Concurrent Zero-Knowledge With Timing, Revisited (Preliminary Version)

Oded Goldreich*
Department of Computer Science
Weizmann Institute of Science
Rehovot, ISRAEL.
oded@wisdom.weizmann.ac.il

November 27, 2001

Abstract

Following Dwork, Naor, and Sahai (*30th STOC*, 1998), we consider concurrent execution of protocols in a semi-synchronized network. Specifically, we assume that each party holds a local clock such that a constant bound on the relative rates of these clocks is a-priori known, and consider protocols that employ time-driven operations (i.e., `time-out` in-coming messages and `delay` out-going messages).

We show that the constant-round zero-knowledge proof for $\mathcal{NP}$ of Goldreich and Kahan (*Jour. of Crypto.*, 1996) preserves its security when polynomially-many independent copies are executed *concurrently under the above timing model*.

We stress that our main result establishes zero-knowledge of interactive *proofs*, whereas the results of Dwork *et. al.* are either for zero-knowledge *arguments* or for a *weak notion* of zero-knowledge (called ϵ-knowledge) proofs.

Our analysis identifies two extreme schedulings of concurrent executions under the above timing model: the first is the case of *parallel execution* of polynomially-many copies, and the second is of concurrent execution of polynomially-many copies such the number of copies that are simultaneously active at any time is bounded by a constant (i.e., *bounded simultaneity*). Dealing with each of these extreme cases is of independent interest, and the general result (regarding concurrent executions under the timing model) is obtained by combining the two treatments.

Keywords: Zero-Knowledge, Parallel Composition, Concurrent Composition, Timing Assumptions, proofs versus arguments, Black-box simulation,

*Supported by a MINERVA Foundation, Germany.

Contents

1 Introduction

Zero-Knowledge proofs, introduced by Goldwasser, Micali and Rackoff [20, 21], are fascinating and extremely useful constructs. Their fascinating nature is due to their seemingly contradictory definition: they are both convincing and yet yield nothing beyond the validity of the assertion being proven. Their applicability in the domain of cryptography is vast: they are typically used to force malicious parties to behave according to a predetermined protocol (which requires parties to provide proofs of the correctness of their secret-based actions without revealing these secrets). Such applications are based on the fact, proven by Goldreich, Micali and Wigderson [17], that any language in $\mathcal{NP}$ has a zero-knowledge proof system, provided that commitment schemes exist.[1] The related notion of a zero-knowledge *argument* was suggested (and implemented) by Brassard, Chaum and Crépeau [5], where the difference between *proofs* and *arguments* is that in the latter the soundness condition refers only to computationally-bounded cheating provers.

1.1 Composition of zero-knowledge protocols

A natural question regarding zero-knowledge proofs (and arguments) is whether the zero-knowledge condition is preserved under a variety of composition operations. Three types of composition operation were considered in the literature, and we briefly summarize what is known about them:

1. *Sequential composition*: Here the protocol is invoked (polynomially) many times, where each invocation follows the termination of the previous one. At the very least, security (e.g., zero-knowledge) should be preserved under sequential composition, or else the applicability of the protocol is highly limited (because one cannot safely use it more than once).

   Although the basic definition of zero-knowledge (as in the preliminary version of Goldwasser *et. al.* [20]) is not closed under sequential composition (cf. [16]), a minor augmentation of it (by auxiliary inputs) is closed under sequential composition (cf. [18]). Indeed, this augmentation was adopted in all subsequent works (as well as in the final version of Goldwasser *et. al.* [21]).

2. *Parallel composition*: Here (polynomially) many instances of the protocol are invoked at the same time and proceed at the same pace. That is, we assume a synchronous model of communication, and consider (polynomially) many executions that are totally synchronized so that the ith round message in all instances is sent exactly (or approximately) at the same time.

   Goldreich and Krawczyk presented a simple protocol that is zero-knowledge (in a strong sense), but is not closed under parallel composition (even in a very weak sense) [16]. At the time, their result was interpreted mainly in the context of *round-efficient error reduction*; that is, the construction of full-fledge zero-knowledge proofs (of negligible soundness error) by composing (in parallel) a basic zero-knowledge protocol of high (but bounded away from 1) soundness error. Since alternative ways of constructing constant-round zero-knowledge proofs (and arguments) were found (cf. [15, 13, 6]), interest in parallel composition (of zero-knowledge protocols) has died. In retrospect, as we argue below, this was a conceptual mistake.

