The Clean Paper

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

一个密码学证明可以通过让缺失的模拟器难以被证明来隐藏其秘密

零知识证明允许某人证明一个陈述而不揭示见证。经典理论说，在完整意义上，你无法用一条消息、无需可信设置和完美可靠性来做到这一点。*Rahul Ilango* 的论文没有使这种不可能性消失。它改变了目标：不再要求模拟器确实存在，而是要求没有选定的证明系统能够高效地证明模拟器不存在。在重要的证明复杂性和密码学假设下，这足以逐属性地恢复零知识的可证伪、博弈式后果。重点不是一个即插即用的互联网原语。它是将数学不可证明性转化为密码学掩护的一种证明论方法。

Written by Cinzia Vaglio · Cross-checked by Vera Rubin · AI-assisted, human-reviewed

Paper: *Gödel in Cryptography: Effectively Zero-Knowledge Proofs for NP with No Interaction, No Setup, and Perfect Soundness*, Rahul Ilango, FOCS 2025 / IACR ePrint 2025/1296 · DOI: 10.1109/FOCS63196.2025.00058

诀窍不是证明秘密被隐藏了

从零知识最简单的版本开始。

Alice 想说服 Bob：某个数独谜题有解。如果她把解发过去，Bob 就信服了，但谜题也就被毁了。她想要的东西更奇怪：一个证明解存在、却不泄露解本身的证明。

这就是零知识证明的承诺。证明者（Alice）说服验证者（Bob）某个陈述为真，同时除了该陈述为真这一点之外不泄露任何东西。

问题在于，这个承诺是有代价的。一个普通的数学证明有两个令人舒适的特征。它是一**条消息**：你把它写下来，交出去，然后走开。而且它是**完美可靠的**：一个假的陈述根本没有任何有效的证明。经典的不可能性结果表明，零知识必须放弃这两个特征——而且不仅仅是二者一起放弃；每一个特征单独来看都是禁区。

首先，零知识证明需要对话。如果 Alice 发送一条消息，而事先没有安排任何可信设置，那么零知识保证就会崩溃——而且无论你愿意为交换而放弃多少可靠性，这一点都成立。

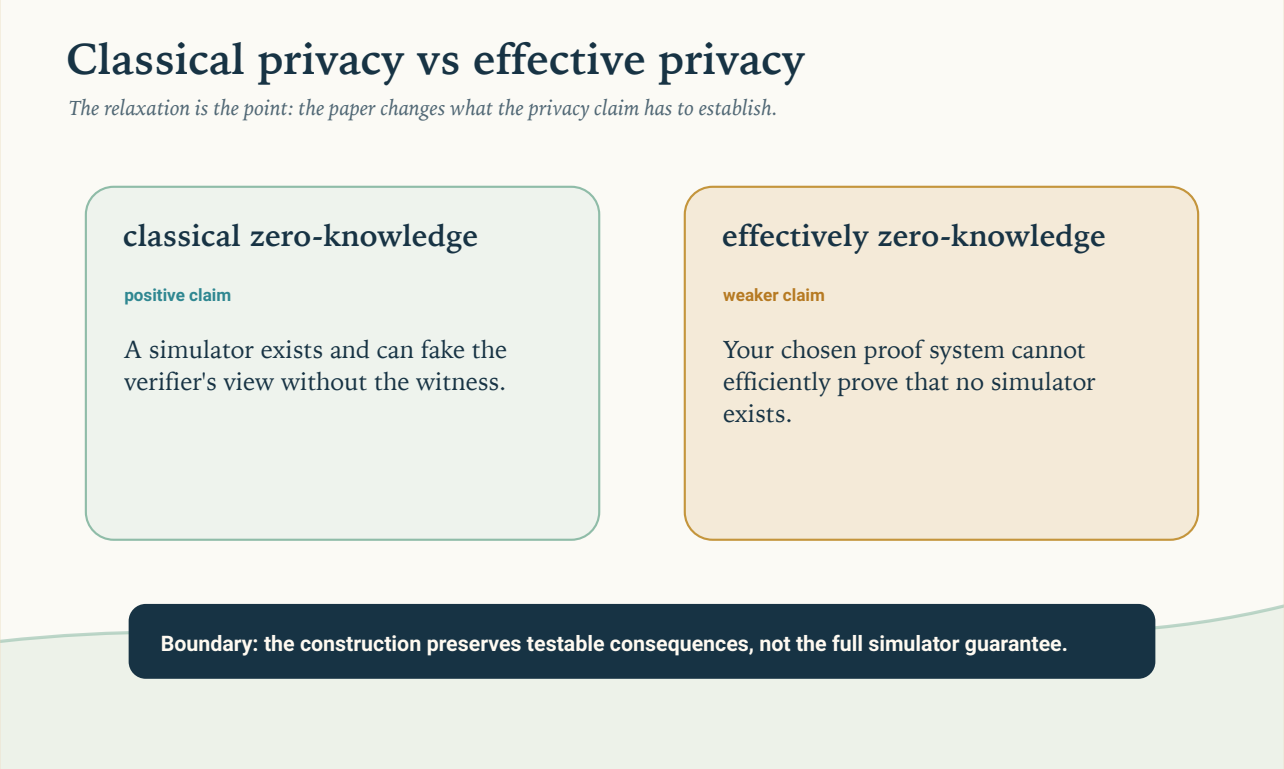
其次，零知识证明需要对错误有一点点容忍。事实证明，要求完美可靠性也会悄悄地摧毁交互：一个无论做出哪些随机选择都永远不会被欺骗的验证者，不如干脆事先把那些选择固定下来——而一旦验证者是可预测的，Alice 就可以在一条消息里回答所有问题，这恰恰就是前面已经崩溃的那种情形。

