Artif Intell Law (2013) 21:221–252
DOI 10.1007/s10506-012-9136-5

Modeling the forensic two-trace problem with Bayesian networks

Simone Gittelson · Alex Biedermann · Silvia Bozza · Franco Taroni

Published online: 12 December 2012

Abstract The forensic two-trace problem is a perplexing inference problem introduced by Evett (J Forensic Sci Soc 27:375–381, 1987). Different possible ways of wording the competing pair of propositions (i.e., one proposition advanced by the prosecution and one proposition advanced by the defence) led to different quantifications of the value of the evidence (Meester and Sjerps in Biometrics 59:727–732, 2003). Here, we re-examine this scenario with the aim of clarifying the interrelationships that exist between the different solutions, and in this way, produce a global vision of the problem. We propose to investigate the different expressions for evaluating the value of the evidence by using a graphical approach, i.e. Bayesian networks, to model the rationale behind each of the proposed solutions and the assumptions made on the unknown parameters in this problem.

Keywords Evaluation of evidence · Value of the evidence · Graphical probability models · Bayesian networks · Two-trace problem

1 Introduction

The two-trace problem, introduced by Evett (1987), is a perplexing inference problem that continues to puzzle many forensic scientists. It considers a scenario where forensic investigators recover two items of a particular category of trace evidence on a crime scene, e.g. two bloodstains, and compare both of these to the

S. Gittelson (✉) · A. Biedermann · F. Taroni
Institut de Police Scientifique, Ecole des Sciences Criminelles, Université de Lausanne,
Quartier Sorge: le Batochime, 1015 Lausanne, Switzerland
e-mail: simone.gittelson@unil.ch

S. Bozza
Dipartimento di Economia, Università Ca' Foscari di Venezia, San Giobbe, Cannaregio 873,
30121 Venezia, Italy

sample taken from a suspect. The question of interest to the court is, ‘How strong is the evidence resulting from these two comparisons in favor of the prosecution or the defence?’

The objective of the forensic scientist’s testimony is to answer this question. The answer to this question takes the form of the *value of the evidence* (e.g., Aitken and Taroni 2004):

$$V = \frac{Pr(\text{evidence}|\text{proposition } 1, I)}{Pr(\text{evidence}|\text{proposition } 2, I)}, \tag{1}$$

where proposition 1 is the proposition advanced by the prosecution, proposition 2 the proposition advanced by the defence, and I the background information consisting of the forensic scientist’s knowledge on the case circumstances prior to observing the evidence. The evidence is an intrinsic trait (e.g., the blood group or DNA profile) of the two traces and the suspect’s sample, observed as a result of the test or analysis performed in the forensic laboratory. Prior to hearing the forensic scientist’s testimony, the prosecution and the defence each take position on the origin of the traces. These views are formalized into the two propositions, that is, into two statements that are each either true or false. As a pair, these propositions must be mutually exclusive,[1] yet there is no requirement for them to be exhaustive[2] (e.g., Robertson and Vignaux 1995; Aitken and Taroni 2004). In this case, the first proposition (advanced by the prosecution) links the suspect to the crime stains, and the second (advanced by the defence) rejects such a link. The fact-finder (a judge or jury member) has a particular degree of belief in the truth of each of these propositions before hearing the forensic testimony. By presenting the value of the evidence V, the forensic scientist’s testimony conveys by how much more or less the evidence supports the first proposition with regard to the second proposition: if $V > 1$, the evidence supports the first proposition; if $V < 1$, the evidence supports the second proposition; and if $V = 1$, the evidence does not provide support for either of the two propositions, meaning that it is irrelevant for discriminating between them. Hence, the value of the evidence allows the fact-finder to update his or her belief in the truth of these propositions, and construct an informed opinion about each party’s account of the events.

1.1 Aim and outline of this paper

With two traces making up the recovered evidence, there are several possibilities for formulating a pair of propositions: they can focus on one of the two traces, or on both, and in the latter case, either specify or not specify which of the two traces originates (or does not originate) from the suspect. What is disturbing for a fact-finder hearing a forensic scientist’s testimony in the context of a two-trace problem, is that the value of the evidence, as given by Eq. (1), is *different* for different pairs of propositions (Meester and Sjerps 2003).

[1] Two propositions are *mutually exclusive* if they cannot both be true at the same time.

[2] A set of propositions is *exhaustive* if it covers all scenarios, so that at least one of its propositions is always true.

The aim of this paper is to investigate the value of the evidence in a two-trace problem with regard to different pairs of propositions, by unifying three different pairs in a single framework. To accomplish this, we will construct a Bayesian network, i.e., a graphical probability model. In forensic contexts, Bayesian networks help examine the reasonableness of the formal derivation of a formula, that is, the assumptions that have been made (Taroni et al. 2006). This allows us to compare the derived values of the evidence for different pairs of propositions. In addition, these models allow the user to perform complex probabilistic calculations that take into account the probability assignments over all of the unknown parameters. In this way, we hope to provide a global model which offers a complete and realistic approach to the valuation of scientific evidence in a two-trace problem. With this model, we hope to draw the reader's attention to the importance of the formulation of a pair of propositions, and increase his/her awareness of the impact that subtle differences in these formulations can have on the value of the evidence.

Besides a brief description of what Bayesian networks are and how they work (Sect. 2), we do not give a detailed explanation on Bayesian networks, and refer the interested reader to one of the many publications on the subject (e.g., Jensen 2001; Kjaerulff and Madsen 2008). Section 3 gives an overview of the two-trace problem as we will treat it in this paper, and Sect. 4 describes the notation we will use. Section 5 explains how we construct the Bayesian network, and Sects. 6 and 7 illustrate the use of this model and the influence of the different parameters through a numerical example. Concluding remarks are in Sect. 8.

2 Bayesian networks

Bayesian networks are graphical probability models, also known as probabilistic expert systems. The key advantage of these models is their capacity of splitting up a complex inference problem into its different parts. They represent random variables as nodes, and dependence relationships between the random variables as arrows connecting the nodes to form a directed acyclic graph. The random variables can be either discrete or continuous, but for the sake of simplicity we will use discrete nodes in this paper. Thus, each random variable will consist of a finite and exhaustive list of mutually exclusive states. The arrows model the probabilistic relationships between the variables by connecting a 'parent' node to a 'child' node. They condition the probability distribution of the child node upon each of its parents with probability tables that allow the user to quantify the probabilistic relationships.

In this way, the Bayesian network decomposes the joint probability distribution of a set of random variables $X_1, \ldots, X_n$ into the product of each of their probabilities conditioned on their parents. This is known as the Markov property:

$$Pr(X_1, \ldots, X_n) = \prod_{i=1}^{n} Pr(X_i | parents(X_i)). \tag{2}$$

It is important to stress that there is no true model, because a model is personal and reflects the constructor's view of the problem and the information available at the time of the construction (Lindley 2000). As our understanding of the issue progresses, the constructed network may evolve to model a situation more accurately, so that several different Bayesian network structures may be accepted as a description of the same scenario (Garbolino 2001).

The relevance of Bayesian networks for applications in forensic science was first recognized—in print—by Aitken and Gammerman (1989). Key publications in legal literature followed (Edwards 1991; Schum 1994; Kadane and Schum 1996), presenting thorough descriptions of the potential of probabilistic models for reasoning about evidence in real cases. Since then, the application of Bayesian networks has covered different aspects of evidential assessment (Taroni et al. 2006), in particular the evaluation of DNA evidence (Dawid et al. 2002, 2007; Mortera et al. 2003), evidence collected in fire debris (Biedermann et al. 2005a, b), firearm evidence (Biedermann and Taroni 2006; Biedermann et al. 2009), and fibre evidence (Garbolino and Taroni 2002), as well as practical considerations on how to present their results in court (Fenton and Neil 2011).

In this study, we constructed the Bayesian networks using the software *Hugin Researcher 7.3*, by Hugin Expert A/S. This program allows the user to construct and use Bayesian networks that contain numerical probability values. It can only carry out numerical propagations between the nodes. To derive the algebraic expressions corresponding to the calculations performed by the model, the user applies Eq. (2). The probabilistic relationships defined by the structure and the probability tables tell the user how the probability of a compound event is broken down into separate conditional probabilities.

3 The two-trace problem

We denote the following pair of propositions 'pair H':

Proposition 1: *At least one of the crime stains comes from the suspect*;

Proposition 2: *Neither of the crime stains comes from the suspect*.

These propositions are called source level propositions according to the hierarchy of propositions defined by Cook et al. (1998), because they describe whether a particular object or person is the source, or origin, of the traces recovered on the crime scene. Source level propositions are different from activity level propositions (describing the activity that led to the transfer of the traces from their source to the crime scene) and crime level propositions (concerned with whether the suspect actually committed the crime under investigation). In this paper we treat only source level propositions. For activity or crime level evaluations of the evidence in a two-trace problem, see Triggs and Buckleton (2003) and Gittelson et al. (2012), and Dawid (2004), respectively.

The value of the evidence with regard to the above propositions depends on the evidence observed. There are three possibilities:

1. If neither of the two crime stains matches the suspect's sample, then the likelihood of the first proposition is 0, and consequently Eq. (1) becomes

$$V = 0; \tag{3}$$

2. If one of the crime stains matches the suspect's sample, Evett (1987) showed that this leads to

$$V = \frac{1}{2\gamma}, \tag{4}$$

where γ represents the *match probability* (Weir 2000) of the matching trait in the relevant population of possible crime stain donors [Note that originally, Evett (1987) did not deduce this expression for the source level propositions as described, but for their equivalents at the crime level, assuming the relevance of both traces to be maximal. A crime level evaluation with maximal relevance produces the same value of the evidence as the source level evaluation presented here (see, e.g., Aitken and Taroni 2004)];

3. And if both of the crime stains match the suspect's sample, most forensic scientists would assume that the two traces come from a single contributor so that Eq. (1) reduces to

$$V = \frac{1}{\gamma}. \tag{5}$$

These assessments are based on the assumption that no laboratory errors are possible, an assumption we maintain throughout this paper. Note, however, that relaxing this assumption may have a considerable effect on the value of the evidence (Thompson et al. 2003).

Among these three ratios, Eqs. (3) and (5) are the same as for a scenario involving a single crime stain. This is because the differentiation between the two traces is not necessary in these cases in order to describe the observed evidence. In these two cases, one can combine the two crime stains into a single group, which we see as either matching (Eq. 5), or not matching (Eq. 3) the suspect's sample. In both of these cases, the reasoning that leads to Eqs. (3) and (5) is the same as that applied to the evaluation of the value of a single crime stain.

This is different for the case involving one matching stain and one non-matching stain (item 2 in the list). This case requires the forensic scientist to distinguish between the two traces by multiplying the traditional value of $\frac{1}{\gamma}$ by a factor of 0.5 (we will discuss the meaning of this additional factor in Sect. 7). This is the case which interests us in this paper.

