A Response to Harel and Porat: On the Inadmissibility of the Aggregated Probabilities Principle

Doron Menashe, Dr.
A Response to Harel and Porat: On the Inadmissibility of the Aggregated Probabilities Principle

Doron Menashe

A. Introduction

This article is a response to an article by Harel and Porat (hereafter: the authors), recently published in the Minnesota Law Review.¹ In the article, the authors argue that, under certain conditions, courts should be permitted to convict a defendant in an unspecified offense. This possibility is meant to address situations in which there is no reasonable doubt that the defendant committed an offense, even though the prosecution failed to prove beyond reasonable doubt that the defendant committed any specific offense of which he was accused.

The authors suggest that this new decision principle be called the Aggregated Probabilities Principle – APP. Under this principle, the court would examine all of the charges brought against the defendant in aggregate, and decides whether there is any reasonable doubt that the defendant committed at least one offense from among those for which he was accused, even if it is impossible to specify which specific offense he committed. The authors distinguish this principle from the prevailing Distinct Probabilities Principle – DPP – by which the court examines whether each individual charge has been proven beyond reasonable doubt.

The authors advance a number of considerations in support of APP. They argue that APP would clearly advance the principle of retributive justice, and would better serve the social objectives of criminal law: deterrence, cost-effective enforcement, and minimizing adjudication errors.

In the course of the article, the authors distinguish APP from the similar-crimes doctrine and the prior-acts doctrine, while providing criteria for deciding when to apply the various doctrines. While doing so, they also suggest principles that, in their view, would resolve the problem of probabilistic interdependence between offenses.

The authors also present situations in which APP is consistent with expressivist theories of punishment, which the authors deem to be the theories that underlie the present reticence of the law to adopt APP.2

In this response, I intend to demonstrate the inadmissibility of the new decision principle advanced by the authors. My argument shall proceed as follows: First, in Part B, I shall point out that, as opposed to the authors’ claim, APP can coexist only with the mathematical probability model of legal proof and of the “beyond-reasonable-doubt” standard of proof. That will refute the authors’ claim that their model is consistent with the inductive probability model.

---

2 In this context, I would note that I am somewhat skeptical of this claim. Some of the retributive approaches emphasize the conduct that is being punished. Such narrow retributivism emphasizes that the perpetrator must be punished because he committed the offense. Thus, for example, Michael Moore writes:

“to achieve retributive justice, the punishment must be inflicted because the offender did the offence. To the extent that someone is punished for reasons other than that he deserves to be punished, retributive justice is not achieved.” Michael S. Moore, Placing Blame: A Theory of Criminal Law 28 (1997).

Like the doubt that arises as to the willingness of expressivist approaches (cf. p. 306) to accept condemnation for all offences without any connection to a specific offense, it is also questionable whether narrow retributivist theories would deem punishment imposed for an undefined offense as meeting the requirements of retributive theory.
of L. Jonathan Cohen or with other critics of the mathematical model of the standard of proof.

Having established that the authors’ are bound to the mathematical model, I will show, in Part C, that such reliance upon the mathematical probability model prevents the practical application of APP. This is because, as a practical matter, there is almost always probabilistic interdependence between the various charges. The use of APP would require that courts assign conditional probability to the various charges. With all due respect, it is hard to imagine that the courts possess the necessary skill, particularly if precise analysis is required, but even if only broad, intuitive estimates of probability were required.

In this context, I will show that the principles that the authors set out for contending with probabilistic interdependence not only fail to contend with it, but also reveal a profound misunderstanding of the concept of probabilistic interdependence and of the concept of non-interdependent events. As we shall see, my disagreement with the authors in this regard is not founded upon opinions, estimates and assumptions, but rather upon a demonstrable mathematical error made by the authors.

In Part D, I will examine several claims made by the authors in support of APP. I will show that it is highly doubtful that APP would lead to minimizing adjudication errors or result in more cost-effective enforcement. Moreover, not only would APP not promote equality, it would place us in a situation of having to decide between infringing equality or harming deterrence. I will also address the ad hoc nature of the solutions that the authors offer for resolving the problems raised by the principle they propose.

B. The theory of proof grounding the APP method
The authors present APP, and distinguish it from DPP, against the background of the following example:\(^3\)

**Example 1:** A person is charged with pickpocketing and rape, two unrelated offenses allegedly committed by him at different times and places. The evidence suggests that the probability that he committed each one of these offenses is .9. Assume that the required probability necessary to satisfy the beyond-a-reasonable-doubt standard is .95. Should the court convict the defendant on either of the offenses?

In regard to this example, the authors argue that if we were to follow the prevailing principle – DPP – according to which each offense must be examined independently, then it would appear that the defendant must be acquitted of both offenses, inasmuch as neither was proven to the extent required by the standard of proof (because 0.95>0.9). As opposed to this, were we to calculate the probability that the defendant perpetrated at least one of the offenses (in accordance with APP), the probability of guilt would rise to 0.99, which lies above the required standard of proof (it should be noted that this calculation is correct only if – and, as we shall see, it is a big “if” – we assume an absence of probabilistic interdependence between the charge of pickpocketing and the charge of rape). In this case, APP would suggest that the defendant be convicted of a non-specific offense, and sentenced in accordance with the less serious crime.