Rahul Ilango 的论文讲的是一种绕过这道双重高墙的方法。不是假装这堵墙不存在，也不是在那个不可能的设定下造出经典的零知识。这一手更微妙：削弱 “不泄露任何东西” 的含义，但以一种能保留密码学家实际上可以检验的安全性质的方式来削弱它。

这个结果被称为**有效零知识**。

零知识在三道门前被挡住——交互、可信设置和不完美的可靠性。Ilango 的构造从另一道门溜了过去：规则手册无法高效地反驳模拟器。

Original diagram —The Clean Paper CC BY 4.0



经典零知识问的是模拟器是否存在；“有效零知识”只问你所选的规则手册能否高效地证明模拟器不可能存在。正是这个更弱的问题，让这个构造得以保留一条消息、无设置和完美可靠性。

Original diagram —The Clean Paper CC BY 4.0

旧的检验：存在一个模拟器

将零知识形式化的经典方法用到了一个虚构的助手，叫作**模拟器**。

想法是这样的：想象一个叫 Jane 的人，她**不**知道 Alice 的秘密。如果 Jane 完全靠她自己就能生成一些看起来和 Bob 本会从 Alice 那里收到的证明一模一样的证明，那么 Alice 的证明并没有教给 Bob 任何新东西。Jane 在没有 Alice 秘密的情况下就已经能伪造出这种体验了。

所以经典零知识要求有一个真正的模拟器。必须存在一个高效的算法，它能在不知道秘密——用行话说就是见证——的情况下生成看起来像模像样的假证明；对数独来说，见证就是那张解好的网格。

这个定义很强大，但它也恰恰是旧的不可能性发作的地方。直觉是这样的。一个真正非交互的证明只是一个字符串。一旦 Bob 拿到了这个字符串，他就可以把它拿给别人看：他获得了向他人证明该陈述的能力，这听起来已经不止是“什么都没有”了。经典定理把这个直觉锐化成了上面那些不可能性。