Evett (1987) was not the only author to treat this problem. After Evett (1987), the case of one matching stain and one non-matching stain gave rise to the formulation of other propositions, which led to evidential values that were not equal to Eq. (4). According to Meester and Sjerps (2003), a pair of propositions worded slightly differently (note that the wording of these propositions has been modified here with regard to their original formulation in Meester and Sjerps (2003), yet their logical meaning remains unchanged), that is,

Proposition 1: *Crime stain 1 comes from the suspect*;

Proposition 2: *Neither of the crime stains comes from the suspect*;
(we denote this pair 'pair H_1'') produced a value of

$$\frac{1}{\gamma}, \tag{6}$$

and the pair

Proposition 1: *Crime stain 1 comes from the suspect*;

Proposition 2: *Crime stain 1 does not come from the suspect*;
(denoted 'pair H_1''') produced a value of[3]

$$\frac{2-\delta}{2\gamma(1-\delta)}, \tag{7}$$

for evidence consisting of a match between crime stain 1 and the suspect's sample, and a non-match between crime stain 2 and the suspect's sample. In Eq. (7), the probability denoted δ represents the prior probability that the suspect was one of two crime stain donors. This probability had to be introduced to correctly evaluate the probability of the evidence given proposition 2 and I. See Sect. 7.3 for further explanations.

The pair of propositions H (on page 4) is related to the above two pairs (pairs H_1' and H_1'') when the forensic scientist observes a match between the suspect's sample and stain 1, and a non-match between the suspect's sample and stain 2: in this case, all three pairs of propositions have identical posterior odds of

$$\frac{\delta}{2(1-\delta)\gamma}, \tag{8}$$

for a prior probability of δ that the suspect was a crime stain donor (Note that we call the odds of a pair of propositions 'prior odds' before observing the evidence, and 'posterior odds' after observing the evidence. We use the terms 'prior probability' and 'posterior probability' in the same way).

Here, the observation of a matching trait in the suspect's sample and stain 1, and a non-matching trait in the suspect's sample and stain 2 has made the three pairs of propositions logically equivalent, since it has become impossible for the suspect to be the donor of stain 2. Algebraically, this comes down to multiplying the value of the evidence by the corresponding pair of propositions' prior odds according to the odds' form of Bayes' theorem (Meester and Sjerps 2003):

$$\underbrace{\frac{\delta}{1-\delta}}_{\textit{prior odds for pair } H} \times \underbrace{\frac{1}{2\gamma}}_{\textit{value of the evidence for pair } H}, \tag{9}$$

[3] Note that Eq. (7) gives the simplified form of the value of the evidence, so that the numerator and denominator of this ratio do not represent the probabilities forming the numerator and denominator in Eq. (1).

$$\underbrace{\frac{\delta/2}{1-\delta}}_{\text{prior odds for pair } H_1'} \times \underbrace{\frac{1}{\gamma}}_{\text{value of the evidence for pair } H_1'}, \tag{10}$$

$$\underbrace{\frac{\delta/2}{1-\delta/2}}_{\text{prior odds for pair } H_1''} \times \underbrace{\frac{2-\delta}{2\gamma(1-\delta)}}_{\text{value of the evidence for pair } H_1''}. \tag{11}$$

Meester and Sjerps (2003) and Meester and Sjerps (2004a, b) conclude from this that forensic scientists should use posterior odds in the place of the value of the evidence to communicate the strength of forensic evidence, a recommendation which makes no attempt at clarifying the logical relationships between the three pairs of propositions and their different values for the same evidence [a point criticized by Dawid (2004)].

The Bayesian network we present in Sect. 5 will illustrate the interrelationships between these three pairs of propositions by modeling them in separate nodes. Before explaining the rationale behind this model, the next section presents the notation we will use in the rest of this paper.

4 Notation

We distinguish between the background information (Sect. 4.1), the propositions (Sect. 4.2), the unknown parameters δ, λ and τ (Sect. 4.3), and the evidence (Sect. 4.4).

4.1 Background information

The background information I is all of the knowledge available *prior* to observing the evidence. This information includes the case circumstances (e.g., the location of the crime scene), the facts surrounding the recovery of the two traces on the crime scene (e.g., their exact locations on the scene), the fact that one suspect has been found from whom a sample has been obtained for comparison with the recovered traces, and the non-scientific information associating this suspect to the crime scene (e.g., witness statements asserting the suspect's presence near the scene). All of the probabilities assessed in a case are conditional probabilities given I. However, for the sake of brevity in the mathematical expressions that follow, we shall hereafter omit I from their notation.

4.2 Propositions

The propositions reflect the viewpoints of the prosecution and the defence. At the time they are formulated, the evidence has not yet been observed, so that these formulations are independent of the evidence, and based solely on the background information. Each proposition depicts the most plausible situation(s) given the party's point of view and the background information (Robertson and Vignaux 1995). Since the background information is case-specific, one pair of propositions may be reasonable in one case, yet unreasonable in another case.

Section 1 introduced four propositions, which we denote with capital letters as follows:

D—at least one of the crime stains comes from the suspect;
$\bar{D}$—neither of the crime stains comes from the suspect;

C_1—crime stain 1 comes from the suspect;
$\bar{C}_1$—crime stain 1 does not come from the suspect.

The horizontal bar over a capital letter means that the proposition described is the negation of the proposition denoted by that letter (i.e., its complement). The number figuring as a subscript to propositions C and $\bar{C}$ indicates which crime stain the proposition refers to. Analogous to C_1 and $\bar{C}_1$, we also formulate:

C_2—crime stain 2 comes from the suspect;
$\bar{C}_2$—crime stain 2 does not come from the suspect.

Meester and Sjerps (2003) considered 3 pairs of propositions, denoted here as pairs $H, H_1', \text{and} H_1''$ (the subscript '1' indicates that the pair contains at least one proposition referring only to crime stain 1). These combine in different ways the four propositions D, $\bar{D}$, C_1 and $\bar{C}_1$ as follows:

pair H: *D—at least one of the crime stains comes from the suspect;*
$\bar{D}$—neither of the crime stains comes from the suspect;
pair H_1': *C_1—crime stain 1 comes from the suspect;*
$\bar{D}$—neither of the crime stains comes from the suspect;
pair H_1'': *C_1—crime stain 1 comes from the suspect;*
$\bar{C}_1$—crime stain 1 does not come from the suspect.

To model a pair of propositions as a node in a Bayesian network, the node must have an exhaustive list of states (see Sect. 2). The propositions in pairs H and H_1'' already form an exhaustive set of possibilities, and can therefore be modeled as nodes with two states. Yet pair H_1' is not exhaustive because it does not consider the possibility that crime stain 2 comes from the suspect. An exhaustive list would need to include all of the possible combinations between $C_1, \bar{C}_1, C_2$ and $\bar{C}_2$, i.e.,

$C_1 \cap C_2$—both crime stains come from the suspect;
$C_1 \cap \bar{C}_2$—crime stain 1 comes from the suspect, and crime stain 2 does not come from the suspect;
$\bar{C}_1 \cap C_2$—crime stain 1 does not come from the suspect, but crime stain 2 comes from the suspect;
$\bar{C}_1 \cap \bar{C}_2$—neither of the crime stains comes from the suspect.

In this list, proposition $\bar{D}$ is equivalent to $\bar{C}_1 \cap \bar{C}_2$, and proposition C_1 to $\{C_1 \cap C_2\} \cup \{C_1 \cap \bar{C}_2\}$. Modeling pair H_1' as a node with exhaustive states in a Bayesian network will therefore require the additional state $\bar{C}_1 \cap C_2$ in this node.

Analogous to pair H_1', we define pair H_2' for the combination of C_2 and $\bar{D}$:

pair H_2': *C_2—crime stain 2 comes from the suspect;*
$\bar{D}$—neither of the crime stains comes from the suspect;

and analogous to pair H_1'', we define pair H_2'' for the combination of C_2 and $\bar{C}_2$:

pair H_2'': C_2—*crime stain 2 comes from the suspect;*
$\bar{C}_2$—*crime stain 2 does not come from the suspect.*

Modeling pair H_2' as a node in a Bayesian network will follow the same reasoning as for pair H_1' by requiring the additional state $C_1 \cap \bar{C}_2$ to make the node's states exhaustive.

4.3 Unknown parameters

The two-trace problem involves three unknown parameters (Table 1):

- δ: The first parameter, δ, we encountered in Eq. (7). This is the prior probability that the suspect is a crime stain donor, i.e.,

$$Pr(D) = \delta,$$

  as defined in Meester and Sjerps (2003). δ describes the probability that a trace recovered on the crime scene comes from the suspect based on the information available prior to the laboratory analyses of the crime stains. This parameter takes into account the background information regarding the suspect's presence on or near the crime scene during the lapse of time when the traces were deposited (for example, witness statements, data from mobile phone providers, and images from surveillance cameras), as well as background information regarding the suspect's ability to transfer the type of trace evidence in question (for example in the case of recovered bloodstains, the fact that the suspect had a scratch, cut or other injury with blood loss at the time when the traces were deposited would increase δ). In this model, the value of δ is based on this background information alone, independent of whether the recovered traces come from a single source or from two different sources. Note however that in some cases this assumption of δ being independent of the total number of crime stain donors may not be reasonable. Notably when the background information described above is very poor or not available, it may be reasonable to assume that δ is greater in the case of two donors than in the case of a single donor (Meester and Sjerps 2004a, b). This situation is not treated in this paper, but it would require an additional dependence relationship in the Bayesian network presented in Sect. 5 (Fig. 2).
- λ: The second parameter describes the uncertainty on the number of donors (Dawid 2004). Defined by Dawid (2004), λ represents the probability that there are two distinct donors. Before observing the evidence, all we know is that there are two traces. A priori, these may come from the same source with a probability of $1 - \lambda$, and from two different sources with a probability of λ.
- τ: The third parameter, τ, considers the conditional probability that crime stain 1 comes from the suspect given that the suspect is one of two crime stain donors, i.e.,

$$Pr(C_1 | 2\,\text{donors}, D) = \tau.$$

From this definition, it follows that $1 - \tau$ is the probability of crime stain 2 coming from the suspect given that the suspect is the source of one of the two traces, i.e.,

Table 1 Definition of the parameters δ, λ and τ

δ	Probability that the suspect is a crime stain donor
λ	Probability that there were two distinct donors
τ	Probability that crime stain 1 comes from the suspect, given that the suspect is one of two donors

$$Pr(C_2|2\,\text{donors}, D) = 1 - \tau.$$

All of these parameters are assessed on the basis of the background information alone, that is, before observing the evidence: the value of δ will depend on the prior information regarding the suspect's connection to the crime scene; and the values of λ and τ are based on the circumstancial information of the case, including the location of each of the traces on the scene, witness reports, and images from surveillance cameras.