As can be seen, the above example, like the other examples brought by the authors, employs Pascalian mathematical probabilities, including the product rule, according to which the probability of the occurrence of independent events (as we shall see, this is an important proviso) is a product of their probabilities.

Nevertheless, the authors are careful to point out that APP can be made to accord with the approach of those who reject the mathematical model of legal proof, and the beyond-reasonable-doubt standard of proof. In particular, the authors point to the positions of the

---

\(^3\) P. 262 of the article, n.1.
members of the traditional opposition, Laurence H. Tribe and L. Jonathan Cohen. But can APP actually be harmonized with the views that reject the mathematical model of the standard of proof?

In my opinion, the answer is a resounding no. In order to understand why, we must take a closer look at the two assumptions regarding the standard of proof that allow the authors to make their distinction between APP and DPP: 1) When a person is charged with a number of offenses, the probability of guilt in regard to at least one offense is equal to or greater than the probability of guilt for each of the other offenses individually (this is a result of the product rule); 2) The mathematical-probabilistic significance of a beyond-reasonable-doubt standard of proof is constant, and it does not change from case to case (e.g., in the example presented by the authors, 0.95).

It is important to note that the second assumption is necessary. If, as opposed to that, we were to express the beyond-reasonable-doubt standard of proof in such a way that its mathematical-probabilistic significance would change from case to case, then the distinction between APP and DPP could not be maintained. This can be demonstrated simply by means of the authors’ Example 1. In that example, the probability of the commission of at least one of the offenses is 0.99. It therefore exceeds the required standard of proof, viz. 0.95, even though the probability of guilt for each of the offences is less than 0.9. However, if the probabilistic significance of the standard of proof were to change in accordance with the circumstances of each case, such that the probabilistic significance for proving pickpocketing and rape individually would be proof beyond 0.95, but the probabilistic significance in regard to the question of whether at least one offence had been committed would be proof in excess of 0.995, then APP and DPP would not necessarily yield different results.

Critics of the Pascalian mathematical model of the standard of proof, particularly those the authors refer to by name – Tribe and Cohen – specifically reject assumption 2, i.e., the

4 P. 267 of the article, n. 17.
assumption that the mathematical-probabilistic significance of the standard of proof is constant (moreover, as we shall see further on, even the product rule, which grounds the first assumption, does not hold in the proof model presented by Cohen).

Thus, for example, Tribe expressly rejects quantifying the standard of proof, inasmuch as it requires a predetermination of an agreed level of probability for errors in convictions. Tribe distinguishes between agreeing in advance to a certain level of wrongful convictions, and being reconciled to such errors after the fact as being unavoidable despite our best efforts to prevent them. The result, in mathematical-probabilistic terms – or if one prefers, the probabilistic significance – of a standard of proof based upon exhausting the means for preventing errors cannot be constant and unchanging for every case. On the contrary, no additional effort is being made to ascertain the defendant’s innocence by the mere fact that we examine the two offenses together. In other words, if there is an additional avenue that was not pursued in regard to the charge of rape (e.g., an alibi claimed by the defendant), that avenue will remain unchecked even if we examine the offenses of rape and pickpocketing together.

Similarly, the authors’ claim that APP can be reconciled with Cohen’s theory seems problematic, to put it mildly. To help to understand this, I will briefly present the basics of Cohen’s theory. One of the central characteristics of Cohen’s theory is refraining from any mathematical-numerical quantification of the standards of proof. Cohen’s objection is not to formal quantification, and it does not derive from the practical difficulty in achieving it. His objection is to quantification itself, whether formal or intuitive, and whether precise or approximate. In regard to criminal law, Cohen based his approach upon the idea that the standard of proof is that of absolute certainty, in terms of inductive probability. This logical approach is expressed in the elimination of all possible alternatives to explain the finding, and in proving the final conclusion to be the only reasonable possibility. In the present context, this approach says that proof “beyond reasonable doubt” does not indicate a standard of measurement for the strength of the evidence, but rather is examined and established by elimination – by subjecting the evidence to refutation. The evaluation of the
evidence is thus the result of a certain thought process that is based upon eliminating doubts by refuting conjectures.\(^5\)

In that light, it becomes clear that Cohen’s approach cannot be reconciled with APP. The reason is simple: There are independent exculpating scenarios for every offense. Under the assumption of non-interdependency of the events – as assumed by the authors (and which will be examined below) – the addition of another offense does not eliminate the exculpating scenarios of the earlier offense. Such an addition only lowers the mathematical probability of the combination, but it does not cancel out the addition of the separate exculpatory scenarios. The only circumstances in which one might imagine such a situation in reliance upon Cohen is one in which, for example, the alibi for one offense is the very perpetration of the second offense, or more generally and precisely: where the only exculpating scenario for a particular offense is the perpetration of another offense. This is not the general case addressed by the authors (who, as noted, assume interdependence), and in any event, such situations would be rare.