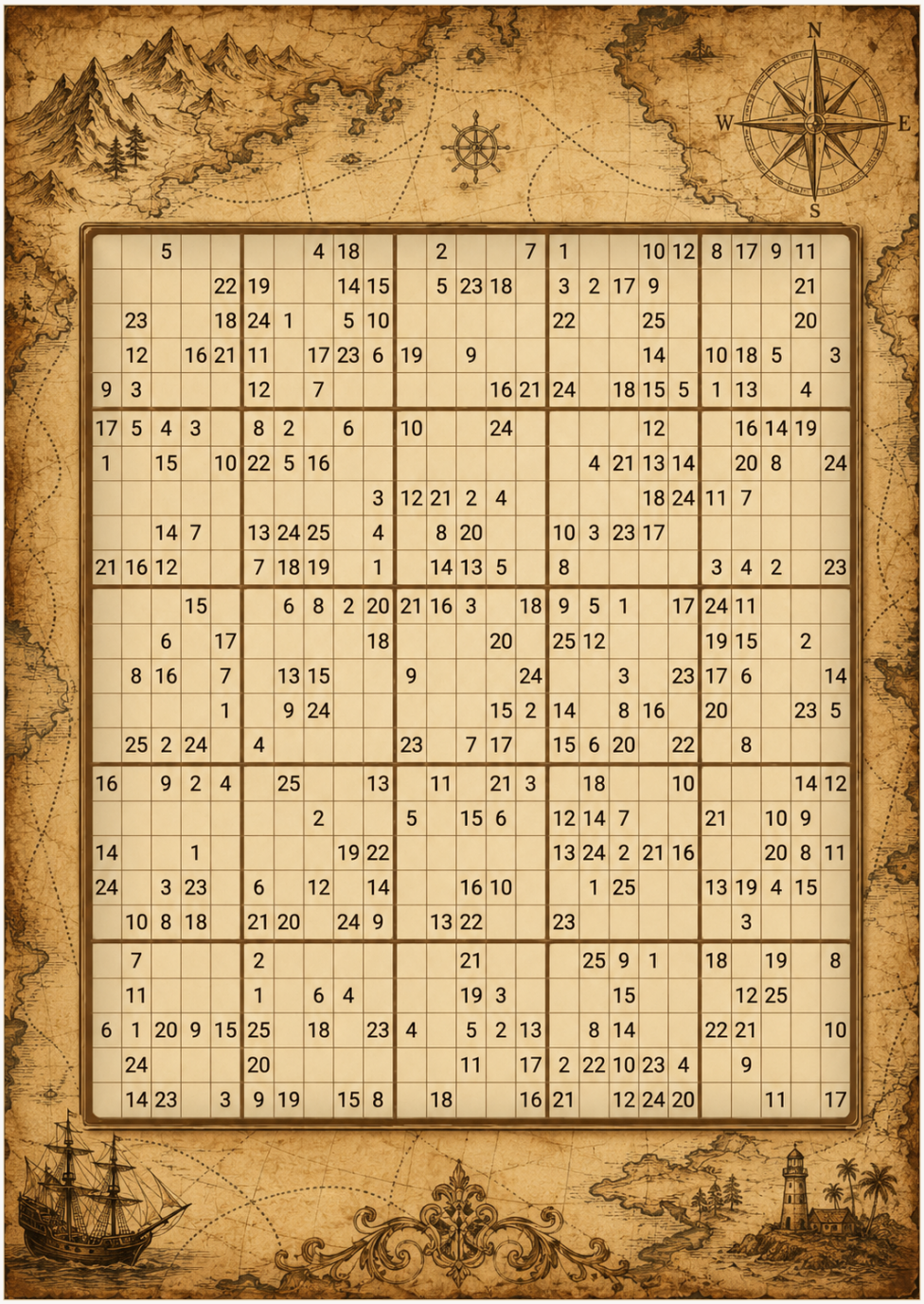
本文坚持的三个性质

论文的标题点出了三个约束：
无交互： Alice 发送一个证明字符串。没有来回往复的协议。
无设置： Alice 和 Bob 不依赖任何可信的公共参考字符串或其他事先安排好的公共随机性。许多被称为 “非交互零知识” 的系统仍然依赖设置；本文所说的是零设置。
完美可靠性： 一个假的陈述没有任何有效的证明。不是 “几乎从不被接受”；而是根本不存在有效的证明。
这三个性质恰恰是普通的书面数学所具备的——而且，正如上面所解释的，经典零知识无法保留它们。

这个区别的超级数独版本

下面是一种刻意简化的、用来感受这个区别的方式。

在这个类比的严肃部分，不要用普通的 9×9 数独。它太小、太有限了：计算机可以直接解出它，或者证明它无解。取而代之，想象一族 **MegaSudoku(n)**（超级数独）谜题。把通常的规则放大：选择一个块大小 n，令 $N = n^2$，构建一个 $N \times N$ 的网格，划分为 $n \times n$ 的块，使用 N 个符号。普通数独只不过是那个很小的 $n = 3$、$N = 9$ 的情形：一个 9×9 的网格、3×3 的块和九个符号。只有当允许 n 增长，并且当网格可以承载额外的小配件（gadget），使它表现得像一个伪装成数独谜题的 SAT 公式时，证明复杂度的故事才真正开始。一个 SAT 公式只不过是一列是/否约束：你能否给这些变量赋上真/假值，使得每一个约束都被满足？



一个 25×25 的数独：它的规则可以在不泄露完成后的网格的情况下被检验——这是一个对隐藏解（即见证）进行验证的证明的视觉替身。

AI-generated editorial thumbnail —The Clean Paper CC BY 4.0

数独与 SAT：同一个谜题的两种装扮

说数独可以“表现得像一个 SAT 公式”，这并不是一个比喻。这种转换在两个方向上都成立，而且容易的那个方向可以完整地写出来。

从数独到 SAT。 SAT 只会说真/假，所以给每个（行、列、值）三元组配一个布尔变量：$x(r,c,v)$ 表示“第 r 行、第 c 列的格子里装着值 v”。一个 4×4 的数独（2×2 的块，值为 1–4）需要 4·4·4 =

64 个变量；经典的 9×9 需要 729 个。于是，每一条数独规则都变成一批子句。（一个子句是若干变量或它们的否定的“或”（OR）；整个公式是它所有子句的“与”（AND）。）
每个格子至少装一个值 ——每个格子一条子句：

x(1,1,1) ⊠ x(1,1,2) ⊠ x(1,1,3) ⊠ x(1,1,4)

每个格子至多装一个值 ——对每一对值都有一条“不能兼得”的子句：

¬x(1,1,1) ⊠ ¬x(1,1,2) ¬x(1,1,1) ⊠ ¬x(1,1,3) … 对全部六对依此类推。

每一行都包含每一个值 ——对第 1 行和值 3 来说：至少出现一次，

x(1,1,3) ⊠ x(1,2,3) ⊠ x(1,3,3) ⊠ x(1,4,3)