These prior assessments determine the prior probabilities of the propositions (see column 3 of Table 2). The probabilities of D and $\bar{D}$ are determined by δ, as described above. The probabilities of C_1, $\bar{C}_1$,C_2 and $\bar{C}_2$ are made up of the probabilities of $C_1 \cap C_2$, $C_1 \cap \bar{C}_2$, $\bar{C}_1 \cap C_2$ and $\bar{C}_1 \cap \bar{C}_2$, which are

$$\begin{aligned}
Pr(C_1 \cap C_2) &= \delta(1-\lambda)\\
Pr(C_1 \cap \bar{C}_2) &= \delta\lambda\tau\\
Pr(\bar{C}_1 \cap C_2) &= \delta\lambda(1-\tau)\\
Pr(\bar{C}_1 \cap \bar{C}_2) &= 1-\delta,
\end{aligned}$$

so that the probabilities of $C_1, \bar{C}_1, C_2$ and $\bar{C}_2$ are

$$\begin{aligned}
Pr(C_1) &= Pr(C_1 \cap C_2) + Pr(C_1 \cap \bar{C}_2)\\
&= \delta(1-\lambda) + \delta\lambda\tau,\\
Pr(\bar{C}_1) &= Pr(\bar{C}_1 \cap C_2) + Pr(\bar{C}_1 \cap \bar{C}_2)\\
&= \delta\lambda(1-\tau) + 1 - \delta,\\
Pr(C_2) &= Pr(C_1 \cap C_2) + Pr(\bar{C}_1 \cap C_2)\\
&= \delta(1-\lambda) + \delta\lambda(1-\tau),
\end{aligned}$$

and

$$\begin{aligned}
Pr(\bar{C}_2) &= Pr(C_1 \cap \bar{C}_2) + Pr(\bar{C}_1 \cap \bar{C}_2)\\
&= \delta\lambda\tau + 1 - \delta.
\end{aligned}$$

The examples in Sects. 6 and 7 will illustrate the impact of parameters δ, λ and τ on the value of the evidence and on the posterior odds of the different pairs of propositions.

4.4 Evidence

The evidence is the new piece of information we observe. It is the compound event of observing the states of the three variables X, Y_1 and Y_2. X denotes the profile of

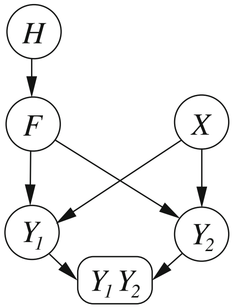

Fig. 1 The Bayesian network presented in Taroni et al. (2006) for a very specific scenario of the two-trace problem. Nodes H, X, Y_1 and Y_2 consist of the states presented in Table 2, and node F of the states $C_1 \cap \bar{C}_2$, $\bar{C}_1 \cap C_2$ and $\bar{C}_1 \cap \bar{C}_2 \equiv \bar{D}$

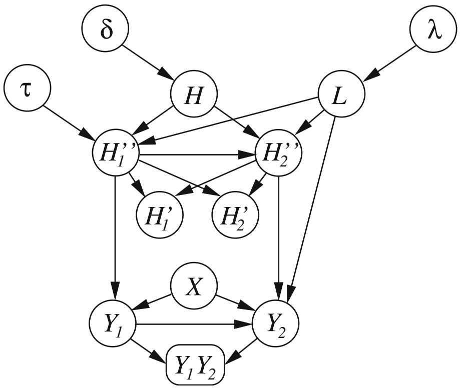

Fig. 2 The extended Bayesian network for the two-trace problem. This model is more flexible and realistic than the Bayesian network shown in Fig. 1, because it models the uncertainty on the number of crime stain donors, and the uncertainty on which trace comes from the suspect if the suspect is one of two donors. It also includes a node for each of the unknown parameters, allowing the user to define a probability distribution for each. Table 1 gives the definitions of the parameters, and Table 2 lists the definitions and probabilities of the states in each of the non-parametric nodes

the suspect's sample, Y_1 the profile of the first of the recovered traces, which we call 'crime stain 1', and Y_2 the profile of the second of the recovered traces, which we call 'crime stain 2'.

We assume that the analysis performed is capable of distinguishing between k different profiles, which we label $\Gamma_1, \Gamma_2, \ldots, \Gamma_k$. Profile Γ_i, $i = 1, 2, \ldots, k$, has a match probability of γ_i in the relevant population of possible crime stain donors. Note that the relevant population is defined on the basis of the background information. Before observing the evidence, X, Y_1, and Y_2 each have a probability of γ_i, $i = 1, 2, \ldots, k$, $\sum \gamma_i = 1$, to have profile Γ_i. After observing the evidence, the states of X, Y_1, and Y_2 are known with certainty. They are each equal to one of the k profiles, $\Gamma_1, \Gamma_2, \ldots, \Gamma_k$.

In the next section, we combine the above evidence, propositions, and parameters in a Bayesian network for the two-trace problem.

Table 2 Description of the states of the non-parametric nodes in the Bayesian network in Fig. 2

Nodes	States	Probabilities	Definitions of the states
H	D	δ	At least one of the crime stains comes from the suspect
	$\bar{D}$	$1-\delta$	Neither of the crime stains comes from the suspect
L	1 donor	$1-\lambda$	The crime stains come from the same source
	2 donors	λ	The crime stains come from two different sources
H_1''	C_1	$\delta(1-\lambda+\lambda\tau)$	Crime stain 1 comes from the suspect
	$\bar{C}_1$	$\delta\lambda(1-\tau)+1-\delta$	Crime stain 1 does not come from the suspect
H_2''	C_2	$\delta[1-\lambda+\lambda(1-\tau)]$	Crime stain 2 comes from the suspect
	$\bar{C}_2$	$\delta\lambda\tau+1-\delta$	Crime stain 2 does not come from the suspect
H_1'	C_1	$\delta(1-\lambda+\lambda\tau)$	Crime stain 1 comes from the suspect
	$\bar{C}_1\cap C_2$	$\delta\lambda(1-\tau)$	Only crime stain 2 comes from the suspect
	$\bar{D}$	$1-\delta$	Neither of the crime stains comes from the suspect
H_2'	C_2	$\delta[1-\lambda+\lambda(1-\tau)]$	Crime stain 2 comes from the suspect
	$C_1\cap\bar{C}_2$	$\delta\lambda\tau$	Only crime stain 1 comes from the suspect
	$\bar{D}$	$1-\delta$	Neither of the crime stains comes from the suspect
X	Γ_1	γ_1	Profile of the suspect's sample
	Γ_2	γ_2	
	$\vdots$	$\vdots$	
	Γ_k	γ_k	
Y_1	Γ_1	γ_1	Profile of crime stain 1
	Γ_2	γ_2	
	$\vdots$	$\vdots$	
	Γ_k	γ_k	
Y_2	Γ_1	γ_1	Profile of crime stain 2
	Γ_2	γ_2	
	$\vdots$	$\vdots$	
	Γ_k	γ_k	
Y_1Y_2	$\Gamma_1\Gamma_1$	$\gamma_1\ \gamma_1$	Profiles of crime stains 1 and 2 (as ordered pairs)
	$\Gamma_1\Gamma_2$	$\gamma_1\gamma_2$	
	$\Gamma_2\Gamma_1$	$\gamma_2\ \gamma_1$	
	$\vdots$	$\vdots$	
	$\Gamma_k\Gamma_k$	$\gamma_k\ \gamma_k$	

The parameters δ, λ, and τ are defined in Table 1

5 Constructing a Bayesian network

The aim of this section is to construct a Bayesian network containing the propositions, the parameters, and the evidence, defined in the previous section. This section contains several technical details of the constructed Bayesian network, and may be skipped by readers interested more in the application of the model than in its construction. For the propositions we create nodes H, H_1', H_1'', H_2' and H_2'', and for the evidence, nodes X, Y_1 and Y_2. Table 2 provides the exhaustive list of the states and probabilities associated to each of these nodes.

Taroni et al. (2006) proposed a model containing some of these nodes for a very specific scenario of a two-trace problem (Fig. 1). In this model, node F contains the inexhaustive list of states $C_1 \cap \bar{C}_2, \bar{C}_1 \cap C_2$ and $\bar{C}_1 \cap \bar{C}_2$. This model sets the profile of Y_1 equal to the profile of X if C_1 is true, and the profile of Y_2 equal to the profile of X if C_2 is true. Concerning the propositions, it specifies that C_1 and C_2 can only be true if D (in node H) is true. However, this model makes the assumption that C_1 and C_2 are equally likely under D, and it does not consider the possibility of C_1 and C_2 being true at the same time (i.e., node F contains an inexhaustive list of states). Node Y_1Y_2 combines the states of Y_1 and Y_2 as ordered pairs, so that the model computes the compound probability of the two crime stain profiles. This node is necessary for evaluating the value of the evidence (see Sect. 7).

We use this model as a starting point to extend and improve it to a more general Bayesian network for evaluating the value of the evidence in a two-trace problem. For this, we examine the following points: the relationship between the propositional nodes (Sect. 5.1), the uncertainty on the number of donors (Sect. 5.2), and the relationship between the propositional and the evidential nodes (Sect. 5.3).

5.1 Relationships between the propositional nodes H, H_1', H_1'', H_2' and H_2''

The postdata equivalence presented in Sect. 3 indicates a relationship between the nodes containing the pairs of propositions H, H_1' and H_1''. To expose the links that exist between these nodes, we analyze the logical relationships between the propositions that form the nodes' states.

The difference between proposition D and propositions C_1 and C_2 is that the former does not specify which trace, or traces, come(s) from the suspect, whereas the latter do. Logically, this means that propositions C_1 and C_2 are two subsets of proposition D, i.e., $C_1 \subset D$ and $C_2 \subset D$. In a Bayesian network, this relationship may be modeled by conditioning the probabilities of C_1 and C_2 on D (Taroni et al. 2006). In other words, we model node H (containing proposition D) as a parent of nodes H_1'' (containing proposition C_1) and H_2'' (containing proposition C_2).

As for nodes H_1' and H_2', their states $\bar{C}_1 \cap C_2$, $C_1 \cap \bar{C}_2$ and $\bar{D}$ ($\equiv \bar{C}_1 \cap \bar{C}_2$) are combinations of C_1, $\bar{C}_1$, C_2, and $\bar{C}_2$. Each of these combinations is a subset of its single components: $\{C_1 \cap \bar{C}_2\} \subset C_1$, $\{C_1 \cap \bar{C}_2\} \subset \bar{C}_2$, $\{\bar{C}_1 \cap C_2\} \subset \bar{C}_1$, $\{\bar{C}_1 \cap C_2\} \subset C_2$, $\bar{D} \subset \bar{C}_1$ and $\bar{D} \subset \bar{C}_2$. Again, we find it convenient to model a subset as a child of its superset. Therefore, we model nodes H_1' and H_2' as children of nodes H_1'' and H_2'', with the conditional probability distributions given in Tables 3 and 4.

The resulting hierarchical ordering, from the parent node to the child node, is therefore:

$$H \rightarrow \{H_1'', H_2''\} \rightarrow \{H_1', H_2'\}.$$

Our Bayesian network will reflect this hierarchy.

Table 3 Probability table for node H_1' in Fig. 2. The states of H_1' are defined by the combinations of the states in nodes H_1'' and H_2''

	H_1'':	C_1		$\bar{C}_1$	
	H_2'':	C_2	$\bar{C}_2$	C_2	$\bar{C}_2$
H_1':	C_1	1	1	0	0
	$\bar{C}_1 \cap C_2$	0	0	1	0
	$\bar{D}$	0	0	0	1

Table 4 Probability table for node H_2' in Fig. 2. The states of H_2' are defined by the combinations of the states in nodes H_1'' and H_2''

	H_1'':	C_1		$\bar{C}_1$	
	H_2'':	C_2	$\bar{C}_2$	C_2	$\bar{C}_2$
H_2':	C_2	1	0	1	0
	$C_1 \cap \bar{C}_2$	0	1	0	0
	$\bar{D}$	0	0	0	1

5.2 Uncertainty on the number of donors

To take into account the possibility that there was only one donor, we add an additional node L made up of the states *'1 donor'* and *'2 donors'*. We use the parameter λ, denoting the prior probability of *'2 donors'*, to introduce the uncertainty on the number of donors into this node.

