) \lg P_1(t_{n_1, n_2} - \tau_{\text{bin}}) + S_0(n_2) \lg P_0(t_{n_1, n_2} - \tau_{\text{bin}}). \quad (9)$$

其中,  $S_1(n_2) = \sum_{n_1=0}^{N_p-1} I(n_1, n_2)$ , 为第  $S_1(n_2)$  个时间仓累加的光子到达数,  $S_0(n_2) = N_p - S_1(n_2)$ . 代入 (5) 式和 (6) 式, 有

$$\begin{aligned} \lg Pr(\tau_{\text{bin}}; I) &= C_1 + \sum_{n_2=0}^{N_b-1} S_1(n_2) \{ \lg(1 - \exp(-T_b \lambda(t_{n_1, n_2} - \tau_{\text{bin}}))) + T_b \lambda(t_{n_1, n_2} - \tau_{\text{bin}}) \} \\ &= C_1 + \sum_{n_2=0}^{N_b-1} S_1(n_2) \lg(\exp\{T_b \lambda(t_{n_1, n_2} - \tau_{\text{bin}})\} - 1), \end{aligned} \quad (10)$$

其中  $C_1 = -N_p T_b \sum_{n_2=0}^{N_b-1} \lambda(t_{n_1, n_2} - \tau_{\text{bin}})$ ,  $N_p$  足够大时,  $C_1$  并不依赖时延参数  $\tau_{\text{bin}}$ . 若定义

$$f(t) = \lg(e^{T_b \lambda(t)} - 1), \quad (11)$$

则似然函数最终为

$$L(\tilde{\tau}_{\text{bin}}; I) = \lg Pr(\tilde{\tau}_{\text{bin}}; I) = C_1 + \sum_{n_2=0}^{N_b-1} S_1(n_2) f(t_{n_1, n_2} - \tilde{\tau}_{\text{bin}}). \quad (12)$$

其中,  $\tilde{\tau}_{\text{bin}}$  为时延  $\tau_{\text{bin}}$  的估计, 即最大似然估计量.

下面给出低信噪比下的近似估计方法. 设信噪比  $R_{\text{SN}}$  较弱, 即  $\lambda_n$  比  $\lambda_p(t)$  大几个量级, 可把函数  $f(t)$  在  $\frac{\lambda_p(t)}{\lambda_n} = 0$  处进行泰勒级数展开

$$f(t) = \lg(e^{T_b \lambda(t)} - 1) \approx C_2 + \frac{T_b}{1 - \exp(-T_b \lambda_n)} \lambda(t), \quad (13)$$

其中  $C_2 = \lg(e^{T_b \lambda_n} - 1)$  与时间参数无关, 把 (13) 式代入对数似然函数 (12) 式, 有

$$L(\tilde{\tau}_{\text{bin}}; I) = C_3 + \frac{N_p T_b}{1 - \exp(-T_b \lambda_n)} \sum_{n_2=0}^{N_b-1} \hat{S}_1(n_2) \lambda(t_{n_1, n_2} - \tilde{\tau}_{\text{bin}}), \quad (14)$$

其中  $\hat{S}_1(n_2) = \frac{1}{N_p} S_1(n_2) = \frac{1}{N_p} \sum_{n_1=0}^{N_p-1} I(n_1, n_2)$ , 为每个时间仓平均的光子到达数, 服从期望为  $P_1(t_{n_1, n_2} - \tau_{\text{bin}})$ 、方差为  $P_1(t_{n_1, n_2} - \tau_{\text{bin}})P_0(t_{n_1, n_2} - \tau_{\text{bin}})/N_p$  的二项分布. (14) 式可以看作是脉冲星到达密度函数的匹配滤波输出. 所求的最大似然估计量就是使下面的代价函数达到最大

$$C(\tilde{\tau}_{\text{bin}}) = \sum_{n_2=0}^{N_b-1} \hat{S}_1(n_2) \lambda(t_{n_1, n_2} - \tilde{\tau}_{\text{bin}}). \quad (15)$$

## 4 克拉美 - 罗限

对于似然函数, 若满足

$$E\left\{\frac{d}{d\tau} \lg Pr(\tau; I)\right\} = 0, \quad (16)$$

则对参数  $\tau$  的任何无偏估计  $\hat{\tau}$ , 有<sup>[14]</sup>

$$Var(\hat{\tau}) \geq CRB(\tau) \equiv \frac{1}{E\left\{\frac{d^2}{d\tau^2} \lg Pr(\tau; I)\right\}}. \quad (17)$$