   We comment that parallel composition is problematic also in the context of reducing the soundness error of arguments (cf. [3]), but our focus here is on the zero-knowledge aspect of protocols regardless if they are proofs, arguments or neither.

[1] Or, equivalently [25, 22], that one-way functions exist.

3. *Concurrent composition*: This notion generalizes both the previous ones. Here (polynomially) many instances of the protocol are invoked at arbitrary times and proceed at arbitrary pace. That is, we assume an asynchronous (rather than synchronous) model of communication.

   In the 1990's, when extensive two-party (and multi-party) computations became a reality (rather than a vision), it became clear that it is (at least) desirable that cryptographic protocols maintain their security under concurrent composition (cf. [10]). In the context of zero-knowledge, concurrent composition was first considered by Dwork, Naor, and Sahai [11]. Their actual suggestions refer to a model of naturally-limited asynchronousness (which certainly covers the case of parallel composition). Essentially, they assume that each party holds a local clock such that the relative clock rates are bounded by an a-priori known constant, and consider protocols that employ time-driven operations (i.e., `time-out` in-coming messages and `delay` out-going messages). This *timing model* is the main focus of the current paper.[2] The previously known main results *for this model* were (cf. [11]):

   - Assuming the existence of one-way functions, every language in $\mathcal{NP}$ has a constant-round concurrent zero-knowledge *argument*.
   - Assuming the existence of two-round perfectly-hiding commitment schemes (which in turn imply one-way functions), every language in $\mathcal{NP}$ has a constant-round concurrent ϵ-knowledge proof, where ϵ-knowledge means that (for every noticeable function $\epsilon : \mathsf{N} \to (0,1]$) a simulator working in time $\mathrm{poly}(n/\epsilon(n))$ can produce output that is ϵ-indistinguishable from the one of a real interaction.[3]

   Thus, no constant-round *proofs* for $\mathcal{NP}$ were previously known to be concurrent zero-knowledge (under the timing model).

1.2 Our results

Our main result closes the gap mentioned above, by showing that a known constant-round zero-knowledge proof for $\mathcal{NP}$ is actually concurrent zero-knowledge under the timing model. That is, we prove that

Theorem 1.1 *The* (five-round) *zero-knowledge proof system for $\mathcal{NP}$ of Goldreich and Kahan [15] is* concurrent zero-knowledge under the timing model.

Thus, the zero-knowledge property of the proof system (of [15]) is preserved under any concurrent composition that satisfies the timing model. In particular, the zero-knowledge property is preserved under parallel composition, a result we consider of independent interest.

Recall that the proof system of [15] relies on the existence of two-round perfectly-hiding commitment schemes (which is implied by the existence of claw-free pairs of functions and implies the existence of one-way functions). Thus, we get:

Theorem 1.2 *Assuming the existence of two-round perfectly-hiding commitment schemes, there exists a* (constant-round) *proof system for $\mathcal{NP}$ that is* concurrent zero-knowledge under the timing model.

[2] We shortly discuss the pure asynchronous model below.

[3] For further discussion of ϵ-knowledge, see Section 1.5.

Using the same techniques, we can show that several other known (constant-round) zero-knowledge protocols remain secure under the concurrent timing-model. Examples include the (constant-round) zero-knowledge *arguments* of Fiege and Shamir [13] and of Bellare, Jakobsson and Yung [4]. The latter protocol (referred to as the BJY-protocol) is of special interest since it is a four-round argument for $\mathcal{NP}$ that relies only on the existence of one-way functions. The above protocols are simpler (and use fewer rounds) than the argument systems previously shown (in [11]) to be concurrent zero-knowledge (under the timing-model), alas their security (under this model) is established by a more complex simulator. (See further details in the appendix.)

1.3 Discussion of some issues

We stress that when we talk of composition of protocols (or proof systems) we mean that the honest users are supposed to follow the prescribed program (specified in the protocol description) that refers to a single execution. That is, the actions of honest parties in each execution are independent of the messages they received in other executions. The adversary, however, may coordinate the actions it takes in the various executions, and in particular its actions in one execution may depend also on messages it received in other executions.