The states of node L add a constraint on the probability distribution over $C_1, \bar{C}_1, C_2$, and $\bar{C}_2$, and on the observed profile of crime stain 2 (Y_2) given the profile of crime stain 1 (Y_1). That is, if there is only 1 donor, then Y_2 must be equal to Y_1, and both C_1 and C_2 must be true or false, together, according to whether D is true or false. If there are 2 donors, then either C_1 or C_2 will be true when D is true, but never both C_1 and C_2. In the case of two donors, the parameter τ (denoting $Pr(C_1|2$ donors, D)) determines the probability distribution over C_1 and C_2 under proposition D. Tables 5 and 6 describe the logical relationships between the propositions $C_1, \bar{C}_1, C_2$, and $\bar{C}_2$ and the propositions D and $\bar{D}$ given the number of donors specified in node L.

5.3 Relationship between the propositional and evidential nodes

As proposed by Taroni et al. (2006), the profile of each crime stain depends on whether that particular crime stain comes from the suspect, i.e., on propositions C_1 and C_2. This means that node Y_1 should be connected with a node containing state C_1, and node Y_2 with a node containing state C_2. The most straightforward way of

Table 5 Probability table for node H_1'' in Fig. 2

	H:	D		$\bar{D}$	
	L:	1 donor	2 donors	1 donor	2 donors
H_1'':	C_1	1	τ	0	0
	$\bar{C}_1$	0	$1 - \tau$	1	1

This probability table contains the parameter $\tau = Pr(C_1|2 \text{ donors}, D)$

Table 6 Probability table for node H_2'' in Fig. 2

	H:	D				$\bar{D}$			
	L:	1 donor		2 donors		1 donor		2 donors	
	H_1'':	C_1	$\bar{C}_1$	C_1	$\bar{C}_1$	C_1	$\bar{C}_1$	C_1	$\bar{C}_1$
H_2'':	C_2	1	n/a	0	1	n/a	0	n/a	0
	$\bar{C}_2$	0	n/a	1	0	n/a	1	n/a	1

Note that the second, fifth and seventh columns describe impossible combinations of states (i.e., in the second column, the suspect is a crime stain donor, and there is only a single donor for both crime stains, yet the suspect is not the donor of crime stain 1; and in the fifth and seventh columns the suspect is not a crime stain donor, yet crime stain 1 comes from the suspect), so that the probability distribution over states C_2 and $\bar{C}_2$ is not defined for these events ('n/a' = not applicable). For an alternative way of modeling this conditional probability distribution over the states of node H_2'' that avoids having these impossible combinations in the conditional probability table, we refer the reader to the work by Fenton et al. (2011)

achieving this in the model is for Y_1 to be a child of H_1'', and Y_2 a child of H_2''. Thus, Y_1 copies the state of X when C_1 is true, and is independent of X when $\bar{C}_1$ is true (see Table 7). The same principle holds for Y_2 (Table 8), with the additional constraint that Y_2 copies the state of Y_1 in every case where both traces come from the same source (defined by node L). Finally, node Y_1Y_2 combines the states of Y_1 and Y_2 as ordered pairs, as proposed by Taroni et al. (2006) (see Table 9). Putting all of these considerations together produces the Bayesian network shown in Fig. 2.

There are two ways to use the Bayesian network, which we will illustrate in the next two sections: the user can either update the prior probability distributions over the propositions to posterior probability distributions given the evidence (see Sect. 6), or the user can use the Bayesian network to evaluate the probabilities forming the ratio of the value of the evidence (Eq. 1) for a given pair of propositions (see Sect. 7). Both of these are useful means for a forensic scientist to convey the value of the evidence to a fact-finder.

6 Using the Bayesian network to update the prior probability distribution to a posterior probability distribution

Fact-finders and lawyers are interested in the probability distribution over the propositions given the forensic scientist's evidence. The Bayesian network

Table 7 Probability table for node Y_1 in Fig. 2. If C_1 is true, the state of Y_1 is equal to the state of X. If $\bar{C}_1$ is true, the probability of observing each profile is equal to that profile's match probability in the relevant population of possible crime stain donors

	H_1'':	C_1			$\bar{C}_1$		
	X:	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}
Y_1:	Γ_1	1	0	0	γ_1	γ_1	γ_1
	Γ_2	0	1	0	γ_2	γ_2	γ_2
	Γ_{other}	0	0	1	$1-\gamma_1-\gamma_2$	$1-\gamma_1-\gamma_2$	$1-\gamma_1-\gamma_2$

presented in Fig. 2 can compute this posterior probability distribution for a given prior probability distribution over the propositions. There are two applications where a forensic scientist testifying in court would use the model in this way: (1) when a fact-finder or lawyer interested in the probabilities of the propositions communicates the information required to define the prior distribution to the forensic scientist, (2) when the forensic scientist wants to illustrate the evidence's effect on several prior probability distributions of different orders of magnitude to show what the posterior probability distribution would be for each.

To specify the prior probability of each proposition, the user must assess the values of δ, λ and τ (see the definitions given in Table 1 and in Sect. 4.3). Practically speaking, the user of the model must enter these values into the Bayesian network, an action called 'instantiating' the corresponding nodes. The Bayesian network then propagates these values to the rest of the network.

To find the posterior probability distribution given the evidence, the user instantiates the observed traits for the suspect's sample and the two traces in nodes X, Y_1 and Y_2, respectively. After entering this evidence, the Bayesian network updates the probability distributions in the remaining nodes according to the laws of probability and the probabilistic relationships specified by the model. The probability distributions indicated in the propositional nodes now correspond to the posterior distributions given the evidence. Mathematically, this updating corresponds to the application of the laws of probability, in particular Bayes' theorem. The following numerical example illustrates this concept.

Example Consider a case where crime scene investigators recover two contact stains on a wall, at a given height above the floor: say, crime stain 1 at 1.5 meters, and crime stain 2 at 1.8 meters from the floor. There are no witness statements asserting whether these two traces come from a single source or from two different sources. We assume that it is, a priori, equally probable for the two traces to come from a single source as it is for them to come from two different sources, and thus set $\lambda = 0.5$. A suspect, with a prior probability assessed at $\delta = 0.1$ of being the source of at least one of the two traces recovered on the crime scene, comes to the attention of the police. This suspect is particular in that he is very short, measuring only 1.6 meters. This information makes a contact between the suspect and the location of crime stain 1 more probable than a contact between the suspect and the location of crime stain 2. In other words, if only one of the two traces comes from

Table 8 Probability table for node Y_2 in Fig. 2. If $L =$ *1 donor*, the state of Y_2 is equal to the state of Y_1. If $L =$ *2 donors*, the probability table for node Y_2 is identical to the probability table for node Y_1 (Table 7)

	L:	1 donor											
	Y_1:	Γ_1						Γ_2					
	H_2''	C_2			$\bar{C}_2$			C_2			$\bar{C}_2$		
	X:	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}
Y_2:	Γ_1	1	1	1	1	1	1	0	0	0	0	0	0
	Γ_2	0	0	0	0	0	0	1	1	1	1	1	1
	Γ_{other}	0	0	0	0	0	0	0	0	0	0	0	0

1 donor						2 donors					
Γ_{other}						Γ_1					
C_2			$\bar{C}_2$			C_2			$\bar{C}_2$		
Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}
0	0	0	0	0	0	1	0	0	γ_1	γ_1	γ_1
0	0	0	0	0	0	0	1	0	γ_2	γ_2	γ_2
1	1	1	1	1	1	0	0	1	$1-\gamma_1-\gamma_2$	$1-\gamma_1-\gamma_2$	$1-\gamma_1-\gamma_2$

2 donors											
Γ_2						Γ_{other}					
C_2			$\bar{C}_2$			C_2			$\bar{C}_1$		
Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}
1	0	0	γ_1	γ_1	γ_1	1	0	0	γ_1	γ_1	γ_1
0	1	0	γ_2	γ_2	γ_2	0	1	0	γ_2	γ_2	γ_2
0	0	1	$1-\gamma_1-\gamma_2$	$1-\gamma_1-\gamma_2$	$1-\gamma_1-\gamma_2$	0	0	1	$1-\gamma_1-\gamma_2$	$1-\gamma_1-\gamma_2$	$1-\gamma_1-\gamma_2$

Table 9 Probability table for node Y_1Y_2. This node combines the states of Y_1 and Y_2 as ordered pairs

	Y_1:	Γ_1			Γ_2			Γ_{other}		
	Y_2:	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}	Γ_1	Γ_2	Γ_{other}
Y_1Y_2:	$\Gamma_1\Gamma_2$	0	1	0	0	0	0	0	0	0
	other	1	0	1	1	1	1	1	1	1

this suspect, it is more probable for this stain to be crime stain 1 than crime stain 2. For this reason, we set $\tau = 0.75$.

The following analysis compares the probability distributions for the three different types of pairs of propositions by considering pairs H, H_1' and H_1'' (Note that the Bayesian network presented in Fig. 2 allows for the same analysis with regard to pairs H, H_2' and H_2'', focusing on crime stain 2 instead of on crime stain 1. Here, however, we will focus on crime stain 1). For this, node H_2' is superfluous in the Bayesian network (Fig. 2). In this section, Fig. 3 has omitted this node to avoid cluttering the Bayesian network's expanded representations.

Figure 3a gives the prior probability distribution over the nodes of the Bayesian network for the above described example. The ratio of the probabilities of the propositions of each pair forms the following prior odds for the three pairs of propositions defined in Sect. 4.2:[4]

$$\frac{\delta}{1-\delta} = \frac{0.1000}{0.9000} = 0.1111 \quad \text{for pair } H, \tag{12}$$

$$\frac{\delta(1-\lambda+\lambda\tau)}{1-\delta} = \frac{0.0875}{0.9000} = 0.0972 \quad \text{for pair } H_1', \text{ and} \tag{13}$$

$$\frac{\delta(1-\lambda+\lambda\tau)}{\delta\lambda(1-\tau)+1-\delta} = \frac{0.0875}{0.9125} = 0.0959 \quad \text{for pair } H_1''. \tag{14}$$

With these numerical calculations, we do not imply that it is possible to attain this level of precision in practice. The precision of the numerical calculations in Eqs. (12)–(14), and in the equations of the rest of this and the next section, is only for the purpose of showing the level of agreement between the Bayesian network's computations and the algebraic equations.

