Discussion on “Analysis of forensic DNA mixtures with artefacts” by Cowell, Graversen, Lauritzen and Mor
tera.

Alessio Farcomeni, Maria Francesca Marino, Monia Ranalli
(Sapienza - University of Rome)

First, we would like to congratulate the authors on a very stimulating paper. We have been provided with a method for modelling many subtleties when dealing with forensic DNA mixtures. Many artefacts and random variability due to data collection and the technology involved are directly taken into account. The inferential procedure is flexible and efficient, and the model is definitely parsimonious.

The authors focus mostly on \( n - 4 \) stutter. We believe different types of stutter (e.g., \( n + 4 \) stutters, where a fake peak appears after the real one) might be incorporated by increasing the number of components in the proposed mixture model. Each additional component involves estimation of one single additional parameter, so the formulation would still be parsimonious.

We wish to stimulate the authors in providing additional developments.

No human population can be safely assumed to be in Hardy-Weinberg equilibrium. We wonder about the implications of this assumption on the estimates obtained under the model. There might for instance be bias for offenders whose mother and father come from different ethnicities. Furthermore, population frequencies should be periodically updated under HW disequilibrium.

There also may be uncertainty in attribution of the offender to a target population. We wonder how much the estimated WoE is sensitive to this. If we believe an offender is Caucasian when he/she is not, how much bias can we incur in the WoE? If we are uncertain about the target population, can we perhaps average the WoE associated with different ones?

Similarly, there might be isolated sub-populations. In Italy for instance
it is acknowledged that Sardinians have a slightly different genetic profile. Similar differences can be seen for Icelanders among Europeans, or for certain ethnic minorities (e.g., Hawaiians in U.S.A.). How much bias is there in assuming Italian allele frequencies for a Sardinian offender? How can we deal with it?

Finally, some parameters could be made marker dependent to deal with variable fractions of DNA across markers. This might be useful when some markers are on the sex chromosomes.

From a more technical perspective, we wonder whether relaxing the assumption that, e.g., \(n_{i2}, \ldots, n_{i6}\), are independent would improve the goodness of fit substantially.

Furthermore, since the efficient estimation procedure proposed requires an approximation of the observed likelihood, we wonder how difficult it would be to work with the inverse Godambe matrix (see e.g. Huber, 1967; White, 1982) to obtain standard errors.

References