Let us motivate the asymmetry between the independence of executions assumed of honest parties but not of the adversary. Coordinating actions in different executions is typically difficult but not impossible. Thus, it is desirable to use composition (as defined above) rather than to use protocols that include inter-execution coordination-actions, which require users to keep track of all executions that they perform. Actually, trying to coordinate honest executions is even more problematic, because one may need to coordinate executions of different honest parties (e.g., all employees of a big cooperation or an agency under attack), which in many cases is highly unrealistic. On the other hand, the adversary attacking the system may be willing to go into the extra trouble of coordinating its attack on the various executions of the protocol.

Auxiliary inputs and non-uniformity: As mentioned above, almost all work on zero-knowledge actually refer to zero-knowledge with respect to (non-uniform) auxiliary inputs. This work is no exception, but (as in most other work) the reference to auxiliary inputs is typically omitted. We comment that zero-knowledge with respect to auxiliary inputs "comes for free" whenever zero-knowledge is demonstrated (like in this work) via a black-box simulator (see below). The only thing to bear in mind is that allowing the adversary (non-uniform) auxiliary inputs means that the computational assumption that are used need to be non-uniform ones. For example, when we talk of computational-binding (resp., computational-hiding) commitment schemes we mean that the binding (resp., hiding) property holds with respect to any family of polynomial-size circuits (rather than with respect to any probabilistic polynomial-time algorithm).

Expected polynomial-time simulators: With the exception of the recent (constant-round) zero-knowledge argument of Barak [1], all previous *constant-round* arguments (or proofs) utilize an *expected* polynomial-time simulator (rather than a *strict* polynomial-time simulator). (Indeed our work inherits this "feature" of [15].) As recently shown by Barak and Lindell [2], this is no coincidence, because all the above (with the exception of [1]) utilize black-box simulators, whereas no *strict* polynomial-time black-box simulator can demonstrate the zero-knowledge property of a constant-round argument system for a language out of $\mathcal{BPP}$.

Black-box simulation: The definition of zero-knowledge (only) requires that the interaction of the prover with any cheating (probabilistic polynomial-time) verifier be simulateable by an ordinary probabilistic polynomial-time machine (which interacts with no one). A *black-box simulator* is one that can simulate the interaction of the prover with any such verifier when given oracle access to the strategy of that verifier. All previous zero-knowledge arguments (or proofs), with the exception of the recent (constant-round) zero-knowledge argument of Barak [1], are established using a black-box simulator, and our work is no exception (i.e., we also use a black-box simulator). Indeed, Barak demonstrated that (contrary to previous beliefs) non-black-box simulators may exist in cases where black-box ones do not exist [1]. However, black-box simulators, whenever they exist, are preferable to non-black-box ones, because the former offers greater security: Firstly, as mentioned above, black-box simulators imply zero-knowledge with respect to auxiliary inputs.[4] Secondly, black-box simulators imply polynomial bounds on the knowledge tightness, where *knowledge tightness* is the (inverse) ratio of the running-time of any cheating verifier and the running-time of the corresponding simulation [14, Sec. 4.4.4.2].[5]

Perspective: the pure asynchronous model. Regarding the pure asynchronous model, the current state of the art is as follows:

- Black-box simulator cannot demonstrated the concurrent zero-knowledge property of non-trivial proofs (or arguments) having significantly less than logarithmically many rounds (cf. Canetti *et. al.* [8]).[6]
- Every language in $\mathcal{NP}$ has a concurrent zero-knowledge proof with poly-logarithmically many rounds, and this can be demonstrated using a black-box simulator (cf. Kilian and Petrank [23] building upon [26]).
- Recently, Barak [1] demonstrated that the "black-box simulation barrier" can be bypassed. With respect to concurrent zero-knowledge he only obtains partial results: constant-round zero-knowledge arguments (rather than proofs) for $\mathcal{NP}$ that maintain security as long as an a-priori bounded (polynomial) number of executions take place concurrently. (The length of the messages in his protocol grows linearly with this a-priori bound.)[7]

Thus, it is currently unknown whether constant-round arguments for $\mathcal{NP}$ may be concurrent zero-knowledge (in the pure asynchronous model).

On the timing model: The timing model consists of the *assumption* that talking about the actual timing of events is meaningful (at least in a weak sense) and of the *introduction of time-driven operations*. The timing assumption amounts to postulating that each party holds a local clock

[4] In contrast, whether or not a non-black-box simulator implies zero-knowledge *with respect to auxiliary inputs*, depends on the specific simulator: In fact, in [1], Barak first presents (as a warm-up) a protocol with a non-black-box simulator that cannot handle auxiliary inputs, and next uses a more sophisticated construction to handle auxiliary inputs.