The comparison of the above results with the prior odds given in Meester and Sjerps (2003) (see Eqs. 9–11) shows that the latter describe a case where $\lambda = 1$ and $\tau = 0.5$. The above expressions relax these assumptions by allowing the user to specify parameters λ and τ so that they reflect the circumstances of the case as accurately as possible. Comparing the prior odds for each of the pairs of propositions with each other reveals that the most general pair of propositions (pair H) has the greatest odds, and the most specific pair of propositions (pair H_1'') has the smallest odds. This is logical since the prior odds for a specific crime stain cannot be

[4] Note that, by definition, pair H_1' consists of two nonexhaustive propositions. This is not problematic in this situation, because the evidence introduced later on will render the remaining proposition impossible.

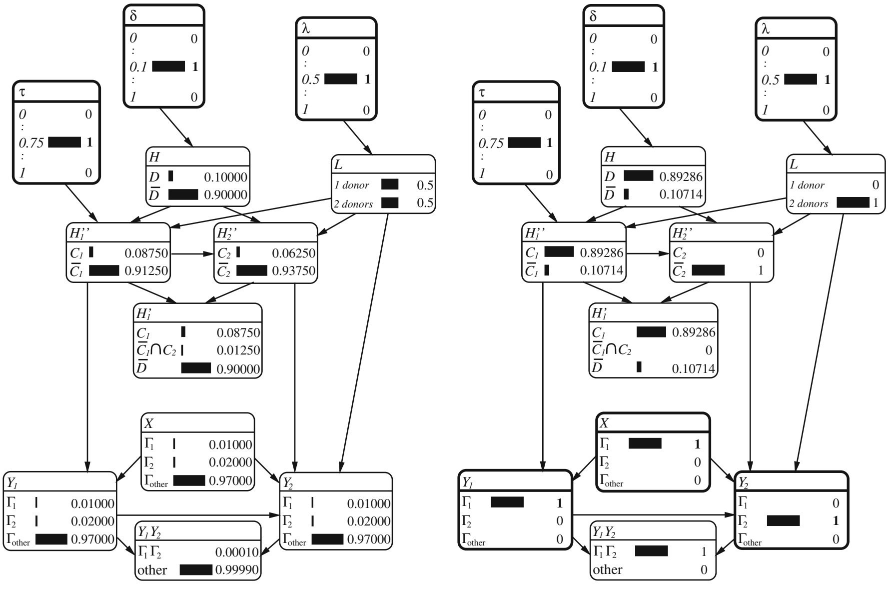

Fig. 3 Expanded representation of the Bayesian network presented in Fig. 2, without node H_2', which has been omitted to avoid cluttering the figure, and to focus the reader's attention on the probability distributions in nodes H, H_1' and H_1''. This Bayesian network updates **a** the prior probability distribution over the propositions, to **b** the posterior probability distribution obtained after observing the traits of X, Y_1 and Y_2. Here, the model is applied to the example described on pages 16 and 18, with $\delta = 0.1$, $\lambda = 0.5$, $\tau = 0.75$, $\gamma_1 = 0.01$ and $\gamma_2 = 0.02$. The observed evidence consists of $X = \Gamma_1, Y_1 = \Gamma_1$ and $Y_2 = \Gamma_2$. This information is communicated to the Bayesian network by instantiating the evidential nodes to these observed states. Here, the instantiated nodes are indicated by a *thicker border*, and the instantiated state with a probability of 1 in *bold*. Instantiating the evidential nodes produces identical posterior odds for the pairs of propositions H, H_1' and H_1''. Note that in **b**, we could also have instantiated node Y_1Y_2 instead of nodes Y_1 and Y_2 and obtained the same outcome

greater than the general prior odds for the suspect being a donor of any one of the crime stains.

Example (*continued*) We now analyze the evidence, and observe $Y_1 = \Gamma_1, Y_2 = \Gamma_2$ and $X = \Gamma_1$, i.e., the suspect's sample matches crime stain 1. In the population of potential sources of the two traces, we assume $\gamma_1 = 0.01$ and $\gamma_2 = 0.02$.

Instantiating the evidential nodes X, Y_1 and Y_2 to their observed traits (Fig. 3b), produces identical posterior odds of

$$\frac{0.89286}{0.10714} \approx 8.3333$$

for all three pairs of propositions. Algebraically, these posterior odds are given by

$$\frac{\delta\tau}{(1-\delta)\gamma_1}. \tag{15}$$

The comparison of this ratio with the posterior odds presented in Meester and Sjerps (2003) (see Eq. 8) shows that Eq. (15) relaxes the assumption of $\tau = 0.5$, assumed in Eq. (8). Equation (15) therefore gives the generalized expression for the posterior odds for any value of τ.

This approach of instantiating the evidential nodes in the Bayesian network is useful whenever one wants to find a posterior probability distribution for a given prior probability distribution. This application is limited to situations where the forensic scientist receives information about the prior probability distribution from an actor in the legal system, or situations where the forensic scientist assigns hypothetical prior distributions to illustrate the evidence's effect on the probabilities of the propositions. However, the forensic scientist's role is not to determine the probability distribution over the propositions. The role of the forensic scientist is to evaluate the value of the evidence (e.g., Lindley 1977; Aitken and Taroni 2004). This means that he/she wants to find out to what extent the observed evidence will affect the probability distribution over the propositions, without knowing what this probability distribution is.

In addition to computing the posterior probabilities seen in this section, the Bayesian network allows its user to evaluate the probabilities forming the value of the evidence for any of the three pairs of propositions. We discuss this use of the Bayesian network in the next section.

7 Using the Bayesian network to evaluate the value of the evidence

The objective of the forensic scientist's testimony is to present the value of the evidence. That is, he/she should present how much more or less probable the evidence is if the first proposition is true than if the second proposition is true. This value depends on the formulation of the two propositions. The value of the evidence for each pair of propositions corresponds to the Bayes factor obtained by dividing the posterior odds by the prior odds (Table 10).

Mathematically, this value is given by Eq. (1). Applying the third law of probability for dependent events according to a suspect-anchored perspective (e.g., Aitken and Taroni 2004) makes this equation equal to

$$\begin{aligned} V &= \frac{Pr(X, Y_1, Y_2 | \text{proposition 1})}{Pr(X, Y_1, Y_2 | \text{proposition 2})} \\ &= \frac{Pr(Y_1, Y_2 | X, \text{proposition 1})}{Pr(Y_1, Y_2 | X, \text{proposition 2})} \times \frac{Pr(X | \text{proposition 1})}{Pr(X | \text{proposition 2})}, \end{aligned}$$

which reduces to

$$= \frac{Pr(Y_1, Y_2 | X, \text{proposition 1})}{Pr(Y_1, Y_2 | X, \text{proposition 2})}, \tag{16}$$

given that the profile of the suspect's sample does not change under the competing propositions, i.e., $Pr(X|\text{proposition 1}) = \Pr(X|\text{proposition 2})$.

So, to find the value of the evidence, the Bayesian network calculates the probabilities forming the numerator and the denominator of Eq. (16). The Bayesian

Table 10 The mathematical expressions used by the Bayesian network in Fig. 2 to compute the prior odds, value of the evidence (Bayes factor) and posterior odds for each of the three pairs of propositions, H, H_1' and H_1''

for pair H:	$\underbrace{\frac{\delta}{1-\delta}}_{\textit{prior odds}} \times \underbrace{\frac{\tau}{\gamma_1}}_{V} = \underbrace{\frac{\delta\tau}{(1-\delta)\gamma_1}}_{\textit{posterior odds}}$
for pair H_1':	$\underbrace{\frac{\delta(1-\lambda+\lambda\tau)}{1-\delta}}_{\textit{prior odds}} \times \underbrace{\frac{\frac{\tau}{1-\lambda+\lambda\tau}}{\gamma_1}}_{V} = \underbrace{\frac{\delta\tau}{(1-\delta)\gamma_1}}_{\textit{posterior odds}}$
for pair H_1'':	$\underbrace{\frac{\delta(1-\lambda+\lambda\tau)}{\delta\lambda(1-\tau)+1-\delta}}_{\textit{prior odds}} \times \underbrace{\frac{\frac{\tau}{1-\lambda+\lambda\tau}}{\gamma_1\frac{1-\delta}{(1-\tau)\lambda\delta+1-\delta}}}_{V} = \underbrace{\frac{\delta\tau}{(1-\delta)\gamma_1}}_{\textit{posterior odds}}$

For the definitions of δ, λ and τ, see Table 1

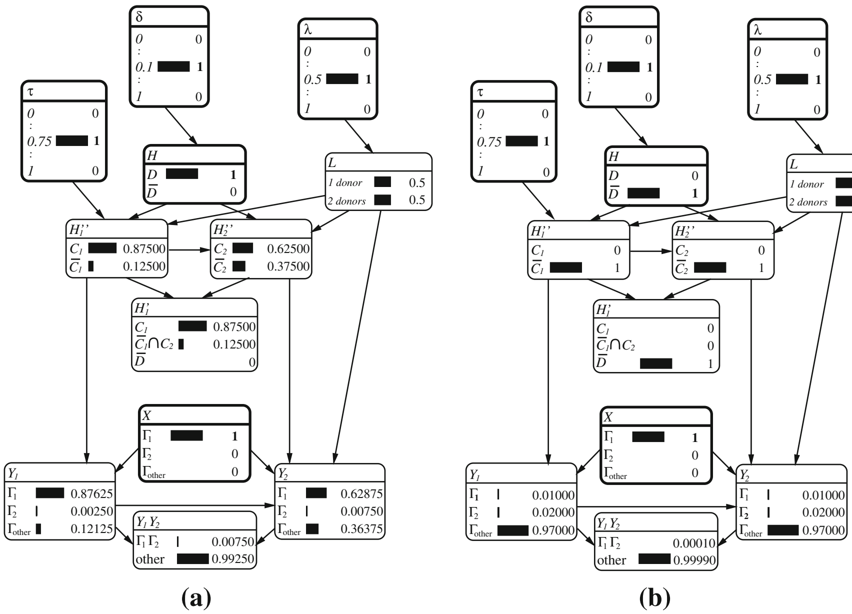

Fig. 4 The Bayesian network computes **a** the numerator, and **b** the denominator of the value of the evidence (Eq. 16) for pair of propositions H, for evidence consisting of $X = \Gamma_1, Y_1 = \Gamma_1$ and $Y_2 = \Gamma_2$. The instantiated nodes are indicated by a *thicker border*, and the instantiated state with a probability of 1 in *bold*. The numerator is the probability of $\Gamma_1\Gamma_2$ in node Y_1Y_2 when Γ_1 is instantiated in node X and D is instantiated in node H. The denominator is the probability of $\Gamma_1\Gamma_2$ in node Y_1Y_2 when Γ_1 is instantiated in node X and $\bar{D}$ is instantiated in node H. The calculations are for the example described in Sect. 6 ($\delta = 0.1$, $\lambda = 0.5$ and $\tau = 0.75$)

network computes the compound probability of Y_1 and Y_2 in node Y_1Y_2. This node indicates the numerator of V for the observed traits of Y_1 and Y_2 when X and 'proposition 1' are instantiated, and the denominator of V when X and 'proposition

2' are instantiated.[5] Figures 4, 5 and 7 illustrate the results obtained in this way for each of the three pairs of propositions, H, H_1' and H_1'', for the numerical example presented in Sect. 6. Again, the expanded representations of the Bayesian network omit node H_2' to avoid cluttering these figures, and to focus the reader's attention on the propositional nodes H, H_1' and H_1''. In the following sections, we discuss each value in turn, and examine how each is affected by the parameters τ, λ and δ.