其中, 符号  $E$  代表求期望,  $Var$  表示  $\hat{\tau}$  的方差,  $CRB$  代表其理论上的克拉美 - 罗限. 实际情况中, 由于克拉美 - 罗限的推导过程中包含复杂的积分和数学期望计算, 一般采用修正形式 (即有关均方误差的确切界限值) 来反映估计的性能<sup>[15]</sup>

$$Var(\hat{\tau}) \geq CRB(\tau) \equiv \frac{1}{E\left\{\left(\frac{d}{d\tau} \lg Pr(\tau; I)\right)^2\right\}}, \quad (18)$$

其中

$$\begin{aligned} CRB^{-1}(\tau) &= E\left\{\left(\frac{d}{d\tau} \lg Pr(\tau; I)\right)^2\right\} \\ &= N_p^2 \sum_{n_2=0}^{N_b-1} E\left(\hat{S}_1(n_2) - P_1(t_{n_1, n_2} - \tau)\right)^2 f'^2(t_{n_1, n_2} - \tau) \\ &= N_p T_b^2 \sum_{n_2=0}^{N_b-1} \frac{\lambda^2(t_{n_1, n_2} - \tau)}{\exp(T_b \lambda(t_{n_1, n_2})) - 1}. \end{aligned} \quad (19)$$

在低信噪比假设下, 上式有

$$CRB^{-1}(\tau) \approx \frac{N_p T_b}{\exp(T_b \lambda_n) - 1} \int_0^P \lambda^2(t) dt \leq \frac{N_p}{\lambda_n} \int_0^P \lambda^2(t) dt. \quad (20)$$

(20) 式的上限是在  $T_b \rightarrow 0$  的情况下, 由于  $\frac{T_b}{\exp(T_b \lambda_n) - 1}$  为单调减函数且趋于  $\frac{1}{\lambda_n}$  而得到的, 则估计方差为

$$\text{Var}(\hat{\tau}) \geq \text{CRB}(\tau) \geq \frac{\lambda_n}{N_p \int_0^P \lambda'^2(t) dt}. \quad (21)$$

因此, 为了在低信噪比下得到最优的估计性能, 应该保证  $T_b \lambda_n \ll 1$ . 否则, 背景噪声将主导探测器输出, 使得估计性能急剧下降.

## 5 仿真

选择 PSR B1821-24 进行 Monte Carlo 仿真, 其中仿真参数分别为光子流密度  $\lambda_p = 5.8 \times 10^{-3}$  ph/s, 脉冲星周期  $P = 3.054 \times 10^{-3}$  s, 时间分辨率  $T_b = 10^{-6}$ . 可以看出, PSR B1821-24 的光子流量非常小, 信号微弱.

仿真中, 把观测时间等分为间隔足够小的时间段 (对应于 X 射线探测器的时间分辨率), 每段的光子到达个数按照 Poisson 分布 ((8) 式) 生成随机变量的仿真值, 然后通过周期折叠生成观测脉冲. 由最小二乘法拟合的归一化脉冲轮廓 (理论上的标准轮廓, 幅度与  $\lambda_p(t)$  成正比) 如图 3 所示. 在给定的  $R_{\text{SN}}$  和  $N_p$  下, 做 1 000 次实验 (设观测脉冲相对于理论脉冲的时延  $\tau_{\text{bin}}$  为 0.001 5 s), 其中某次实验的脉冲轮廓如图 4, 1 000 次实验得到的所有波形的平均脉冲轮廓如图 5 所示. 每次实验利用最大似然估计的搜索算法求解 TOA, 然后通过统计平均计算均方根误差. 其中, 单次实验的估计方差采取对脉冲周期归一化形式 (相位估计误差), 由下式计算:

$$\sigma_i^2 = \begin{cases} (|\tau - \hat{\tau}|/P - 1)^2 & |\tau - \hat{\tau}|/P > 0.5 \\ (|\tau - \hat{\tau}|/P)^2 & |\tau - \hat{\tau}|/P \leq 0.5 \end{cases}. \quad (22)$$

另外, 估计方差的理论值可由其 CRB 下限即 (21) 式的数值积分获得.