[5] That is, a protocol is said to have knowledge tightness $k: \mathsf{N} \to \mathsf{R}$ if for some polynomial p and every probabilistic polynomial-time verifier V^* the interaction of V^* with the prover can be simulated within time $k(n) \cdot T_{V^*}(n) + p(n)$, where T_{V^*} denotes the time complexity of V^*. In fact, the running-time of the simulator constructed by Barak [1] is polynomial in T_{V^*}, and so the knowledge tightness of his protocol is not bounded by any polynomial.

[6] By *non-trivial* proof systems we mean ones for languages outside $\mathcal{BPP}$, whereas by *significantly less than logarithmic* we mean any function $f: \mathsf{N} \to \mathsf{N}$ satisfying $f(n) = o(\log n / \log \log n)$.

[7] We are quite sure that Barak's arguments remain zero-knowledge under concurrent executions that satisfy the timing model. But since these are arguments (rather than proofs) such a result will not improve upon the previously known result of [11] (which uses black-box simulations).

and knows a global bound, denoted $\rho \geq 1$, on the relative rates of the local clocks.[8] Furthermore, it is postulated that the parties know a (pessimistic) bound, denoted Δ, on the message-delivery time (which also includes the local computation and handling times). In our opinion, these timing assumptions are most reasonable, and are unlikely to restrict the scope of applications for which concurrent zero-knowledge is relevant. We are more concerned about the effect of the time-driven operations introduced in the timing model. Recall that these operations are the `time-out` of incoming messages and the `delay` of out-going messages. Furthermore, typically (and in fact also in our work), the delay period is at least as long as the time-out period, which in turn is at least Δ (i.e., the time-out period must be at least as long as the pessimistic bound on message-delivery time so not to disrupt the proper operation of the protocol). This means that the use of these time-driven operations yields slowing down the execution of the protocol (i.e., running it at the rate of the pessimistic message-delivery time rather than at the rate of the actual message-delivery time, which is typically much faster). Still, in absence of more appealing alternatives (i.e., a constant-round concurrent zero-knowledge protocol for the pure asynchronous model), the use of the timing model may be considered reasonable. (We comment than other alternatives to the timing-model include various set-up assumptions; cf. [7, 9].)

On parallel composition: Given our opinion about the timing model, it is not surprising that we consider the problem of parallel composition almost as important as the problem of concurrent composition in the timing model. Firstly, it is quite reasonable to assume that the parties' local clocks have approximately the same rate, and that drifting is corrected by occasional clock synchronization. Thus, it is reasonable to assume that the parties have approximately-good estimate of some global time. Furthermore, the global time may be partitioned into phases, each consisting of a constant (e.g., 5) number of rounds, so that each party wishing to execute the protocol just delays its invocation to the beginning of the next phase. Thus, concurrent execution of (constant-round) protocols in this setting amounts to a sequence of (time-disjoint) almost-parallel executions of the protocol. Consequently proving that the protocol is parallel zero-knowledge suffices for concurrent composition in this setting.

1.4 Techniques

To discuss our techniques, let us fix a timing assumption (i.e., a-priori bound on local clock rates) and consider a c-round protocol that utilizes appropriately selected `time-out` and `delay` mechanisms (which depend on the above bound as well as on a bound on normal message-delivery time). The reader may think of the bound on local clock rates as being close to 1 (or even just 1; i.e., equal rates), and of c as being a constant (in fact, we will use $c = 5$). Furthermore, suppose that all prover's actions in the protocol are time-driven (by the `time-out` and `delay` mechanisms).

A key observation underlying our work is that a concurrent scheduling (of such protocol instances) under the timing model can be decomposed into a sequence of parallel executions (or blocks) such that the number of simultaneously active parallel blocks is bounded by $O(c)$, where the constant in the O-notation depends on the a-priori known bound on the relative clock rates. (Indeed, it is instructive to consider a fixed scheduling, but the observation extends to schedulings that are dynamically-chosen by the adversary.) This (simple) observation applies whenever the timing model is used (and is not restricted to the context of zero-knowledge), and it may be useful towards the analysis of the concurrent execution of any set of protocols under the timing model.