7.1 The value of the evidence for pair H

According to Fig. 4, the value of the evidence for pair H is equal to

$$V = \frac{0.0075}{0.0001} = 75.$$

Algebraically, this value is given by

$$V = \frac{\lambda\tau\gamma_2}{\lambda\gamma_1\gamma_2} = \frac{\tau}{\gamma_1}. \tag{17}$$

The numerator describes the probability of observing the evidence given that at least one of the crime stains comes from the suspect (proposition D). In this case, the observation of the evidence is only possible when the two traces come from two different donors (for which the probability is λ), of which the suspect is the donor of the first trace (probability τ), and someone with trait Γ_2 the donor of the second trace (probability γ_2). The denominator describes the probability of observing the evidence given that neither of the crime stains comes from the suspect (proposition $\bar{D}$). In this case, the observation of the evidence corresponds to the event that the two traces come from two different donors (probability λ), of which one has trait Γ_1 (probability γ_1) and the other trait Γ_2 (probability γ_2).

For this pair of propositions, V reduces to a linear function of τ, ranging from a minimum of 0 when $\tau = 0$ (i.e., when it is a priori impossible for the suspect to be the source of trace 1 in a case where the suspect is the source of one of the two traces), to a maximum of $\frac{1}{\gamma_1}$ for $\tau = 1$ (i.e., when it is a priori certain that the suspect is the source of trace 1 in the case that the suspect is the source of one of the two traces). In the latter case, the value of the evidence is the same as in a one-trace problem, because, just as in a one-trace problem, it becomes certain to observe a match between crime stain 1 and the suspect's sample if proposition D is true.

When $\tau = 0.5$, this means that it is equally likely for either of the two traces to come from the suspect in a case where the suspect is one of two crime stain donors. This is the additional factor multiplied by $\frac{1}{\gamma_1}$ to produce Eq. (4) derived by Evett (1987) for the value of one matching stain and one non-matching stain. Underlying Eq. (4) is therefore the assumption that each of the two traces is equally likely to come from the suspect if the suspect is the source of one of the two traces. Yet, as

[5] Note that the Bayesian network presented here models the probability of Y_1 and Y_2 as independent of the suspect's sample given 'proposition 2'. This makes the probability of Y_1 and Y_2 when 'proposition 2' and X are instantiated identical to the probability of Y_1 and Y_2 when only 'proposition 2' is instantiated, so that the instantiation of X is not absolutely necessary in this case.

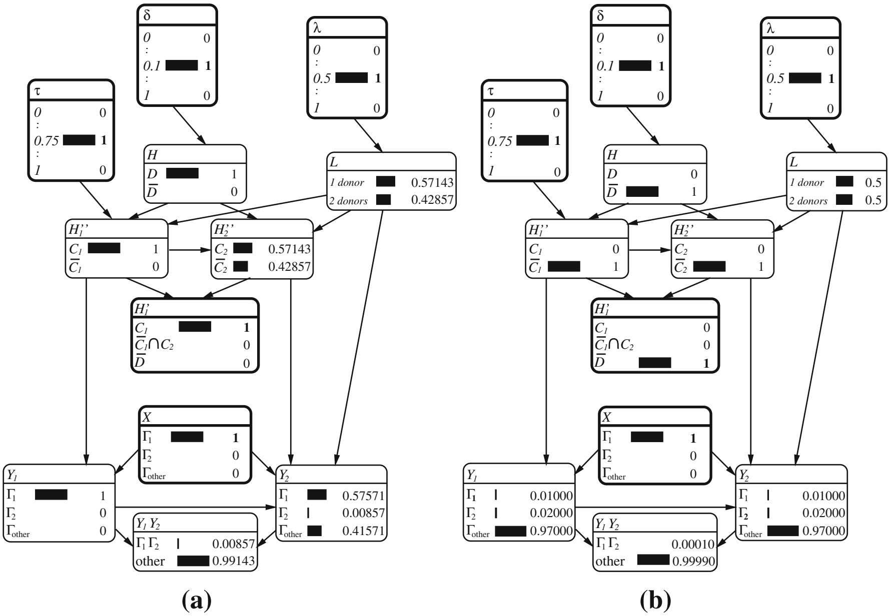

Fig. 5 The Bayesian network computes **a** the numerator, and **b** the denominator of the value of the evidence (Eq. 16) for pair of propositions H_1', for evidence consisting of $X = \Gamma_1, Y_1 = \Gamma_1$ and $Y_2 = \Gamma_2$. The nodes with the *thicker borders* are the instantiated nodes, with the instantiated state indicated with a probability of 1 in *bold*. The numerator is the probability of $\Gamma_1\Gamma_2$ in node Y_1Y_2 when Γ_1 is instantiated in node X and C_1 is instantiated in node H_1', and the denominator is the probability of $\Gamma_1\Gamma_2$ in node Y_1Y_2 when Γ_1 is instantiated in node X and $\bar{D}$ is instantiated in node H_1'. These calculations are for the scenario described in Sect. 6 ($\delta = 0.1$, $\lambda = 0.5$ and $\tau = 0.75$)

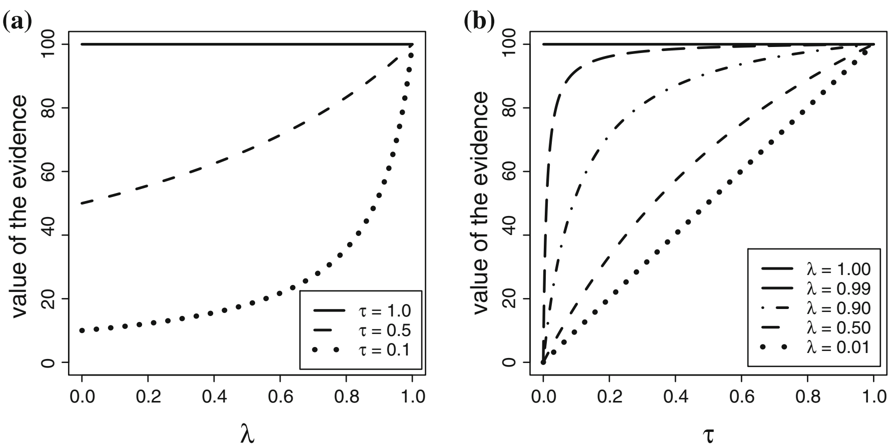

Fig. 6 The value of the evidence V for pair H_1' as a function of **a** λ and **b** τ. Here, $\gamma_1 = 0.01$. V is an increasing function of λ and τ, attaining a maximal value of $\frac{1}{\gamma_1}$ when $\lambda = 1$ or $\tau = 1$, for $\lambda \neq 0$, $\tau \neq 0$

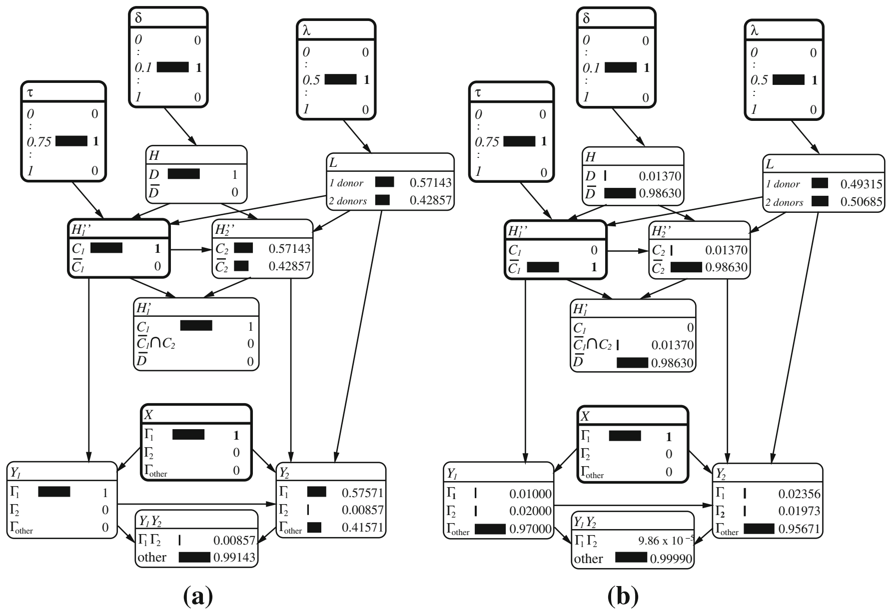

Fig. 7 The Bayesian network computes **a** the numerator and **b** the denominator of the value of the evidence (Eq. 16) for pair of propositions H_1'', for evidence consisting of $X = \Gamma_1, Y_1 = \Gamma_1$ and $Y_2 = \Gamma_2$. The nodes with the *thicker borders* are the instantiated nodes, with the instantiated state indicated with a probability of 1 in *bold*. The numerator is the probability of $\Gamma_1\Gamma_2$ in node Y_1Y_2 when Γ_1 is instantiated in node X and C_1 is instantiated in node H_1'', and the denominator is the probability of $\Gamma_1\Gamma_2$ in node Y_1Y_2 when Γ_1 is instantiated in node X and $\bar{C}_1$ is instantiated in node H_1''. These calculations are for the scenario described in Sect. 6 ($\delta = 0.1$, $\lambda = 0.5$ and $\tau = 0.75$)

seen in the example on page 16, the two crime stains may not have the same prior probability of coming from the suspect if they were recovered at different locations on the crime scene. In this case, it is necessary to replace Eq. (4) with Eq. (17), and assign a more adequate value for τ based on the circumstances of the case.

7.2 The value of the evidence for pair H_1'

For pair H_1', Fig. 5 shows that the value of the evidence is equal to

$$V = \frac{0.00857}{0.0001} = 85.7,$$

which, algebraically, corresponds to

$$V = \frac{\frac{\lambda\tau}{1-\lambda+\lambda\tau}\gamma_2}{\lambda\gamma_1\gamma_2} \tag{18}$$

$$= \frac{\frac{\tau}{1-\lambda+\lambda\tau}}{\gamma_1}. \tag{19}$$

Here, the probability in the numerator is the probability of observing the evidence given that crime stain 1 comes from the suspect (proposition C_1). A priori, there are two possibilities if proposition C_1 is true: either both traces come from the suspect (for which the probability is $1 - \lambda$), or only crime stain 1 comes from the suspect (for which the probability is $\lambda\tau$). The observation of one matching and one non-matching trace is impossible if both traces come from the suspect. The probability of observing the evidence is therefore the normalized probability for the latter case times the probability that the donor of the second trace has trait Γ_2 (probability γ_2), as shown in the numerator of Eq. (18). The denominator remains the same as for pair H. For this pair of propositions, V is an increasing function of both τ and λ for all $\tau < 1$ and $\lambda < 1$ (Fig. 6). V attains the maximum value of $\frac{1}{\gamma_1}$ when at least one of these parameters is equal to 1:

- When $\lambda = 1$, it is certain that the two traces come from two different sources. In this case, it is, a priori, certain that crime stain 2 does not come from the suspect given proposition C_1, and the only possibility left under this proposition is that the suspect is the source of only crime stain 1. The normalized probability in the numerator of Eq. (18) therefore reduces to 1, so that the probability of observing the evidence given C_1 and $X = \Gamma_1$ is equal to $1 \times \gamma_2$. With the denominator, this reduces the numerator of V to 1.
- When $\tau = 1$, it is a priori certain that trace 1 comes from the suspect if exactly one of the traces comes from the suspect. The normalized probability in the numerator of Eq. (18) thus reduces to λ. The probability of observing the evidence is therefore equal to the probability that the two traces come from two different sources times the probability that the other donor has trait Γ_2, i.e., $\lambda \times \gamma_2$. With the denominator, this reduces the numerator of V to 1.