以及至多出现一次：¬x(1,1,3) ⊠ ¬x(1,2,3)，并对该行中每一对格子依此类推。
列和块 ——同样的一批批子句；只是格子的分组变了。对左上角的块和值 2 来说：

x(1,1,2) ⊠ x(1,2,2) ⊠ x(2,1,2) ⊠ x(2,2,2)

再加上成对的“不能兼得”子句。
印好的提示数字 ——最简单的部分：每个提示都是一条只含单个变量的子句。左上角印着的一个 3 就变成子句

x(1,1,3)

所有这些的“与”（AND）恰好在数独有解时可满足——而且一个可满足的赋值就是那个解：读出哪些 x(r,c,v) 为真，然后把网格填上。对一个 9×9 来说，这总共是 729 个变量和几千条子句，一个现代 SAT 求解器在几毫秒内就能处理完。注意那条提示子句 x(1,1,3)：它说的是“这个格子恰好等于 3”，而不是“这些格子两两不同”——正是这种不对称性，会迫使下面那个协议说明里对提示格子采用额外的技巧。
从 SAT 到数独。论文需要的是相反的、更难的那个方向：给定一个任意的 SAT 公式，构造一个恰好在该公式有解时才有解的超级数独。数独原生的规则只会说“这些格子两两不同”，所以任意的逻辑约束必须被构造出来——而这正是那些 gadget 所做的事情。一个 gadget 是一小簇预先制造好的格子，公式的每一条子句对应一簇，其中指定的一些格子扮演变量的角色（它们所装的符号编码真或假），而这簇格子的内部约束被精心设计，使得它唯一合法的填法对应于满足该子句的那些赋值。这是 NP 完全性证明中的标准手艺；对于广义数独，它由 Yato 和 Seta 在 2003 年完成。
把这两个方向合起来，就是说 N×N 数独和 SAT 是同一个问题穿着不同的戏服。正是这一点，才准许这篇文章——以及这篇论文——用网格和符号来讲一个关于整个 NP 的故事。

见证仍然很容易想象。Alice 知道超级数独的一种完整的合法填法。Bob 想要被说服这样一种填法存在，但 Alice 不想把它泄露出来。如果她把整个填法发过去，Bob 就信服了，但秘密也就没了。

在经典零知识的版本里，Alice 和 Bob 进行交互。一个老派的思维模型用到了盖住的方块。Alice 把解好的网格藏起来，在每一轮之前偷偷地给符号重新命名，然后让 Bob 检查一个随机选中的局部约束：一行、一列、一个宫，或者一个 gadget。如果被揭开的格子显示出的符号两两不同，Bob 就增加了几分信心。然后一切又被盖起来，符号也被重新命名一遍。（有一个小麻烦：谜题给定的提示需要一个额