[8] The rate should be computed with respect to reasonable intervals of time; for example, for Δ as defined below, one may assume that a time period of Δ units is measured as Δ' units of time on the local clock, where $\Delta/\rho \leq \Delta' \leq \rho\Delta$.

Let us restate the above observation in concrete terms, assuming (for sake of simplicity) that the clock rates are all equal, that the prover utilizes equal delays between its messages, and that these delays are four times the length of the time-out period: Then, any scheduling of executions (of protocol instances) that respect this timing model can be decomposed into *sub-schedules* such that the following hold:[9]

1. Each sub-schedule consists of an almost-parallel execution of instances of the protocol. That is, each sub-schedule can be partitioned into c disjoint time intervals such that round i in each instance take place within the i^{th} time-interval. (Here we allow to arbitrary delay messages sent by the verifier (as long as the time-out condition is not violated), since this only increases the adversary's power.)

2. The number of sub-schedules that are active (i.e., have some active protocol instance) at any given time is at most $8c$.

In view of the above, it is quite natural to conclude that in order to handle the concurrent timing model it suffices to deal with two extreme schedulings: the parallel scheduling and the bounded-simultaneity scheduling. Indeed, this conclusion is essentially correct in our case.

Handling parallel composition. At first glance, one may be tempted to say that the techniques used for proving that the Goldreich–Kahan (GK) protocol is zero-knowledge [15] extend to showing that it remains zero-knowledge under parallel composition. This would have been true if we were handling *coordinated* parallel executions of the GK-protocol (where the prover would abort all copies if the verifier decommits improperly in any of them). However, this is not what we are handling here (i.e., parallel composition refers to uncoordinated parallel execution of many copies of the protocol). Consequently, a couple of new techniques are introduced in order to deal with the parallel composition of the GK-protocol. We consider these simulation techniques to be of independent interest.

Handling bounded-simultaneity concurrent composition. Experts in the area may not find it surprising that the GK-protocol remains zero-knowledge under bounded-simultaneity concurrent composition. In fact, previous work (e.g., [11]) suggest that the difficulty in simulating concurrent executions of the GK-protocol arises from the case in which a large number of instances is executed in a "nested" (and in particular simultaneous) manner.[10] Furthermore, the work of Richardson

[9] For example, we can place protocol instances in sub-schedules according to their invocation time. Specifically, let us consider the length of the time-out period as one time unit (and so the delays have length 4). Next, consider (consecutive and non-overlapping) time-intervals having unit length, and place an instance in the i^{th} sub-schedule if it is invoked during the i^{th} time-interval. Clearly, the i^{th} and j^{th} sub-schedules are simultaneously active (at some time) only if $|i-j| < 4c$, and so Condition 2 holds. Observe that the i^{th} sub-schedule can be partitioned to c intervals such that Condition 1 holds: Suppose (as is the case in our protocol) that the prover sends the first message and that $c = 2d + 1$. Observe that, for any instance placed in the i^{th} sub-schedule and $k = 1, ..., d$, the $k+1^{\text{st}}$ prover message (coming after k delays) is sent at time $t + 4k$, where $t \in [i, i+1)$ is the invocation time of the instance, and so it is sent within the time interval $[i + 4k, i + 4k + 1)$. By virtue of the time-out mechanism, the $k+1^{\text{th}}$ verifier message is received during the time interval $[t + 4k, t + 4k + 1) \subseteq [i + 4k, i + 4k + 2)$, and by some delaying we may assume it occurs within the time interval $[i + 4k + 2, i + 4k + 3)$. Thus, a partition as in Condition 1 is obtained. Notice that the above analysis extends to the case that the clock rates are within a factor of $1 + (1/2c)$ of one another, because the clock drift accumulated during the execution of one protocol instance is less than one time unit.

[10] In fact, even if each level of nesting only multiplies the simulation time by a factor of 2, we get an exponential blow-up.

and Kilian [26] suggests that certain (related) protocols may be zero-knowledge under bounded-simultaneity concurrent composition. Still, to the best of own knowledge, such a technically-appealing result has not been proven before. We prove the result by using a rather straightforward approach, which nevertheless requires careful implementation. We stress that not every zero-knowledge protocol remains zero-knowledge under bounded-simultaneity concurrent composition: Goldreich and Krawczyk presented a simple (constant-round) protocol that is zero-knowledge, but parallel execution of two instances of it is not zero-knowledge [16].

