

Appendix S1. Updating of latent encounter history frequencies in **multimark**

Brett T. McClintock¹

National Marine Mammal Laboratory
Alaska Fisheries Science Center
NOAA National Marine Fisheries Service
Seattle, Washington, U.S.A.
¹*Email:* brett.mcclintock@noaa.gov

March 27, 2015

1 The feasible set of latent encounter histories (\mathbf{Y}) is explored in **multimark** using
2 an extension of the MCMC algorithms proposed by Bonner & Holmberg (2013)
3 and McClintock *et al.* (2013, 2014) that were originally conceived for a different
4 application by Link *et al.* (2010). The new algorithm conditions on the observed
5 data (thus reducing the dimension of the problem) and only proposes updates with
6 non-negative latent encounter history frequencies, $\mathbf{f} = (f_1, f_2, \dots, f_{5^T})$. Let \mathbf{r} denote
7 the set of $4^T - 2^{T+1} + 1$ indices for latent encounter histories that spawn >1 observed
8 history, and let $f_{j(1)}$ and $f_{j(2)}$ denote the corresponding frequencies for type 1 and
9 type 2 histories that arise from latent encounter history $j \in \mathbf{r}$. Referring back to
10 Table 2 with $T = 2$, $\mathbf{r} = \{4, 8, 9, 12, 14, 16, 17, 18, 19\}$ and one potential update would
11 involve frequencies for latent history ‘31’ (f_{17}) and its progeny ‘11’ ($f_{17(1)} = f_7$) and
12 ‘20’ ($f_{17(2)} = f_{11}$).

13 When conditioning on the observed encounter histories, the size of the prob-
 14 lem is typically greatly reduced because many of the potential latent histories and
 15 corresponding moves are not permissible. For example, with $T = 2$, if encounter
 16 history ‘11’ was never observed, then $f_7 = f_9 = f_{17} = f_{19} = 0$ can be ignored and
 17 $j \in \{9, 17, 19\}$ can be removed from \mathbf{r} for subsequent computations.

18 Starting from a permissible \mathbf{f} conditional on the observed encounter histories, the
 19 algorithm proceeds as follows:

- 20 1. Randomly draw a latent encounter history index $r \in \mathbf{r}^*$, where \mathbf{r}^* is the subset
 21 of \mathbf{r} with corresponding frequencies that satisfy $\min(f_1, f_j) + \min(f_{j(1)}, f_{j(2)}) > 0$
 22 for $j \in \mathbf{r}$.
- 23 2. Randomly draw c_r from the integer set $\{-\min(f_1, f_r), \dots, -1, 1, \dots, \min(f_{r(1)}, f_{r(2)})\}$.
- 24 3. Propose $f_r^* = f_r + c_r$, $f_{r(1)}^* = f_{r(1)} - c_r$, and $f_{r(2)}^* = f_{r(2)} - c_r$.
- 25 4. Apportion \mathbf{f}^* to individuals following McClintock *et al.* (2014), and accept
 26 proposed move based on the Metropolis-Hastings ratio described therein [pp.
 27 2470-2472, steps 9(b)-9(c)].

28 Note that because **multimark** uses “semi-complete” data likelihoods that con-
 29 dition on the number of unique individuals encountered at least once (n), the di-
 30 mension of the data-augmented encounter histories (M) described in McClintock
 31 *et al.* (2014) is determined by the number of observed encounter histories (i.e.,
 32 $M = n_1 + n_2 + n_{known}$), such that $f_1 = M - n$ and $n = \sum_{j=2}^{5^T} f_j$. Any addi-
 33 tional constraints, such as those resulting from encounter histories being designated

34 as known with certainty using the *known* argument in *processdata()*, are accounted
35 for by simple modifications to steps 1-2.

36 In terms of mixing, it can sometimes be advantageous to explore more than one
37 move at a time. At each iteration of the chain, the argument *maxnumbasis* specifies
38 how many times to perform steps 1-3 in sequence before evaluating step 4. The
39 default for *multimarkCJS()* and *multimarkClosed()* is *maxnumbasis=1*.

40 References

41 Bonner, S.J. & Holmberg, J. (2013) Mark-recapture with multiple, non-invasive
42 marks. *Biometrics*, **69**, 766–775.

43 Link, W.A., Yoshizaki, J., Bailey, L.L. & Pollock, K.H. (2010) Uncovering a latent
44 multinomial: analysis of mark-recapture data with misidentification. *Biometrics*,
45 **66**, 178–185.

46 McClintock, B.T., Bailey, L.L., Dreher, B.P. & Link, W.A. (2014) Probit models for
47 capture-recapture data subject to imperfect detection, individual heterogeneity
48 and misidentification. *The Annals of Applied Statistics*, **8**, 2461–2484.

49 McClintock, B.T., Conn, P.B., Alonso, R.S. & Crooks, K.R. (2013) Integrated mod-
50 eling of bilateral photo-identification data in mark-recapture analyses. *Ecology*,
51 **94**, 1464–1471.