外的技巧，因为给符号重命名也会把它们藏起来。下面的说明解释了经典协议如何解决这个问题；对于接下来的内容，这个玩具式的图景已经足够了。）

> **经典协议究竟如何处理提示格子**
>
> 重命名这个技巧有一个盲点。行、列、宫的规则都说“这些格子两两不同”，而两两不同在符号的任何重命名下都能保持。但一个提示说的是“这个格子里恰好装着 5”，而在重命名之后，Bob 只看到 $\sigma(5)$——某个被掩盖的符号——却不知道重命名 σ。他什么都检查不了。如果不加修补，Alice 就可以证明某个合法网格存在，却完全无视印好的提示，而这对这个谜题什么都证明不了。经典文献里有两种标准的修补办法。
>
> **调色板**。给隐藏的网格加上额外的一行 N 个格子——一块调色板，Alice 按照固定的、公开的顺序把符号 1…N 填进去，然后和其他一切一起重命名，于是它装着 $\sigma(1)…\sigma(N)$。现在 Bob 的随机挑战多了一个选项。除了挑一行、一列、一个宫或一个 gadget 来揭开之外，他还可以挑**这块调色板加上一个提示格子**。Alice 把两者都揭开；调色板揭示了那一轮的重命名，Bob 检查那个提示格子显示的恰好是印好的提示经过重命名后的版本。这仍然是零知识的，因为 Bob 只学到 σ——它每一轮都是新抽的，单独来看毫无价值——以及一个他早已从谜题中知道其值的格子的值。关于秘密格子的任何信息都没有泄露，而一个模拟器可以通过抽一个随机的 σ 来伪造这个视图。它是可靠的，因为一个作弊的 Alice 每一轮都以固定的概率被抓住，而且轮次会一直重复到疑虑可以忽略为止。
>
> **把提示编译掉**。一个更具结构性的变体，不是增加而是移除这个特殊的挑战。与其验证提示的值，不如用差异约束把它强制出来：把提示格子和除了承载它自己那个值之外的每一个调色板格子都连起来——“与 $\sigma(1)$ 不同、与 $\sigma(2)$ 不同、……、与除 $\sigma(5)$ 以外的一切都不同”。这个格子唯一能合法装的符号就是提示的那个值。现在每一条约束又都是“这两个不同”那一类了——在重命名下不变，可以完全像检查一行那样来检查。这与经典图着色协议中对预先着色的顶点所用的手法是同一个，也正是上面 *gadget* 一词的精神所在：在“把超级数独当作 SAT”的图景里，提示和其他每一条约束一样，都被编译成了不等式 gadget。
>
> **物理协议**。数独的现实世界纸牌协议（Gradwohl、Naor、Pinkas 和 Rothblum，2007）完全不用重命名，并且在藏起来之前就把提示搞定了。对每一个格子，Alice 放下三张写着该格子值的相同的牌——秘密格子的牌面朝下，但**提示格子的牌面朝上**，这样在牌被翻过来之前，Bob 就亲眼看到提示得到了遵守。然后每个格子拿出一张牌放进它所在行的包，一张放进它所在列的包，一张放进它所在宫的包；每个包都被洗牌然后揭开，Bob 检查它包含全部 N 个符号。洗牌摧毁了位置信息（这就是零知识的所在），但提示在发牌的时候就已经被钉死了。
>
> 无论哪种办法，教训都是这篇文章反复回到的那一个：一个零知识协议就是一份小心翼翼的账目，记着哪些事实在隐藏之后还能存活。重命名保住了“两两不同”，抹去了“等于 5”——所以“等于 5”必须靠其他手段偷偷带回来。

这并不是论文里的协议。它是经典零知识的思维模型：

- Alice 和 Bob 来回往复。
- Bob 选择随机的检查。
- Alice 只揭示局部的一致性，而不是整个解。
- 隐私性的证明是通过表明 Bob 所看到的视图本可以在没有 Alice 秘密解的情况下被生成来实现的。

所以经典零知识是围绕一个肯定的事实建立起来的：

> 一个模拟器真的存在。

现在把那些令人舒适的部分去掉。Alice 发送一个证明字符串，然后走开。没有可信设置，没有事先准备好的共享随机字符串，而且 Bob 绝不能接受一个假的谜题。这就是经典零知识无法存活的那个设定。

在这个诀窍之前，还需要一个角色。固定一个**规则手册**：逻辑学家意义上的一个形式证明系统——一组固定的公理，加上用来检查书面数学证明的机械规则。ZFC，即数学的标准公理，就是那个典范的例子。从这里开始的一切，都是相对于一个事先选定的规则手册来陈述的，而且这个选择是灵活的：这个构造对你固定的任何规则手册都成立，包括 ZFC。

（一个关于用词的说明，借自论文本身：这里的“证明系统”始终指的是这个规则手册——那个检查数学证明的形式系统——而绝不是 Alice 发送的消息。Alice 和 Bob 的那套机制被称为“证明者和验证者”。）

哥德尔风格的版本保留了超级数独的故事，但改变了证明。

挑选第二个显示尺寸相同的约束系统，把它叫作 **D**。就故事而言，S 和 D 是两个格式相同的 MegaSudoku(n) 谜题。在幕后，D 可能一开始是一个尺寸不同的困难逻辑公式；如果需要，可以用一些无害的虚设约束把它填充起来，让它适配同一个网格。D 是由一个实际上**不可满足**的逻辑公式构造出来的：不存在任何一种可能的赋值能让它所有的约束都为真，就像一个坏掉的谜题没有合法的完成网格一样。一个玩具式的例子会是一个既要求“X 为真”又要求“X 为假”的公式。所以 D 没有合法的填法。

但 D 不能是一个容易被揭穿的坏掉的谜题。上面那个玩具例子就不合格：任何规则手册都能用一行反驳“X 且非 X”。D 必须以一种所选规则手册无法用简短论证来证实的方式为假。如果规则手册能用一个简短的证明反驳 D，下面的故事就会崩塌：那条本可能在没有 Alice 秘密的情况下产生证明的替代路线，就可以被形式地排除掉，隐私保证也随之一起排除掉。所以 D 是从一个固定规则手册无法高效反驳的族中挑选出来的：在那个规则手册内部，不存在关于 D 无解的简短证明。

于是 Alice 的一条消息证明所针对的是一个二选一的陈述：

> 要么真正的超级数独 S 有解，要么诱饵 D 有解。

这就是那个逻辑上的联系。D **并不是**以某种能让 S 为真的魔法方式生成的。这个证明并不是在论证“D 无解，因此 S 有解”。它证明的是析取式 **S 或 D**。完美可靠性说的是一个假的析取式不可能有有效的证明。既然 D 在现实中为假——它没有解——那么这个析取式能为真的唯一途径就是 S 为真。所以如果证明被接受，S 就必然有解。诱饵没法让一个假的 S 变成真的。