Combining the two techniques, we show that the GK-protocol is concurrent zero-knowledge under the timing model. This is shown by using the abovementioned decomposition, and applying the bounded-simultaneity simulator to the sub-schedules while incorporating the parallel-composition simulator inside of it. Note that the bounded-simultaneity simulator handles the special case in which each sub-schedule contains a single copy, and does so by employing the single-copy simulator. Capitalizing on the high-level similarity of the parallel-composition simulator and the single-copy simulator, we just need to extend the bounded-simultaneity simulator by incorporating the former simulator in it. (Our presentation of the bounded-simultaneity simulator uses terminology that makes this extension quite easy.)

1.5 Zero-knowledge versus ϵ-knowledge

Recall that ϵ-knowledge means that for every noticeable function (i.e., a reciprocal of a positive polynomial) $\epsilon : \mathsf{N} \to (0,1]$ there exists a simulator working in time $\mathrm{poly}(n/\epsilon(n))$ that produces output that is ϵ-indistinguishable from the one of a real interaction, where ϵ-indistinguishability of the ensembles $\{X_\alpha\}$ and $\{Y_\alpha\}$, means that for every efficient procedure (e.g., a polynomial-time algorithm) D,

$$|\Pr[D(\alpha, X_\alpha)=1] - \Pr[D(\alpha, Y_\alpha)=1]| \;<\; \epsilon(|\alpha|) + \mu(|\alpha|)$$

where μ is a negligible function.[11]

Indeed, as mentioned in [11], ϵ-knowledge does provide some level of security. However, this level of security is lower than the one offered by the standard notion of zero-knowledge, and more so when compared to simulators with bounded knowledge tightness (as discussed above; cf. [14, Sec. 4.4.4.2]). Furthermore, unlike zero-knowledge, the notion of ϵ-knowledge is not closed under sequential composition (i.e., t sequential executions of a ϵ-knowledge protocol yield a $t \cdot \epsilon$-knowledge (rather than ϵ-knowledge) protocol).

Expected polynomial-time simulators versus ϵ-knowledge. The above discussion applies also to the comparison of ϵ-knowledge and zero-knowledge via expected polynomial-time simulators (rather than via strict polynomial-time simulators). Furthermore, simulation by an expected polynomial-time simulator implies ϵ-knowledge simulator (running in strict time inversely proportional to ϵ).[12] The converse does not hold (e.g., consider a prover that, for $i = 1, 2...$, with probability 2^{-i} sends the result of a $\mathrm{BPTime}(2^{2^i})$-complete computation).

[11] Indeed, the standard notion of computational indistinguishability [19, 27] is a special case obtained by setting $\epsilon \equiv 0$.

[12] This can be seen by truncating all runs of the original simulator that exceed its expected running-time by a factor of $1/\epsilon$ (or so).

1.6 Organization

In Section 2, we recall some basic notions as well as review the *GK-protocol* (i.e., the five-round zero-knowledge proof system of Goldreich and Kahan [15]). In Section 3 we prove that the GK-protocol remains zero-knowledge under parallel composition. In Section 4 we prove that the GK-protocol remains zero-knowledge under bounded-simultaneity concurrent composition. In Section 5, we augment the GK-protocol with adequate `time-out` and `delay` mechanisms, and prove that the resulting protocol is concurrent zero-knowledge under the timing model.

2 Background

Recall that an *interactive proof system* for a language L is a (randomized) protocol for two parties, called verifier and prover, allowing the prover to convince the verifier to accept any common input in L (*completeness*), while guaranteeing that no prover strategy may fool the verifier to accept inputs not in L (*soundness*), except than with negligible probability. The prescribed verifier strategy is always required to be probabilistic polynomial-time. Furthermore, like in other application-oriented works, we focus on prescribed prover strategies that can be implemented in probabilistic polynomial-time given adequate auxiliary input (e.g., an NP-witness in case of NP-languages). Recall that the latter refers to the prover prescribed for the completeness condition, whereas (unlike in argument systems [5]) soundness must hold no matter how powerful the cheating prover is. Zero-knowledge is a property of some prover-strategies. Loosely speaking, it means that anything that is feasibly computable by (possibly improperly) interacting with the prover, can be feasibly computable without interacting with the prover. That is, the most basic definition of zero-knowledge (of a prover P w.r.t L) requires that, for every feasible verifier strategy V^*, there exists a feasible simulator M^* so that the following two probability ensembles are computationally indistinguishable:

1. $\{\langle P, V^* \rangle(x)\}_{x \in L} \stackrel{\text{def}}{=}$ the output of V^* when interacting with P on common input $x \in L$; and
2. $\{M^*(x)\}_{x \in L} \stackrel{\text{def}}{=}$ the output of M^* on input $x \in L$.