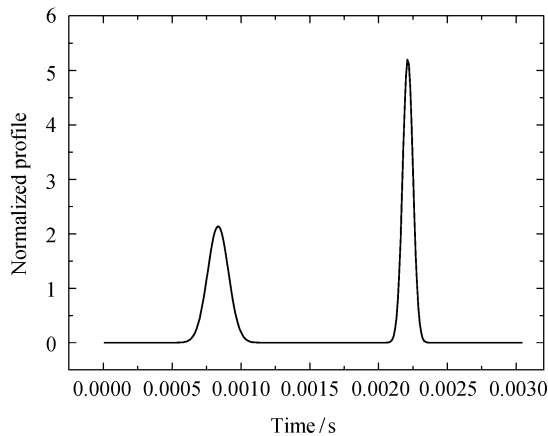


图 3 PSR B1821-24 的解析脉冲轮廓  
Fig. 3 Analytic pulse profile of PSR B1821-24

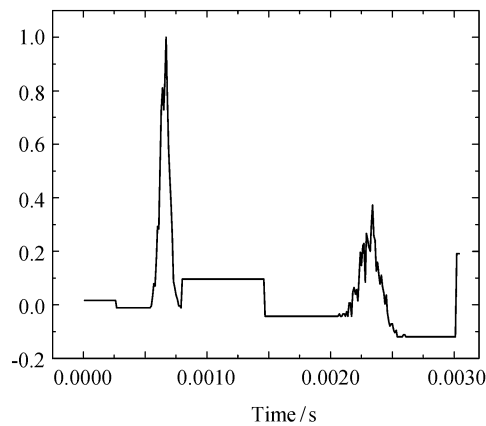


图 4 某次实验产生的脉冲轮廓 (时延为 0.001 5 s)  
Fig. 4 Pulse profile from one simulation experiment ( $\tau_{\text{bin}}=0.0015$  s)

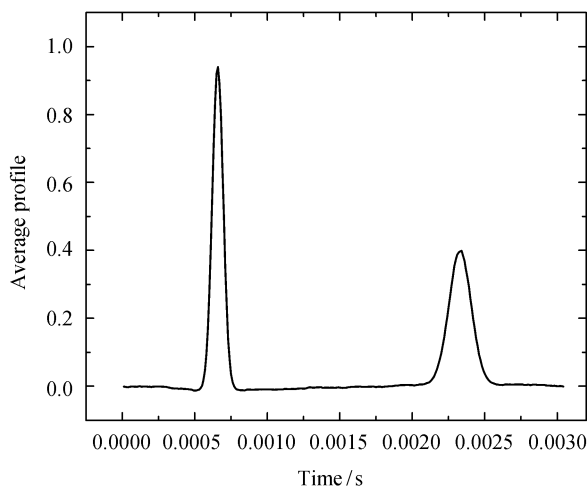


图 5 某 1 000 次实验产生的平均脉冲轮廓

Fig. 5 Average pulse profile from 1000 simulation experiments

令信噪比从 -100 dB 到 -40 dB 变化, 设观测周期  $N_p$  分别为  $10^6$  和  $10^8$ , 仿真得到的估计方差以及相应的 CRB 如图 6 所示. 从图 6 可以看出, 相对于 CRB, 最大似然估计的定时误差表现为 3 种特征 (对应图中的 3 个区域): (1) 信噪比相对较低时偏离 CRB 的区域; (2) 以 CRB 为有效和紧密下界的平稳区域; (3) 信噪比相对较高时的饱和区域. 偏离和吻合区域的边界即为最大似然算法的信噪比门限. 在  $N_p = 10^8$  时, 信噪比门限约为 -90 dB;  $N_p = 10^6$  时, 门限约为 -77 dB. 门限效应是由于似然函数最大化中假设噪声处于过零点的邻域, 其导数近似线性; 若噪声足够高, 接收数据将表现为非线性. 另外, CRB 在相对高的信噪比时到达性能底线, 这是由于脉冲信号内在的噪声特性决定的, 通过增加观测周期可以减小该效应. 同时, 通过仿真发现, 更低的信噪比将使偏离区域的定时方差趋于饱和, 而且不管选择多大的观测周期, 这个初始的误差值 (饱和值) 是不变的 (对本仿真的 PSR B1821-24 来说, 值约为 0.28), 这应该是由 X 射线脉冲信号  $\lambda_p(t)$  的周期平稳性决定的. 总之, 低于门限的定时误差不再是高斯和无偏的, 而与 CRB 吻合的平稳区域是似然函数适用的有效区域, 决定了采用似然算法的接收机工作范围.

可见在吻合区域, 最大似然方法在一定信噪比条件下能够有效地逼近其性能限, 表明了该估计的有效性. 一旦到达时刻以较高的精度估计出来后, 对计算、测量与其有关的其它信息具有重要意义, 如我们可以简单估计脉冲星导航的定位精度  $\sigma_p = c \cdot \sigma_t$ , 其中  $c$  为光速,  $\sigma_t$  为到达时间估计精度; 也可以对不同脉冲星的定时品质进行排序, 为工程中选择高性能的脉冲星提供支持.