但对于零知识风格的那一部分，问一问如果 D 确实有解会发生什么。那个诱饵的解会充当一个替代的见证。它会让某人在不知道 Alice 真正的超级数独解的情况下也能产生证明——换句话说，就是一个模拟器。在现实中 D 没有解，所以这条模拟器路线是封闭的。关键在于，规则手册无法高效地证明它是封闭的。

所以 D 有两项工作。对于**可靠性**，D 为假，因此一个关于“S 或 D”的有效证明强制 S 成立。对于**有效零知识**，D 难以反驳，因此规则手册无法迅速排除那条本会让模拟成为可能的诱饵路线。

所以安全性的检验不再是：

> 我们能否证明一个模拟器真的存在？

它变成了：

> 你的规则手册能否高效地证明这个模拟器不可能存在？

如果答案是否定的，就会得出某种出人意料地强的结论：每一个满足如下两条的安全保证——(a) 可以通过运行一个测试来观察到，并且 (b) 在那个规则手册内部，可证明地从模拟器的存在推出——实际上都成立。对它们中任何一个的成功攻击，本身就相当于那个缺失的简短反驳，而那个缺失的简短反驳并不存在。这就是 “有效零知识” 里 “有效” 的那一部分。

所以课堂上的对比是：

经典零知识：这些证明是安全的，因为存在一个模拟器。

哥德尔风格的有效零知识：对于可观察的安全性测试，这些证明被当作是安全的，因为规则手册无法高效地证明这个模拟器不可能存在。

第二个断言更弱。它也正是论文能够保留那三个打垮了经典版本的特征的原因：一条消息、无设置和完美可靠性。

新的检验：你无法证明模拟器不存在

Ilango 的松弛改变了问题。

经典零知识问的是：

> 是否存在一个模拟器？

有效零知识问的是某种更弱的东西：

> 你所选的规则手册能否高效地证明不存在任何模拟器？

这听起来像是一个技术性的躲闪，但它正是核心思想。这个构造处于一种奇怪的状态：模拟器实际上并不存在——论文对这一点说得很明确——但你所固定的那个规则手册无法高效地证明它不存在。如果你所关心的每一个坏后果都需要这样一个反驳，那么对于那些后果来说，这个系统仍然表现得像零知识。

这里就是哥德尔登场的地方。不是作为装饰，也不是 “哥德尔让密码学变安全”。这个联系是证明论意义上的。一个规则手册被称为**最优的**，如果它在一个精确的意义上是可能的最好的那一个：每当任何一个规则手册能用一个简短的证明反驳某种相关类型的公式时，那个最优的规则手册也能，用一个至多多项式地更长的证明。Krajíček 和 Pudlák 在 1989 年猜想**不存在最优证明系统**：无论你固定哪个规则手册，总有另外某个规则手册能远为简洁地证明某一族真陈述。这是证明复杂度的核心未解猜想之一，也是哥德尔不完备性定理在有限的、复杂度理论意义上的表亲：有些真陈述在你所固定的规则手册里没有简短的证明——不是因为它们原则上不可证明，而是因为每一个固定的规则手册都会留下一些简短的真理，它们没有简短的证明。

论文假设这个猜想（以一个稍强的“无穷频繁”形式，这是把猜想用于密码学时的标准做法）。根据 Krajíček 和 Pudlák 的一个定理，回报是具体的：对每一个规则手册，都存在一个公式序列，它们确实不可满足，规则手册无法用简短的证明反驳它们——而且，至关重要的是，一个高效的算法能够生成它们。最后那个性质，即一致性（uniformity），正是把整个想法从一个存在性断言变成 Alice 真正能运行的一个算法的东西：她的诱饵 D 是从流水线上下来的，而不是凭空变出来的。

密码学上的这一手，就是把那种证明能力的短缺利用起来。

这个构造在做什么

下面是论文的构造，被剥到只剩它的骨架形状。

固定一个规则手册——比如说 ZFC。在证明复杂度假设之下，存在一个可高效生成的公式序列，它们实际上不可满足，但规则手册没有关于它们不可满足的简短证明。

现在构建一个如下形式的一条消息证明：

> 要么真正的陈述可满足，要么这个特殊的困难公式可满足。

这个特殊的困难公式是不可满足的。所以如果底层的证明机制是完美可靠的，接受这条消息仍然意味着真正的陈述为真。这就给出了完美可靠性。

但对于类似零知识的安全性，想象这个特殊的困难公式确实可满足。那么它的见证就可以被用来在不知道真正见证的情况下模拟证明。这个公式在现实中不可满足——但规则手册无法高效地证明这一点。所以它无法高效地证明这个模拟器不可能存在。