2.1 Expected polynomial-time simulation and black-box simulation

As discussed in the introduction, we allow the simulator to run in *expected* probabilistic polynomial-time (rather than *strict* probabilistic polynomial-time), but require it to be implemented by a universal machine that get oracle access to the strategy V^*. See [14, Sec. 4.3.1.6] (resp., [14, Sec. 4.5.4.2] and [1]) for further discussion of the first (resp., second) issue.

Definition 2.1 (black-box zero-knowledge):

- Next message function: *Let B be an interactive Turing machine, and x, z, r be strings representing a common-input, auxiliary-input, and random-input, respectively. Consider the function $B_{x,z,r}(\cdot)$ describing the messages sent by machine B so that $B_{x,z,r}(\overline{m})$ denotes the message sent by B on common-input x, auxiliary-input z, random-input r, and sequence of incoming messages $\overline{m}$. For simplicity, we assume that the output of B appears as its last message.*
- Black-box simulator: *We say that an* expected *probabilistic polynomial-time oracle machine M is a* black-box simulator for the prover P and the language L *if for every polynomial-time interactive machine B, every probabilistic polynomial-time oracle machine D, every*

polynomial $p(\cdot)$, all sufficiently large $x \in L$, and every $z, r \in \{0,1\}^*$:

$$\left|\Pr\left[D^{B_{x,z,r}}(\langle P, B_r(z)\rangle(x))=1\right] - \Pr\left[D^{B_{x,z,r}}(M^{B_{x,z,r}}(x))=1\right]\right| < \frac{1}{p(|x|)}$$

where $B_r(z)$ denotes the interaction of machine B with auxiliary-input z and random-input r.

- *We say that P is* black-box zero knowledge *if it has a black-box simulator.*

As discussed by Canetti *et. al.* [8], the above definition is too restrictive for definition of (unbounded) composition, where the adversary B may invoke a (polynomial) number of sessions with P but this polynomial is not a-priori known. One solution is to consider for each polynomial a different universal simulator that can handle all adversaries that invoke at most a number of sessions (with B) that is bounded by that polynomial. For simplicity, we adopt this solution here.

2.2 Parallel and concurrent zero-knowledge and the timing model

The definition of parallel and concurrent zero-knowledge are derived from Definition 2.1 by considering appropriate adversaries (i.e., adversarial verifiers). In case of parallel zero-knowledge, we consider adversaries that simultaneously initiate a polynomial number of copies of P and interact with this multitude of copies in a synchronized way (i.e., send their i^{th} message to all copies at the same time). In case of concurrent zero-knowledge, we consider adversaries that initiate a polynomial number of copies of P and interleave their interaction with this multitude of copies in an arbitrary way. In case of concurrent zero-knowledge under the timing model, the interleaving of executions by the adversary must satisfying the timing model. (Without loss of generality we may assume that the adversary never violates the time-out condition; it may instead send an illegal message at the latest possible adequate time.) Furthermore, without loss of generality, we may assume that all the adversary's messages are delivered at the latest possible adequate time.[13]

2.3 The Goldreich–Kahan (GK) protocol [15]

Loosely speaking, the Goldreich–Kahan (GK) proof system for Graph 3-Colorability (G3C) proceeds in four steps:

1. The verifier commits to a challenge (i.e., sequence of edges).
2. The prover commits to a sequence of values (i.e., the colors of each vertex under several random relabelings of a 3-coloring of the graph).
3. The verifier decommits (to the edge-sequence).
4. If the verifier has properly decommits then the prover decommits to a subset of the values as indicated by the decommitted challenge. Otherwise the prover sends nothing.

   (Specifically, the query is a sequence of edges each associated with an independently selected 3-coloring of the graph, and the prover decommits to the values corresponding to the endpoints of the i^{th} edge with respect to the i^{th} committed coloring.)