Even more surprisingly in this regard, the authors would appear to be aware of the fact that Cohen does not believe that the problem of conjunction in civil-cumulative cases and criminal-cumulative cases applies to the inductive probability approach that he proposed, inasmuch as the product rule does not hold in that system. Indeed, as the authors note,\(^6\) the conjunction problem is one of Cohen’s arguments against the validity of Pascalian mathematical probability in legal contexts, and one of the arguments he gives in favor of inductive probability as a descriptive model for legal proof. It thus seems amazing that the authors imagine that Cohen’s approach might be squared with the claim that the probability

\(^5\) The model proposed by Cohen is based upon the foundations of modal logic, as opposed to the axiomatic evaluation of set theory, which grounds Pascalian probability theory, particularly that of Bayes. The system advocated by Cohen as an alternative to mathematical probability locates inductive probabilities on an axis beginning with “no-proof of X” and terminating with “proof of X”.

\(^6\) N. 38 of the article.
that the defendant perpetrated at least one offense increases only because of the addition of more charges, which is the very same aspect – the product rule – that Cohen rejects.

Perhaps the following statement by the authors might appear to maintain their claim:

…the APP is about substantive criminal law rather than fact-finding. This Article focuses on the commission of an unspecified offense and does not discuss how to establish whether any particular offense has been committed.

But this proviso cannot save the day. While the APP doctrine that the authors present comprises a substantive-law aspect – viz., the fact that APP makes it possible to convict for an unspecified offense – in order to prove the existence of that unspecified offense, the authors must make recourse to the product rule, which, as earlier noted, is inconsistent with Cohen’s approach.

C. The problem of interdependence and a critique of the authors’ approach thereto

As we have seen, APP is committed to the Pascalian mathematical model of legal proof, and to the beyond-reasonable-doubt standard of proof. But is it practicable even in that framework? In my view, the answer is no. In brief, before presenting the argument in detail, I am of the opinion that, as opposed to the position adopted by the authors: 1) practically speaking, the problem of interdependency is always present, even where we are concerned with completely different offenses (like rape and pickpocketing); 2) in cases of interdependence, the product rule cannot be employed, and recourse must be made to the conditional probabilities formula, which, even according to the authors, would make using APP too complicated.7

7 P. 284 of the article.
I will now set out my arguments in detail. First, I will explain the concept of conditional probability, and show how it affects our subject.

As may be recalled, the authors employ the mathematical formula according to which the product of two independent events is multiple of their probability (or more precisely, independent events are defined as events for which the following formula is true):

\[ P(A \cap B) = P(A) \cdot P(B) \]

Thus, for example, in Example 1 the authors calculate the probability that the defendant did not commit any of the offenses by multiplying the probability that he did not commit the offense of pickpocketing by the probability that he did not commit the offense of rape.\(^8\)

The problem is that throughout their article, the authors ignore the fact that the said equation is valid only in regard to independent events. As opposed to this, in the case of interdependent events, the following, more general equation must be employed:

\[ P(A \cap B) = P(A / B) \cdot P(B) = P(B / A) \cdot P(A) \]

In order to demonstrate the difference, I will put aside the pickpocketing/rape example – to which I shall return in order to refute the authors’ claim that those events are not interdependent – in favor of the following example, which the authors present later in the article:

**Example 5 - Differing Interdependent Doubts:**

---

\(^8\) This is how they explain the computation in n. 3 of the article: “The probability that the defendant committed each one of the offenses is .9, and therefore the probability, for each one, that he did not commit the offense is \(1 - .9 = .1\). Consequently, the probability that he did not commit any offense is (.1)^2 = .01, and the probability that he committed at least one of the offenses is \(1 - .01 = .99\).”
A person is charged with two offenses of sexual assault allegedly committed by him at different times and places against two different victims. Examined separately, the evidence in each case indicates a probability of .7 that the defendant committed each one of the offenses. When the pattern-of-behavior doctrines are applied, this probability increases to .9. The reason the court is not persuaded that the defendant committed each of the offenses is that, in each case, there is a different lone eyewitness whose reliability is questionable. Assume that the probability required for the beyond-a-reasonable-doubt standard in the legal system is .95. Should the court convict the defendant of any of the offenses?²⁹

This is a clear example of interdependent events, even in by the authors’ criteria, inasmuch as the probability of the second offense increases given the evidence for the existence of the first offense.

Nevertheless, surprisingly, the authors find no difficulty in applying APP to this example, by means of a simple multiplication of probabilities, as long as it is not applied in tandem with the pattern-of-behavior doctrine (the authors include both the prior-acts doctrine and the similar-crimes doctrine under this rubric). They even perform the calculation by multiplying the probability that the defendant did not commit the first offense (1-0.7=0.3) by the probability that he did not commit the second (again, 1-0.7=0.3), and thus conclude that the probability that the defendant did not commit both offences is 0.09 (0.3*0.3), and therefore the probability that he committed at least one is 0.91 (1-0.09).