这就是那个枢纽。这个系统不是通过产生一个经典的模拟器来隐藏秘密的。对于一大类可观察的安全性测试来说，它把秘密藏在了规则手册无力证实模拟器不存在的背后。

论文主张了什么

主定理是分层的。核心结果是这样的：

在一个标准的密码学假设——**非交互见证不可区分证明**的存在，这是一类被充分研究过的对象，可从若干已确立的假设包中推出——之下，以及在**不存在（无穷频繁）最优证明系统**这个证明复杂度猜想之下，论文为每一个规则手册的选择都构造了一个针对 NP/SAT 的一条消息证明者和验证者，它具有**完美可靠性**、无设置，并且相对于那个规则手册是有效零知识的。（NP/SAT 是各种谜题式问题的标准“最难公分母”；超级数独是它所穿的一件戏服。）

对于那个关于保留可证伪安全性质的更宽泛的主张，论文再添加一个标准假设，即去随机化的信念 **P = BPP**（大致是说：随机性没有给算法带来任何本质上的额外能力）。

把定理的语言翻译过来：

- 证明是一条消息。
- 没有可信设置。
- 假的陈述无法被证明。
- 这个证明者不是经典零知识的——它没有模拟器。
- 但经典零知识的每一个可证伪的、基于博弈的安全性后果，都可以在这个设定中实现。

“可证伪”这一点很重要。它意味着一次安全失败可以通过在一场博弈中运行一个对手来测试。许多密码学的安全定义都有这种形式：对手能否区分两个密文、求逆一个函数、恢复一个见证，或者赢得某个指定的实验？这个定理为每一个可证伪的性质各给出一个证明者，一次一个。一个同时享有每一个可证伪性质的单一证明者很可能是不可能的——旧的可重用性攻击（“Bob 可以把证明拿给别人看”）本身就是一个可证伪的性质，而它在这里确实失败了。论文的提议是，一个单一的证明者有理由能够涵盖所有自然的可证伪性质——那些在密码学实践中真正出现的性质——但那一部分是一个条件性的定理，依赖于一个非正式的“自然”概念，外加一个明确的猜想。这个保证针对的是可观察的失败，而不是保密性的每一种哲学意义或基于模拟的意义。

有一个具体的推论值得点名：这个构造给出了第一个带有一致证明者的非交互见证隐藏证明——“一个谜题的证明并不能帮你找到它的解”，没有交互也没有设置——这是一个听起来不起眼、却抵抗了数十年构造尝试的对象。

这不是在说什么

这是让整篇文章保持诚实的那一节。

它**并不**是说旧的不可能性定理是错的。这个构造是通过改变定义来绕开它们的。

它**并没有**给出无交互、无设置和完美可靠性的普通的、经典的零知识。论文明确地说，所构造的证明者没有模拟器。

它**并不**意味着这个证明不能被重用。一条消息的证明仍然可以拿给别人看；论文并没有保留可否认性那一类的性质。（带可信设置的非交互零知识也有同样的局限。）

它**并不**意味着这是一个可以部署的实用协议。这是复杂度理论和密码学基础。这个结果依赖于来自证明复杂度和密码学的重大假设，而这个构造关乎的是原则上什么是可能的。

它**并没有**把“哥德尔”变成一个魔法般的安全原语。哥德尔的联系是通过证明系统、最优证明系统以及不完备性的有限类比来实现的。可用的直觉不是“不完备性保护你的密码”。而是：如果一个规则手册无法高效地证明一个模拟器不可能存在，那么那些需要该证明的攻击就可以在安全定义的层面被挡住。

为什么它无论如何都很有趣

密码学常常把困难性转化为安全性。因数分解是困难的，所以 RSA 式的假设变得有用。格问题是困难的，所以格密码学变得有用。这里的困难性更奇怪：不是“难以计算出一个秘密”，而是“难以证明某个证明对象不可能存在”。

这就是为什么这篇论文让人感觉不同寻常。它几乎把公理和规则手册当作密码学资源来对待。通常的不可能性说的是，可靠性和模拟之间存在一种张力。Ilango 这一手，是把这种张力放到一道证明论的帷幕后面：模拟器不存在，但形式系统无法高效地揭露这种不存在。

对一个读者来说，令人惊讶的部分并不是这将取代今天的零知识系统。它大概不会，至少不会直接取代。令人惊讶的部分在于，一个来自数理逻辑的局限性可以被建设性地使用：不只是作为一堵墙，而是作为一种掩护。

证据有多强?