According to the reasoning in Meester and Sjerps (2003), Eq. (6) was obtained for $\lambda = 1$ and $\tau = 0.5$. In this case, V is maximum for this pair of propositions because of $\lambda = 1$. However, the assumption of $\lambda = 1$ can only be made in very specific cases. By definition, this assumption must be made *before* observing the evidence, so it can only be based on other information in the case. For example, one could imagine a case with a crime scene in a location under high surveillance and cleaned on a regular and scheduled basis: here, a surveillance camera showing two unidentifiable individuals on the scene on the day the traces were deposited, where individual 1 was only present in the location of the recovery of crime stain 1, and individual 2 only in the location of the recovery of crime stain 2, might justify an assumption of $\lambda = 1$. Other than these very particular circumstances, it is difficult to imagine a scenario where such an unmitigated assumption could be made.

To justify an assumption of $\tau = 1$, the circumstances must be just as particular. In this case, they must be such that they make it impossible for the suspect to be the source of only crime stain 2. This could be the case when it is physically impossible for the suspect to have been in contact with the surface of crime stain 2. However, even in these cases it is difficult to justify $\tau = 1$ for DNA traces in situations where secondary transfer is possible (e.g., Goray et al. 2010).

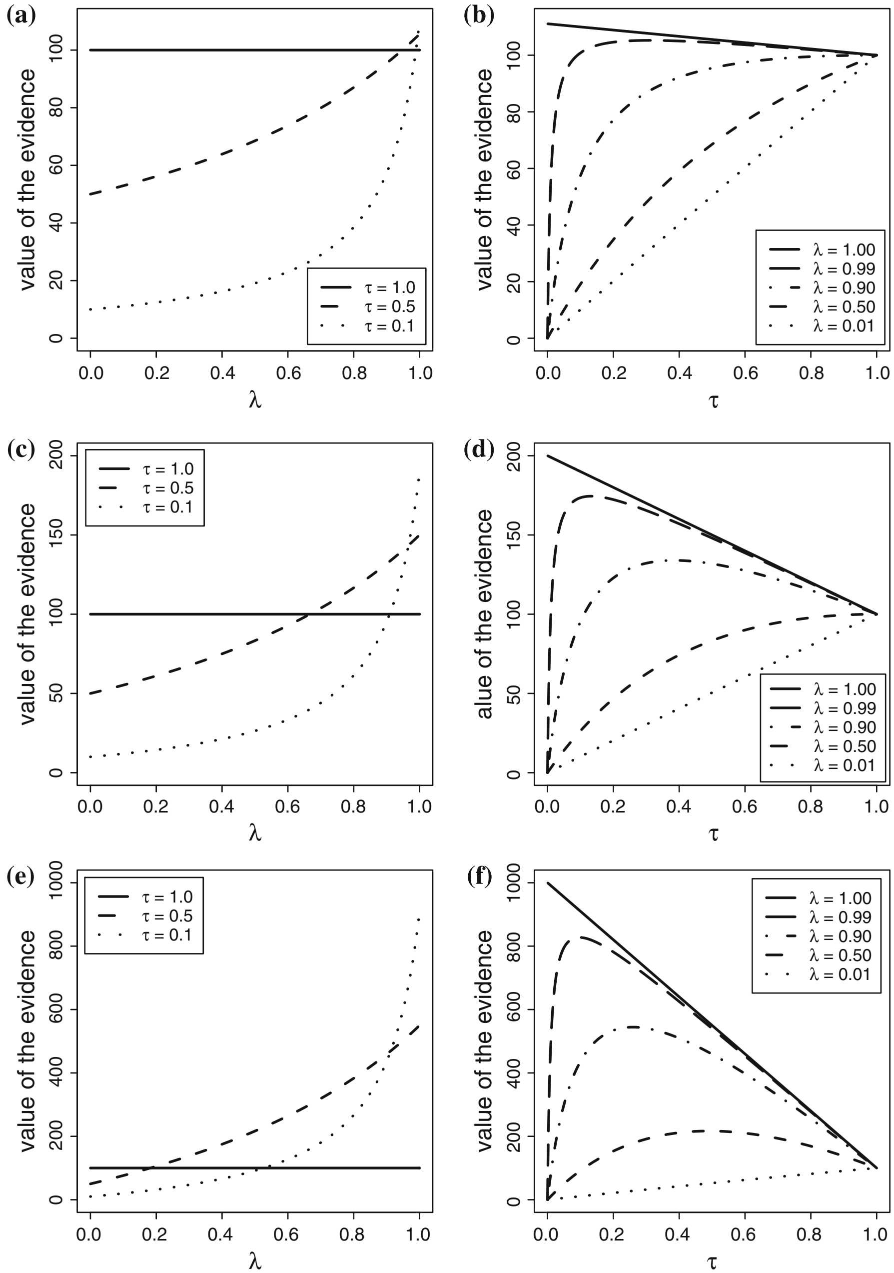

Fig. 8 The value of the evidence V for pair H_1'' in function of: **a** λ for $\delta = 0.1$, **b** τ for $\delta = 0.1$, **c** λ for $\delta = 0.5$, **d** τ for $\delta = 0.5$, **e** λ for $\delta = 0.9$, and **f** τ for $\delta = 0.9$. Here, $\gamma_1 = 0.01$. V is an increasing function of λ and δ, and equal to $\frac{1}{\gamma_1}$ for $\tau = 1$. It tends towards a maximum of $\frac{1}{\gamma_1(1-\delta)}$ for $\lambda = 1$ and $\tau \to 0, \tau \neq 0$

In most cases, V will therefore be less than $\frac{1}{\gamma_1}$. For $\tau = 0.5$, the exact value will lie somewhere on the dashed curve of Fig. 6a below the maximum point at $\lambda = 1$.

According to Fig. 6a, the range of values obtained for V for different values of λ is smaller for high values of τ. This is because a large value of τ leads to a high prior probability that trace 1 comes from the suspect, regardless of whether there was one donor or two donors. That is, if the suspect was the only donor, then it is certain that trace 1 comes from the suspect, and if the suspect was one of two donors, then the prior probability that trace 1 comes from the suspect (=τ) is also high. Therefore a large value for τ leads to a high probability in the numerator of V (Eq. 19), regardless of the value of λ.

This is no longer the case for small values of τ. If τ is small, the prior probability that trace 1 comes from the suspect will be determined mostly by the probability that both traces come from the suspect, i.e., $1 - \lambda$. The numerator of V (Eq. 19) will therefore vary greatly according to the value of λ. The greater λ, the smaller the prior probability of a single donor. Since the evidence is such that it rejects the hypothesis of a single donor, the probability of observing the evidence given proposition C_1 (i.e., the numerator of V) is greater when the prior probability of a single donor is small. That is, a small prior probability for a single donor increases the normalized probability of the event that only crime stain 1 comes from the suspect, figuring in the numerator of V. Thus, the overall value of the evidence is an increasing function of λ.

Figure 6b shows that the range of values obtained for V for different values of τ remains 0 to $\frac{1}{\gamma_1}$, regardless of the value of λ. This is because the evidence (a match with crime stain 1 and a non-match with crime stain 2) is such that its value will always be 0 in a case where it is impossible for crime stain 1 to come from the suspect, given that there were two different donors (i.e., when $\tau = 0$), and equal to $\frac{1}{\gamma_1}$ whenever it is certain that crime stain 1 comes from the suspect, given that there were two different donors (i.e., when $\tau = 1$). Thus, the value of the evidence is an increasing function of τ.

7.3 The value of the evidence for pair H_1''

According to Fig. 7, the value of the evidence for pair H_1'' is equal to

$$V = \frac{0.00857}{9.86 \times 10^{-5}} = 86.9,$$

which is computed by

$$V = \frac{\frac{\lambda\tau}{1-\lambda+\lambda\tau}\gamma_2}{\lambda\gamma_1 \frac{1-\delta}{(1-\tau)\lambda\delta+1-\delta}\gamma_2} \tag{20}$$

$$= \frac{\frac{\tau}{1-\lambda+\lambda\tau}}{\gamma_1 \frac{1-\delta}{(1-\tau)\lambda\delta+1-\delta}} \tag{21}$$

Here, the probability in the numerator is the same as for pair H_1'. The probability in the denominator is the probability of observing the evidence given that crime stain 1 does not come from the suspect (proposition $\bar{C}_1$). A priori, there are two possibilities if proposition $\bar{C}_1$ is true: either the suspect is only the source of crime stain 2 [for which the probability is $(1 - \tau)\lambda\delta$], or neither of the two traces comes from the suspect (for which the probability is $1 - \delta$). The observation of the evidence is only possible in the latter case. Therefore, the probability of the evidence is the probability that the two traces come from two different donors, of which one has trait Γ_1 and the other trait Γ_2, i.e., $\lambda\gamma_1\gamma_2$, times the normalized probability that neither of the traces comes from the suspect (Eq. 20).

For this pair of propositions, V is a function of τ, λ and δ (Fig. 8). Just like for pair H_1', V is equal to $\frac{1}{\gamma_1}$ whenever $\tau = 1$. In this case, the numerator of V reduces to $\lambda\gamma_2$ as explained above for pair H_1', and the denominator of V becomes equal to $\lambda\gamma_1\gamma_2$, because the possibility of the suspect being the source of crime stain 2 when there are two different crimes stain donors becomes impossible. With the numerator, the denominator of V therefore reduces to γ_1.

Yet, unlike for pair H_1', $\lambda = 1$ no longer produces $V = \frac{1}{\gamma_1}$ (e.g., Fig. 8a, c, e). This is because $\lambda = 1$ (i.e., there were two different donors) does not, a priori, exclude the possibility that the suspect is the source of the second trace given that crime stain 1 does not come from the suspect (proposition $\bar{C}_1$). For $\lambda = 1$, V is actually a decreasing function of τ, attaining a minimum of $\frac{1}{\gamma_1}$ when $\tau = 1$ (Fig. 8b, d, f). This is because the possibility of the suspect being only the source of crime stain 2 becomes less probable as τ increases, thus increasing the normalized probability of the event that neither of the traces comes from the suspect. This increases the denominator of V, and decreases the whole value of the evidence. However, when $\tau \to 0$, $\tau \neq 0$, the probability of the suspect being only the source of crime stain 2 increases, which decreases the normalized probability of neither trace coming from the suspect. This decreases the denominator of V, and increases the whole value of the evidence. Thus the maximum of V for this pair of propositions is *greater* than $\frac{1}{\gamma_1}$ (which is the maximum value for the other two pairs of propositions):

$$\text{when} \quad \tau \to 0 \quad \text{and} \quad \lambda = 1, \qquad V \to \frac{1}{\gamma_1(1-\delta)}. \tag{22}$$

In other words, for $\tau < 1$, the possibility that the suspect is the source of crime stain 2 is not excluded. Yet, if the suspect is not the source of crime stain 1 (proposition $\bar{C}_1$), the evidence is only possible when neither stain comes from the suspect (probability of $1 - \delta$), so that the factor $1 - \delta$ has an increasing influence in the denominator of V for $\tau \to 0$, $\tau \neq 0$.