From the brief explanation presented above, the reader should already understand that we are confronted with a mathematical error. In the case of interdependent events, we cannot multiply the probability of the two events in order to determine the probability that both occurred. Rather, we must make recourse to conditional probabilities in accordance with the formula:

\[ P(A \cap B) = P(A \mid B) \cdot P(B) = P(B \mid A) \cdot P(A) \]

²⁹ P. 284 of the article.
This formula must be employed even if the pattern-of-behavior doctrine is not used in parallel.

The reason for this should be clear even intuitively, inasmuch as the probability that the defendant committed the second rape increases given the fact that he committed the first rape, so clearly, given that the defendant did not commit the first rape, the probability that he committed the second must be smaller than 0.7. Indeed, the conclusion must be that one

\[ P(A) = P(A \cap B) + P(A \cap B_C) \rightarrow P(A) = P(A/B) \cdot P(B) + P(A/B_C) \cdot P(B_C) \]

from which it derives that the conditional probability of A, given not-B (in other words $B_C$), is:

\[ P(A/B_C) = \frac{P(A) - P(A/B) \cdot P(B)}{1 - P(B)} \]

We assumed that the conditional probability of A given B is greater than the probability of A. From this assumption we can derive the conclusion that the conditional probability of A given not-B is smaller than the probability of A, thus:

\[ \left[ P(A/B) > P(A) \right] \rightarrow \left[ -P(A/B) < -P(A) \right] \rightarrow \left[ -P(A/B) < -\frac{P(A) \cdot P(B)}{P(B)} \right] \rightarrow \left[ -P(A/B) < \frac{P(A) \cdot P(B)}{P(B)} \right] \rightarrow \left[ P(A) - P(A/B) \cdot P(B) < P(A) \right] \rightarrow \]

\[ P(A/B_C) < P(A) \]

Q.E.D.

To the authors’ credit, it should be noted that we cannot necessarily know from the aforesaid whether or not $P(A/B_C)$ is substantially smaller than $P(A)$. It may, therefore be that in certain cases $P(A)$ is a reasonable approximation of $P(A/B_C)$. But even if we could know in a given case that it was a reasonable approximation, the authors do not limit themselves to such cases. The condition that the authors impose is whether the doubts are similar or dissimilar. But that condition is unrelated to the question of whether we have a reasonable approximation or not (which is the only important question in this regard). There may be similar doubts for which the said approximation will be reasonable, and there may be cases of dissimilar doubts for which the approximation will not be reasonable (for example, let us assume that studies show that the probability that a pedophile will commit only one offense and will immediately be apprehended approaches zero. Following a comprehensive examination of all of the opportunities that defendant A had to commit offences of pedophilia, only two instances were found in which it is possible that he committed an offence with a probability of 0.7 for each instance. It is nearly certain that there were no other instances in which the defendant had an opportunity to commit offences of pedophilia. The doubt in the two cases is that there is a questionable witness in both cases. According to the authors’ definition, we have two separate doubts, just as in Example 5, in which they claim that APP can be applied as long as there is no parallel use of the pattern-of-behavior doctrine. But clearly in this case the probability that defendant A committed just one of the acts approaches zero – in other words, the probability for one of the offences, given innocence in regard to the second, approaches zero. In this case, it is therefore clear that $P(A)$ is not a reasonable approximation of $P(A/B_C)$. 
may not use the product of independent events, unrelated to the question of whether or not there was parallel use of the pattern-of-behavior doctrine.

What the above means is that the authors were mistaken in concluding that “interdependence may (although need not be) a reason to apply the APP”. Quite the opposite: *interdependence among the offenses always* rules out the possibility of using the product rule. If the authors wish to apply APP to such situations, they must use the conditional probabilities formula, but the authors agree that such a course would be too complicated from a practical point of view.\(^\text{11}\)

It should also be clear that wherever it is possible to employ the pattern-of-behavior doctrine (which, by definition, is based upon this assumption of interdependence), it is absolutely impossible to use a simple multiplication of probabilities, and resort must be made to the conditional probabilities formula (even if there is no parallel use). This fact renders superfluous the authors’ categorization of the cases in which parallel or alternative use of APP and the pattern-of-behavior doctrine can be made.

Moreover, a careful examination of the categorization proposed by the authors reveals additional errors. The authors summarize the various cases in which the question of the relationship between the pattern-of-behavior doctrine and APP may arise, as follows:

The following conclusions can be drawn from the above discussion:

1. The APP should not be applied when identical doubts exist with respect to all of the alleged offenses. This is true regardless of whether the pattern-of-behavior doctrines are applied.

\(^{11}\) P. 284, where the authors write: “If one eyewitness in the Example is a liar (or had made a mistake), then not only would acquittal of the relevant offense be justified, but also the probability of the defendant’s guilt of the other offense would decrease from .9 to .7. Thus, even though aggregating the probabilities in such cases is theoretically possible, it is impractical.”
2. The APP should not be applied when the offenses raise differing doubts if the pattern-of-behavior doctrines are applied and the probability of the defendant's guilt on each charge after applying these doctrines depends on his guilt of the other alleged offenses.