这是一篇定理论文，所以“证据”的含义不同于一篇生物学或天文学论文。问题不在于一个实验是否被复现。问题在于那些定义、假设和证明链条是否支持这个主张。

证明是形式化的，而且论文对它的假设说得很明确。这些假设不是随随便便的。非交互见证不可区分证明是密码学中的标准对象，可从若干已确立的假设包中推出。不存在最优证明系统这个猜想是证明复杂度中的一个核心猜想。P = BPP 是一个标准的去随机化信念，只用于那个更宽泛的可证伪性质定理。

论文还论证说，这些假设是恰当的代价，而不是一个随意搭起来的脚手架：它证明了一个逆命题，表明这些假设本质上是必要的——如果像这样的构造根本存在，那么非交互见证不可区分证明就必然存在，并且（在承认标准单向函数的前提下）不可能存在最优证明系统。而且这些假设是“双赢”的：反驳它们中的任何一个，本身就会是证明复杂度、密码学或复杂度理论中的一个里程碑式的发现。

但由于这个结果是条件性的，它的可信度也是条件性的。如果那些假设不成立，定理的解读就会改变。而且即便那些假设成立，这个保证也不是完整的经典零知识；它是论文那个松弛过的、证明论意义上的版本。

所以恰当的可信度是：对于论文确立了一个自洽的条件性可能性结果这一点，是高的；对于它的假设描述了我们实际所处的密码学世界这一点，是中等的；而对于任何即时的实用后果，则是低的。

为什么这很重要

这篇论文打开了一条本应封闭的路线。

经典理论说：完整的零知识不可能是无设置的一条消息，也不可能是完美可靠的。Ilango 的论文说：如果我们要的是零知识那些可以在安全博弈中被测试的后果，并且如果我们允许安全定义取决于一个规则手册能或不能高效反驳什么，那么许多有用的行为就可以被恢复——用一条消息、无设置和完美可靠性。

这并不是一个小小的定义上的微调。它是一种思考密码学保证的不同方式。不要只问什么存在，而要问你的规则手册能排除什么。不要把不可证明性当作一种哲学上的烦扰，而要把它当作一种结构来使用。

实用的世界可能不会在明天就改变。但概念上的地图变了。现在有了一种形式上的意义，在这种意义

下，“没有人能高效地证明秘密泄露了” 可以强到足以恢复许多我们本想从 “秘密没有泄露” 中得到的、基于博弈的保护。

这就是为什么哥德尔理应出现在标题里。

干净的总结

零知识证明让一个证明者能够说服一个验证者某个陈述为真，而不泄露见证。经典的不可能性结果说，零知识不可能被塞进无设置的一条消息里，也不可能具有完美可靠性。Rahul Ilango 的论文并没有反驳那些不可能性。它定义了一个更弱的概念，即有效零知识：它不是要求一个模拟器真的存在，而是要求一个所选的证明系统——一个像 ZFC 那样的形式规则手册——无法高效地证明不存在任何模拟器。在来自密码学（非交互见证不可区分证明）和证明复杂度（不存在最优证明系统）的重大假设之下，论文为 NP/SAT 构造了无设置、完美可靠的一条消息证明者，它们逐个性质地实现了零知识那些可证伪的、基于博弈的后果。一个单一的证明者涵盖所有 “自然的” 这类性质，是一个更进一步的、部分带有猜想性的扩展——而要涵盖字面意义上每一个可证伪的性质，则很可能是不可能的，因为证明仍然是可重用的。这个结果是理论性的、条件性的，不是一个已部署的原语，但它展示了一种把证明论意义上的不可证明性用作密码学资源的新方法。

不扯淡核查

论文展示了什么：在所陈述的假设之下，人们可以为 NP/SAT 构建一条消息、无设置、完美可靠的证明者，它们相对于任何所选的证明系统都是有效零知识的，并且实现了经典零知识每一个可证伪的、基于博弈的后果。

什么是有道理但并未被无条件证明的：所需要的证明复杂度和密码学假设成立。它们是严肃的、被充分研究过的假设——而且论文表明它们不仅充分，还本质上是必要的——但仍然是假设。

它没有展示什么：无交互、无设置和完美可靠性的经典零知识；一个可以部署的实用系统；证明的可否认性或不可重用性；或者哥德尔不完备性定理本身就能保障密码学的安全。

主要局限：这个保证是零知识的一个松弛；最宽泛的版本依赖于多个假设；单一通用证明者的那些主张仍然部分带有猜想性；而且这个结果主要是基础性的。

一个普通读者应该有多少信心？如果接受那些定义，那么对于这是一个重要的条件性理论结果这一点，信心是高的。对于这些假设抓住了现实这一点，是中等的。对于即时的实用部署，是低的。安全的要点是：这篇论文并没有打破零知识的不可能性；它找到了一种新的、证明论意义上的方法，绕过它们当中对许多安全博弈来说重要的那些部分。

来源

Ilango, IACR ePrint 2025/1296
https://eprint.iacr.org/2025/1296

FOCS 2025, DOI 10.1109/FOCS63196.2025.00058
https://doi.org/10.1109/FOCS63196.2025.00058