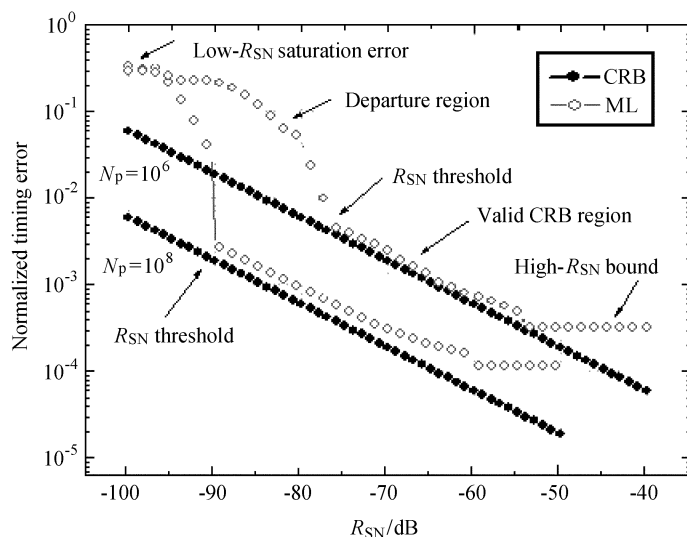


图 6 PSR B1821-24 的最大似然估计精度

Fig. 6 Timing error performance of maximum-likelihood (ML) estimate for PSR B1821-24

## 6 结论

精确的脉冲 TOA 估计对于 X 射线脉冲星的定时和其它应用非常重要. 本文对光子到达模型采用非齐次 Poisson 过程以及循环周期平稳特性进行建模, 给出了到达时刻的最大似然估计法. 在不同的信噪比即脉冲星辐射和背景辐射流量不同的情况下, 对脉冲星进行了脉冲轮廓生成以及相位估计的仿真. 在弱信噪比 (探测器面积一定) 情况下, 必须增加观测时间才能克服信噪比门限效应, 接近 CRB 的理论预测性能限. 在巡天观测阶段, 该方法可以用于对已知脉冲轮廓的脉冲星进行有效的定时性能评估, 有助于进一步建立 X 射线脉冲星的基本信息数据库, 根据不同应用需求筛选脉冲星.

## 参 考 文 献

- [1] 仲崇霞, 杨廷高. 天文学报, 2007, 48: 228
- [2] Zhong C X, Yang T G. ChA&A, 2007, 31: 443
- [3] 柯熙政, 李孝辉, 杨廷高, 等. 脉冲星观测与研究. 乌鲁木齐: 新疆人民出版社, 2000: 67-72
- [4] 仲崇霞, 杨廷高. 天文学报, 2009, 50: 425
- [5] Rodin A E. ChJAS, 2006, 6: 157
- [6] Charles P A, Seward F D. Exploring the X-ray Universe. Cambridge: Cambridge University Press, 1995: 155-190
- [7] Li J X, Ke X Z. ScChG, 2009, 52: 303

- [8] 李建勋, 柯熙政. 天文学报, 2008, 49: 394
- [9] Li J X, Ke X Z. ChA&A, 2009, 33: 158
- [10] Stairs I H. JPhG, 2006, 32: S259
- [11] Taylor J H. CQGra. 1993, 10: S167
- [12] Josep S, Andreu U, Xavier V, et al. Feasibility Study for a Spacecraft Navigation System Relying on Pulsar Timing Information, ESA: ARIADNA Study 03/4202, 2004: 50-57
- [13] Golshan A R, Sheikh S I. Proc. ION 63rd Annual Meeting, Cambridge, MA, April 23-25, 2007
- [14] Steven M K. Fundamentals of Statistical Signal Processing: Estimation Theory. NJ: Prentice Hall PTR, 1993: 219-226
- [15] D'Andrea A N, Mengali U, Reggiannini R. ITCOM, 1994, 42: 1391

## The Maximum-likelihood Estimate of TOA Accuracy for X-ray Pulsar Signal Based on Poisson Process

LI Jian-xun    KE Xi-zheng

(College of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048)

**ABSTRACT** The cycle-stationary Poisson model of photon arrival of X-ray pulsar and pulse time of arrival (TOA) are discussed. Based on this model, the maximum-likelihood estimate (MLE) of TOA is presented as well as the Cramer-Rao boundary (CRB), especially for the low- $R_{SN}$  (signal to noise ratio) cases. Using its analytic pulse profile, the Monte Carlo simulation of MLE for PSR B1821-24 is carried out and the timing error performance is analyzed and compared with the theoretical CRB under different observational durations and different  $R_{SN}$ . The  $R_{SN}$  threshold effect is observed and its significance is discussed. Results show that this method is an effective analysis tool to calculate the TOA precision of X-ray pulsar and would help to evaluate the corresponding performance in other applications.

**Key words** pulsars: general, pulsars: individual, X-rays: binaries, methods: data analysis