3. The APP should be applied when a different doubt arises with respect to each offense and the pattern-of-behavior doctrines are not applied.

4. The APP should be applied when the doubts differ with respect to each offense, even if the pattern-of-behavior doctrines are applied, as long as the respective probabilities of the defendant's guilt in each offense after the doctrines have been applied are not impacted by whether he is guilty of the other offenses.\footnote{Pp. 284-285 of the article.}

Several questions can be raised as to this categorization. I will proceed from the simpler or the more complex:

First, it is not clear why the authors devote separate discussions to the category of identical doubts in the first paragraph, and, in the second paragraph, to the category of differing doubts where the probability of the defendant’s guilt on one charge depends on his guilt of the other offenses. After all, we are concerned with two separate instances of probabilistic interdependence. Indeed, the source of the probabilistic interdependence is different in each case: in the first, the source of the interdependence is in that a shared doubt (e.g., whether the defendant’s enemy accused him of an offense that he did not commit), whereas the second case concerns interdependence deriving from a different source. But the source of the probabilistic interdependence does not change the fact that both are instances of probabilistic interdependence.

This first criticism may initially appear somewhat quibbling, and I would not raise it were it not that it leads to further errors by the authors, which, indeed, it does. Had the authors recognized that they were concerned with two individual instances of the same phenomenon, they would not have thought to argue that APP is never applicable in the first
case, whereas it should be applied in the second (as they write in paragraph 3), providing there is no parallel use of the pattern-of-behavior doctrine.

Another point that should be noted concerns paragraph 4, which permits parallel application of the pattern-of-behavior doctrine and APP in cases in which doubts differ, and the probability of the defendant's guilt in each offense is not affected by whether he is guilty of the other offense. This category of cases appears to suffer from an internal contradiction, inasmuch as if guilt in regard to each offense is not affected by guilt in regard to any other offense, how can the pattern-of-behavior doctrine be applied at all? In order to understand how the authors believe it is possible, let us examine the example they present:

**Example 4 - Independent Doubts:** A person is charged with two offenses of sexual assault allegedly committed by him at different times and places against two different victims. The evidence in each case examined separately indicates a probability of .7 that the defendant is guilty of each alleged offense. Assume now that after applying the pattern-of-behavior doctrines, that probability increases to .9. The reason the court is not fully persuaded that the defendant committed each one of the offenses is because it is not clear that the complainants did not give their consent. The absence of victim consent is a precondition for convicting the defendant under prevailing law. At the same time the court is completely persuaded that the defendant committed all the acts attributed to him and also that he had the requisite intention for committing both sexual assaults. Assume that the required probability for satisfying the beyond-a-reasonable-doubt standard in the legal system is .95. Should the court convict the defendant of any one of the offenses?

Now, if we attempt to uncover the assumptions (which the authors do not explicitly state) that explain how the authors understand the case, we find that they assume: 1) that the actual physical act that would constitute rape, and the defendant’s awareness of the physical act, were proven with *absolute certainty* as a consequence of the pattern-of-behavior doctrine; 2) that the absence of consent in one case does not affect the probability of an absence of consent in the second.
The first assumption is necessary because the two physical acts are interdependent (were it not so, the pattern-of-behavior doctrine could not be applied), and therefore, in each case in which there is no absolute certainty of the commission of both physical acts, the probability that neither was committed must be examined, for which purpose the conditional probabilities formula must be applied (inasmuch as we assumed that the events are interdependent in order to apply the pattern-of-behavior doctrine).

The second assumption is required because otherwise we are concerned with interdependent events, and the product rule cannot be applied.

Both assumptions are highly questionable. The first assumption is questionable because it is not clear when one can arrive at *absolute mathematical certainty* in regard to questions of fact. The second assumption is questionable because how can it be that a lack of consent in one case does not affect the probability of a lack of consent in the second case? This is particularly so if the defendant is required to be aware, or at least negligent, in regard to the absence of consent. Clearly, a lack of consent in one case affects (at least minimally) the probability of a lack of consent in the second case, and thus the cases are interdependent.

Up to this point, I have addressed the second part of my argument in this subsection, *viz*., the claim that “in cases of interdependence, the product rule cannot be employed, and recourse must be made to the conditional probabilities formula, which, even according to the authors, would make using APP too complicated”. I will now proceed to address the first part of the argument, according to which, “practically speaking, the problem of interdependency is always present, even where we are concerned with completely different offenses (like rape and pickpocketing)”.