This effect becomes more pronounced as δ increases (Fig. 8b, d, f). A larger value of δ produces a smaller probability in the denominator of V, and therefore a greater value of V.

The value of the evidence proposed by Meester and Sjerps (2003) for this pair of propositions (Eq. 7) is equal to

$$\frac{1}{\gamma_1 \frac{1-\delta}{\frac{1}{2}\delta+1-\delta}}.$$

Again, this value assumes $\lambda = 1$ and $\tau = 0.5$. Its application is therefore just as limited by the assumption $\lambda = 1$ as the value of the evidence they propose for pair H_1' (see page 24). This value is the point at $\tau = 0.5$ on the solid lines in Fig. 8b, d, f. With $\lambda < 1$, V would be smaller, lying on one of the other curves in these graphs.

7.4 Comparison of the values of the evidence

As Meester and Sjerps (2003) concluded, pairs of differently formulated propositions for the two-trace problem lead to different values of the evidence. For the example presented, the value of the evidence is greatest for pair H_1'', and smallest for pair H. This is because the probability of observing a match with stain 1 and a non-match with stain 2 is greatest given proposition C_1 and smallest given proposition $\bar{C}_1$.

The derived formulae for calculating the value of the evidence show that this value is a function of τ for all three pairs of propositions, a function of λ for two of the three pairs (pairs H_1' and H_1''), and a function of δ for one pair (pair H_1''). In the two-trace problem, the value of the evidence is therefore not based solely on the analytical results provided by the laboratory analyses of the collected evidence, that is, on the match probabilities of these results in the relevant population of possible sources. In addition, the value depends on parameters assessed on the basis of the case circumstances prior to observing the evidence. The more specific the competing pair of propositions are, the more parameters will determine the value of the evidence for these propositions. That is, propositions focusing only on one of the two traces require additional information regarding the total number of donors on the crime scene and/or the prior assumption on the suspect's implication as a donor of any of the traces on the scene. To accurately evaluate the value of the evidence in a two-trace problem, an evaluator's knowledge must therefore extend beyond the observations made on the evidence, to the facts regarding the case circumstances.

8 Discussion and conclusions

The role of the forensic scientist is to evaluate the value of the evidence (e.g., Evett 1998). In the forensic two-trace problem, this has been somewhat perplexing since three different formulations of the competing pair of propositions lead to three different quantifications of this value (Meester and Sjerps 2003).

In this paper, we have provided a more general vision of the entire two-trace problem by constructing a Bayesian network that includes each of the three pairs of propositions as a separate node in the model. Through an illustrative example, we demonstrate how to use the network to evaluate the value of the evidence for each pair of propositions. The different structural relationships between each of the pairs and the evidence inevitably leads to different values of the evidence, each addressing the two-trace problem from a different angle.

The flexibility of the value of the evidence to adapt to each pair of propositions is an advantage, not an inconvenience. A forensic scientist's task is to evaluate the relative support provided by the evidence for one proposition with respect to an alternative proposition (i.e., the value of the evidence) (e.g., Evett 1998). And this, he/she must do with regard to the very particular framework of circumstances that reflects the case, and for the precise propositions of interest to the court. Therefore, it is important that the propositions be chosen and formulated with care, and that these be based on the particular circumstances related to the case. The different values of the evidence then complement each other, providing the scientist with a range of formulae from which he/she can select the most appropriate in view of the pair of propositions of interest to the court. The crucial issue is to understand what assumptions lie behind each formula, in order to correctly use it in the context of the case. In this respect, the Bayesian network offers transparency through its graphical representation of the dependence relationships among the variables. In particular, it models the dependency of each of the random variables on three unknown parameters:

- δ, the probability that at least one trace comes from the suspect,
- λ, the probability that the two traces come from two different donors (Dawid 2004), and
- τ, the probability that trace 1 comes from the suspect in a case where the suspect is one of two different donors.

The value of the evidence is a function of τ for all three pairs of propositions, a function of λ for the two pairs where the prosecutor's proposition relates only to one trace, and [as presented in Meester and Sjerps (2003)] a function of δ for the pair where the defence's proposition relates only to one trace. To accurately evaluate the value of the evidence, an evaluator is therefore obliged to have information on the case circumstances. If it is difficult to obtain precise assessments for the unknown parameters, the Bayesian network environment allows the user to specify subjective probability distributions over each parameter space.

Note that the model presented in this paper is still based on several assumptions, notably on the independence between the three unknown parameters. The validity of this assumption will depend on the circumstantial information available in a case, and on the evaluator's personal assessments of the parameters. The results of this work justify a careful examination and further study on the dependence relationships between these parameters in cases where the assumption of independence no longer holds.

Notwithstanding, the major advantage of using the Bayesian network is when the evidence of the two traces must be combined with other types of evidence. The fundamental structure of this Bayesian network allows for an extension to more than two traces, as well as an extension to address activity level propositions (Gittelson et al. 2012). Thanks to its graphical architecture, this model can be inserted as a component part in a larger network for a more complex inference problem. Given that most forensic cases involve numerous traces of different types of evidence, this possibility is an indispensable property for all practical applications. The generic Bayesian network presented in this paper therefore offers a transparent and practical tool for tackling two-trace problems in forensic casework.

Acknowledgments This research was supported by the Swiss National Science Foundation grant n° 100014-122601/1.

References

Aitken CGG, Gammerman A (1989) Probabilistic reasoning in evidential assessment. J Forensic Sci Soc 29:303–316

Aitken CGG, Taroni F (2004) Statistics and the evaluation of evidence for forensic scientists, 2nd ed. Wiley, Chichester

Biedermann A, Taroni F, Delemont O, Semadeni C, Davison A (2005a) The evaluation of evidence in the forensic investigation of fire incidents (Part I): an approach using Bayesian networks. Forensic Sci Int 147:49–57

Biedermann A, Taroni F, Delemont O, Semadeni C, Davison A (2005b) The evaluation of evidence in the forensic investigation of fire incidents (Part II): practical examples of the use of Bayesian networks. Forensic Sci Int 147:59–69

Biedermann A, Taroni F (2006) A probabilistic approach to the joint evaluation of firearm evidence and gunshot residues. Forensic Sci Int 163:18–33

Biedermann A, Bozza S, Taroni F (2009) Probabilistic evidential assessment of gunshot residue particle evidence (Part I): likelihood ratio calculation and case pre-assessment using Bayesian networks. Forensic Sci Int 191:24–35

Cook R, Evett IW, Jackson G, Jones P, Lambert J (1998) A hierarchy of propositions: deciding which level to address in casework. Sci Just 38:231–239

Dawid AP (2004) Which likelihood ratio? Comment on 'Why the effect of prior odds should accompany the likelihood ratio when reporting DNA evidence' by R. Meester and M. Sjerps. Law Prob Risk 3:65–71

Dawid AP, Mortera J, Pascali V, van Boxel D (2002) Probabilistic expert systems for forensic inference from genetic markers. Scand J Stat 29:577–595

Dawid AP, Mortera J, Vicard P (2007) Object-oriented Bayesian networks for complex forensic DNA profiling problems. Forensic Sci Int 169:195–205

Edwards W (1991) Influence diagrams, Bayesian imperialism, and the *Collins* case: an appeal to reason. Cardozo Law Rev 13:1025–1079

Evett IW (1987) On meaningful questions: a two-trace transfer problem. J Forensic Sci Soc 27:375–381

Evett IW (1998) Towards a uniform framework for reporting opinions in forensic science casework. Sci Just 38:198–202

Fenton NE, Neil M (2011) Avoiding probabilistic reasoning fallacies in legal practice using Bayesian networks. Aust J Leg Philos 36:114–151

Fenton NE, Neil M, Lagnado D (2011) Modelling mutually exclusive causes in Bayesian networks. Working paper. http://www.eecs.qmul.ac.uk/~norman/papers/mutual_IEEE_format_version.pdf. Accessed 8 Oct 2012

Garbolino P (2001) Explaining relevance. Cardozo Law Rev 22:1503–1521

Garbolino P, Taroni F (2002) Evaluation of scientific evidence using Bayesian networks. Forensic Sci Int 125:149–155

Gittelson S, Biedermann A, Bozza S, Taroni F (2012) Bayesian networks and the value of the evidence for the forensic two-trace transfer problem. J Forensic Sci 57:1199–1216

Goray M, Eken E, Mitchell RJ, van Oorschot RAH (2010) Secondary DNA transfer of biological substances under varying test conditions. Forensic Sci Int Genet 4:62–67

Jensen F (2001) Bayesian networks and decision graphs. Springer, New York

Kadane JB, Schum DA (1996) A probabilistic analysis of the Sacco and Vanzetti evidence. Wiley, New York

Kjaerulff UB, Madsen AL (2008) Bayesian networks and influence diagrams: a guide to construction and analysis. Springer, New York

Lindley DV (1977) Probability and the law. Statistician 26:203–220
Lindley DV (2000) The philosophy of statistics. Statistician 49:293–337
Meester R, Sjerps M (2003) The evidential value in the DNA database search controversy and the two-stain problem. Biometrics 59:727–732
Meester R, Sjerps M (2004a) Why the effect of prior odds should accompany the likelihood ratio when reporting DNA evidence. Law Prob Risk 3:51–62
Meester R, Sjerps M (2004b) Response to Dawid, Balding, Triggs and Buckleton (concerning: why the effect of prior odds should accompany the likelihood ratio when reporting DNA evidence. Law, Prob. and Risk 3:51–62). Law Prob Risk 3:83–86
Mortera J, Dawid AP, Lauritzen S (2003) Probabilistic expert systems for DNA mixture profiling. Theor Popul Biol 63:191–205
Robertson B, Vignaux GA (1995) Interpreting evidence. Wiley, Chichester
Schum DA (1994) The evidential foundations of probabilistic reasoning. Wiley, New York
Taroni F, Aitken CGG, Garbolino P, Biedermann A (2006) Bayesian networks and probabilistic inference in forensic science. Wiley, Chichester
Thompson WC, Taroni F, Aitken CGG (2003) How the probability of a false positive affects the value of DNA evidence. J Forensic Sci 48:47–54
Triggs C, Buckleton JS (2003) The two trace transfer problem re-examined. Sci Just 43:127–134
Weir BS (2000) Statistical analysis. In: Siegel J, Saukko P, Knupfer G (eds) Encyclopedia of forensic science. Academic Press, San Diego, pp 545–550