Is it indeed the case that the problem of interdependence is limited, and that it does not affect the application of APP in cases like that in Example 1, in which the defendant is charged with pickpocketing and rape? The answer is clearly no. There is almost always some degree of probabilistic interdependence. It is hard to imagine a situation in which there is
absolutely no interdependence. After all, at the very least, the offenses have one thing in common: they are connected by the perpetrator, who has a personality and tendencies, etc.\textsuperscript{13}

Consider, for example, the authors’ Example 1, in which the defendant is accused of two completely different offenses: pickpocketing and rape. The authors believe that this example does not involve interdependence,\textsuperscript{14} but in my opinion, this reveals a lack of understanding of the concept of interdependence. Is there no interdependence between the two offenses? Does committing the offense of pickpocketing not raise the probability of the commission of the rape even minimally?

In order to appreciate the unreasonableness of this assumption of an absence of interdependence, even in regard to completely different offenses, consider the following example: Assume that two people, A and B, are standing before you, one of whom definitely committed rape. We have no information about them other than the fact that A has perpetrated an offense of pickpocketing. You have to wager which one you think is most probably the rapist. Now, to say honestly that there is absolutely no interdependence between the offense of rape and that of pickpocketing is equivalent to saying that there is no way to choose between them, except arbitrarily. But I question whether a person can honestly say that he thinks both wagers are absolutely equal, and that he would have to toss a coin in order to decide which wager to prefer. Understandably, most people would say that there is

\textsuperscript{13} Indeed, the only clear case in which one might consider an absence of probabilistic interdependence among the offenses attributed to the same person is in a case in which the defendant lost consciousness after perpetrating the first offense, and when he resumed consciousness, he lacked his awareness and sense of his former personal identity. In any other case, due to the existence of the trait of permanent personality, such a lack of interdependence cannot be assumed. It must further be remembered, as derives from the general principles and degree of proof of criminal law, that any doubt in this area must serve to the defendant’s benefit. Thus, the basic assumption must be that there is strong interdependence, which would appear to mean that in cases of multiple offenses attributed to a single defendant, all that can be established is that the probability that the defendant did not commit any offense is equal to the lowest probability of his innocence of any of the offenses, but that probability still, according to the assumption, leads to acquittal. Thus, even according to the concept of aggregating probabilities according to the (mathematically sound) approach of multiplication of conditional probabilities, there is no basis for convicting a defendant faced with multiple charges, where each charge is proved to a degree slightly less than the criminal standard of proof.

\textsuperscript{14} P. 281 of the article.
no great difference between A and B, but if they were forced to choose, then A would be the better bet, because he has a criminal record. That is equivalent to saying that there is interdependence.

Now, if the probability of perpetrating the rape is even somewhat greater given the fact of the pickpocketing, then, as stated above, given innocence of pickpocketing, the probability of rape decreases. It is therefore clear that the product rule cannot be used, but rather the conditional probabilities formula must be adopted, which will yield a lower probability of guilt than the product rule.

Even if the use of evidence of a criminal tendency or a criminal past cannot be used against the defendant, nevertheless it is clear that such evidence is of probative value. This probative value cannot be ignored if it can aid the defendant.

We thus see that the central assumption upon which the APP proposal rests is the lack of interdependence among different events attributed to the defendant, but it is hard to imagine when such a condition might actually be met. On the contrary, the situations in which the authors see an advantage to APP are those of repeat offenders (“professionals”), which are situations that constitute prime examples of interdependence among the various offenses.

It should be noted that even the practical advantage of APP is questionable in those cases. Precisely because of the interdependence of the offences, we can make use of similar crimes and prior acts, so the defendant will not escape justice.

The proper way to address the rare situation of multiple charges, where each offense is proven slightly below the required minimum standard of proof, is to check the “explanatory variable” of the situation. People are not like balls randomly pulled out of a sack, such that any exceptional combination is merely a product of the fact of a large number of repetitions. It would be more reasonable to conjecture that the personality factor (criminal tendency,

\[15\] P. 290 of the article.
etc.) is the variable, or is the most prominent variable, that explains it (a fact that may lead to a conviction for more than one offense), or alternatively, the variable that explains the result is a tendency toward innocent involvement (a scapegoat of a criminal gang, a non-criminal hanger-on of a group of criminals, a person who has been framed, etc.).

It should be emphasized that I am not arguing that such an examination will always reveal the explanatory variable or variables, but rather that a simple multiplication of probabilities, without performing such an examination, is unjustified, particularly in light of the fact that the phenomenon of multiple offences on the threshold of proof is, itself, rare.

D. Examination of the authors claims in favor of APP

As has been demonstrated up to this point, APP as presented by the authors is fundamentally flawed for non-Pascalian mathematical approaches. Moreover, we have seen that the authors’ treatment of the problem of interdependence is flawed even in regard to Pascalian mathematical approaches, while a correct handling of the problem renders APP impractical.

I will now examine the reasons given by the authors for adopting APP.

Does DPP infringe equality?

The authors argue that the APP method would lead to fewer wrongful convictions, and to fewer acquittals of guilty defendants. In justifying this outcome, the authors point out that APP achieves it simply by correcting a discriminatory element of DPP that favors defendants who are charged with multiple offenses over those accused of committing a single offense. In their own words:
Suppose, now, that a person is accused of four offenses of similar severity and for each, there is a probability of .9 that he is guilty. Adhering to the .95 threshold would mandate convicting the defendant for two offenses. If the legal system acquits a person of all four offenses (as required by the DPP) it will in effect endorse a principle under which it is better to have 9999 guilty people acquitted than a single innocent person convicted. Aside from its evident absurdity, this outcome highlights the discriminatory effect of the DPP as opposed to the APP: upholding the DPP implies an unfair preference for people accused of committing a series of offenses over people accused of committing a single offense. The requisite probability of guilt necessary to convict a defendant in the former case (the case involving a series of offenses) is much higher than what is required in the latter case (the case involving a single offense). A person who has committed an offense beyond a reasonable doubt will be set free simply because the particular offense that he committed could not be proven beyond a reasonable doubt. The APP requires that the presumption of innocence should apply in a nondiscriminatory way.\textsuperscript{16}

With all due respect, this argument begs the question, that is, it assumes that the important question is whether or not the defendant committed at least one offense. Only by proceeding from that assumption can one argue, as do the authors, that DPP favors a person accused of four offenses with a probability of guilt of 0.9 – such that the probability that he committed any offense is 0.9999, and he is not convicted – whereas a person accused of an offense with a probability of 0.95 is convicted, even though the probability that he committed any offense is only 0.95.

The problem, however, is that DPP rejects that very assumption. In other words, the supporters of the DPP approach take the view that the important question is not whether or not the defendant committed any offense, but whether he committed a specific offense. On that basis, there is no discrimination, inasmuch as the probability of being guilty of a specific offense to the degree required for a conviction is the same for both defendants.

\textsuperscript{16} P. 287 of the article.
On the contrary, as we shall see below, in attempts to avoid undermining deterrence, it is APP that appears to infringe equality when it is judged on *its own terms* (i.e., in light of the assumption that the important question is whether or not the defendant committed at least one offence).

**APP may lead to minimizing adjudication errors**

One of the claims made by the authors in favor of APP is that APP may result in minimizing adjudication errors or to a lessening of their cost. But even the authors do not establish this as a certainty, but merely assert it as a possibility. Moreover, examining the “realistic assumptions” grounding this possibility raises serious doubt as to how realistic it really is.¹⁷

One of the “realistic” assumptions is that there is an upper limit to the punishment that the state can impose upon defendants (like a total number of years). Therefore, an increase in the number of convictions as a result of applying APP would leave the state with two options: either lowering the number of years of punishment, or raising the standard of proof, which would lead to fewer wrongful convictions.

The problem is that it is not clear why an assumption in regard to a quantitative limit upon penalties is realistic. It would be one thing if we knew that APP would lead to a *massive* increase in the number of convictions that would create a budgetary crisis. But there is no reason to assume that the number of cases in which APP would bring about a change in the result, and lead to a conviction where there would not have been one otherwise, is so very great.

---

¹⁷ Pp. 288-289 of the article.
APP must harm either deterrence or equality

The authors claim that APP is also preferable due to considerations of deterrence, as the chances that repeat offenders will evade punishment is lessened. This advantage is even greater if we assume that the expected sanction required to deter repeat offenders is greater, and given the fact that repeat offenders are often “professionals”, and therefore more sophisticated. Such criminals are also more responsive to sanctions.\(^{18}\)

Nevertheless, the authors are aware of the fact that consistent application of APP might impair deterrence. They provide the following example of such a situation:

In that Example, a defendant is charged with two offenses and the probability that he committed each offense is .95. In such a case, we suggested that under the APP the person ought to be convicted in one rather than two offenses as the probability he committed both offenses is lower than .95. Yet, if applied to cases involving different trials, the APP could result in the absurd outcome that a person who was convicted in the past is "free" (or, at least freer) to commit a crime with no punishment. For this and other reasons we do not recommend using the APP across different trials.

Unfortunately, the recommended solution – limiting the use of APP across different trials\(^ {19}\) – is problematic, inasmuch as it infringes the principle of equality before the law: there is no real difference between people charged of two offences in one trial and people how are charged of the same offences in two separate trials that are tried in parallel. Similarly, there is no difference between defendants who were convicted in the past and whose past files, according to the APP approach, should not be considered in calculating the aggregate probability, and defendants who are accused of different events in a single trial. Nevertheless, although there are no such differences, in a case in which every charge can be

\(^{18}\) Pp. 289-290 of the article.

\(^{19}\) The authors also justify this limitation by reason of the obstacles to sharing information across trials, as well as for reasons of double jeopardy, see p. 271, n. 30.
proved beyond reasonable doubt, the defendant will be better off if he is charged with all the offenses in a single trial, since the greater the number of charges the lower the aggregate probability that the defendant committed all of the offenses. The decision whether to charge a defendant in a single trial or to file more than one is usually a result of chance, and occasionally of the arbitrary discretion of the prosecution. Such technical considerations should not decide the fate of defendants. Clearly, the authors’ proposal would significantly undermine equality before the law.

Somewhat paradoxically, while the authors seek to show that DPP is discriminatory (and fail by begging the question), it would appear that it is actually APP that infringes equality when it tries to avoid undermining deterrence as a result of the consistent application of APP. It should be emphasized that as opposed to DPP, APP infringes equality when it is judged on its own terms.

Costs of enforcement

The authors try to garner support for adopting APP with the claim that it presents an advantage in terms of cost effectiveness. According to their approach, the marginal cost of every additional piece of evidence is higher than the cost of the items of evidence that precede it. By adopting APP, the police will not have to continue to gather expensive evidence to complete a single file in regard to a specific offense, but can open a second file for which new, cheaper evidence can be obtained for another offense, and both files will then serve together to achieve a single conviction by means of APP.20

This line of argument is somewhat hard to understand. Besides the inconsistency of the authors’ approach, which objects to such manipulative police conduct,21 and the serious

20 P. 291 of the article.

21 Pp. 292-294 of the article.
problem caused by encouraging the police to save resources while achieving unclear convictions for non-specific offenses, the underlying assumption of cost effectiveness is incorrect. The authors claim that, typically, the marginal cost of gathering each additional item of evidence increases. For example, the cost of the tenth piece of evidence is higher than the cost of the ninth piece of evidence, and so on. This would not appear to be a reasonable assumption. It is no secret that there are initial costs involved in opening an investigation. It is by no means certain that the police would prefer opening a new file, bearing in mind the cost involved in pursuing a new avenue of investigation (including time to study the new file, plan it and, of course, carry out the action itself). It is, therefore, not clear that opening a new file would not be more expensive than gathering additional evidence in an existing file, for which the initial costs have already been expended. This is all the more true when we consider that the authors condition the opening of a new file upon achieving a 0.5 probability of guilt. In other words, the new file cannot consist merely of some minor evidence pointing in a new investigative direction, but requires a serious investigation.

I should stress that I am not stating that the cost relationship is actually the reverse of that claimed by the authors. I am merely arguing that the relationship cannot be predicted, since it depends upon the nature of the evidence. Thus, for example, one might think that the cost of interrogating a witness increases (although not necessarily linearly) as the interrogation enters its more advanced stages, due to the need to confront the testimony with the evidence that has been collected. But that is not necessarily the case. Not every witness’ testimony needs to be checked against all the evidence collected in the file, and carefully studying the file at the outset, makes it easier to assimilate evidence acquired at a later stage.

For similar reasons, I do not agree with the author’s claim that adopting APP would lower the cost of litigation.\textsuperscript{22}

\textsuperscript{22} Pp. 297-298 of the article.
The ad hoc character of the authors’ solutions

The authors’ approach suffers from a number of ad hoc patches in addition to the ad hoc solution that they adopt in limiting the use of APP to charges in a single trial.

One such ad hoc patch can be found in their answer to one of the counterclaims that they examine. The authors consider the argument that APP creates a fear of prosecutorial abuse that would result in the conviction of defendants through combining a number of weak files. In order to resolve this issue, the authors note, inter alia, that the evidence in each file must be specific, and condition the combining of files upon a minimum probability of guilt of 0.5 in each file.

First, with all due respect, I would note, incidentally, that it is not clear what manipulation or abuse the authors fear. If the authors’ approach is adopted, then in what sense could the prosecution be accused of manipulation or abuse if it adds an additional offense to the charge sheet, for which it has only minimal evidence, in order to ensure a conviction of some offense? After all, according to the authors, what is important is whether some offense was committed. Therefore, the prosecution’s decision not to make an effort to find additional evidence in one file, upon the assumption that the additional evidence in another file will lead to a conviction on the basis of APP must be deemed a legitimate consideration in the framework of APP, otherwise the system is not coherent. The authors’ approach to the problem, and their attempts to resolve it, appear to reveal that they are not entirely comfortable with their own approach.

But the main problem is that the solution that the authors provide for “abuse” (if one might view honestly following the authors’ approach in good faith as abuse) is ad hoc. It is not clear why the conjoining files must be made conditional upon a probability of guilt of 0.5 (rather than 0.4 or 0.6, for example) for each of the offenses.
Moreover, the authors’ demand for specific evidence also seems somewhat problematic. As we have seen, APP is only consistent with an approach that quantifies probabilities, viz., it requires a Pascalian mathematical approach. That being so, why must the evidence be specific?

E. Conclusion

In this article, I have tried to demonstrate the inadmissibility of the APP approach. The article shows that APP is premised upon a Pascalian mathematical approach to the standard of proof. In that framework, the authors’ treatment of interdependence is mathematically in error. A correct mathematical approach to probabilistic interdependence among offenses would render APP impractical.

Finally, the article shows that the various arguments that the authors bring in support of APP are unpersuasive. Inter alia, the article raises doubts as to whether adopting APP would actually lead to a reduction of enforcement costs or to minimizing adjudication errors, and shows that adopting APP would require that we choose between equality and deterrence.