[13] In general, the prover may be modified so that a time-out condition applied to the verifier's message is always followed by at least an similar delay of the next prover message. (Indeed, this may slow down some executions in which the verifier is honest, but never slows them down by more than can be caused by a cheating verifier.) We comment that this modification is unnecessary in our protocol (of Section 5), since it already satisfies the above convention.

The specific details of each of these steps are not important to our analysis. Still, for sake of clarity, we reproduced these details in Construction 2.2 (below). We highlight a couple of points that are relevant to the analysis: Firstly, the prover's commitment is via a commitment scheme that is (only) computationally-hiding, and so commitments to different values are (only) computationally-indistinguishable (which considerably complicates the analysis; cf. [15]). Secondly, the verifier's commitment is via a commitment scheme that is (only) computationally-binding, and so it is (only) infeasible for it to properly decommits in two different way (which slightly complicates the analysis).

Implementation Details: The Goldreich–Kahan protocol [15] utilizes two "dual" commitment scheme (see terminology in [14, Sec. 4.8.2]). The first commitment scheme, denoted C, is used by the prover and has a perfect-binding property. For simplicity, we assume that this scheme is non-interactive, and denote by $C(v)$ a random variable representing the output of C on input v (i.e., a commitment to value v).[14] The second commitment scheme, denoted $\mathcal{C}$, is used by the verifier and has a perfect-hiding property. Such a scheme must be interactive, and we assume that it consists of the receiver sending a random index, denoted α, and the committer responds by applying the randomized process $\mathcal{C}_\alpha$ to the value it wishes to commit to (i.e., $\mathcal{C}_\alpha(v) = \mathcal{C}(\alpha, v)$ represents a commitment to v relative to the receiver's message α). Consequently, Step 1 in the high-level description is implemented by Steps P0 and V1 below.

Construction 2.2 (The GK zero-knowledge proof for G3C):

Common Input: *A simple* (3-colorable) *graph* $G=(V,E)$.

Let $n \stackrel{\text{def}}{=} |V|$, $V = \{1,...,n\}$, *and* $t \stackrel{\text{def}}{=} 2n \cdot |E|$.

Auxiliary Input to the Prover: *A 3-coloring of* G, *denoted* ψ.

Prover's preliminary step (P0): *The prover invokes the commit phase of the perfectly-hiding commitment scheme, which results in sending to the verifier a message* α.

Verifier's commitment to a challenge (V1): *The verifier uniformly and independently selects a sequence of* t *edges,* $\overline{e} \stackrel{\text{def}}{=} ((u_1,v_1),...,(u_t,v_t)) \in E^t$, *and sends to the prover a random commitment to these edges. Namely, the verifier uniformly selects* $s \in \{0,1\}^{\text{poly}(n)}$, *and sends* $c \stackrel{\text{def}}{=} \mathcal{C}_\alpha(\overline{e}, s)$ *to the prover.*

Motivating Remark: *At this point the verifier is committed to a sequence of* t *edges.* (This commitment is of perfect secrecy.)

Prover's commitment step (P1): *The prover uniformly and independently selects a sequence of* t *random relabeling of the 3-coloring* ψ, *and sends the verifier commitments to the color of each vertex under each of these colorings. That is, the prover uniformly and independently selects* t *permutations,* $\pi_1,...,\pi_t$, *over* $\{1,2,3\}$, *and sets* $\phi_j(v) \stackrel{\text{def}}{=} \pi_j(\psi(v))$, *for each* $v \in V$ *and* $1 \leq j \leq t$. *It uses the perfectly-binding commitment scheme to commit itself to the colors of each of the vertices according to each 3-coloring. Namely, the prover uniformly and independently selects* $r_{1,1},...,r_{n,t} \in \{0,1\}^n$, *computes* $c_{i,j} = C(\phi_j(i), r_{i,j})$, *for each* $i \in V$ *and* $1 \leq j \leq t$, *and sends* $c_{1,1},...,c_{n,t}$ *to the verifier.*

[14] Non-interactive perfectly-binding commitment schemes can be constructed using any one-way *permutation*. In case one wishes to rely here only on the existence of one-way *functions*, one may need to use Naor's two-round perfectly-binding commitment scheme [25]. This calls for minor modification of the description below.

Verifier's decommitment step (V2): *The verifier decommits the sequence* $\overline{e} = ((u_1, v_1), ..., (u_t, v_t))$ *to the prover. Namely, the verifier send* $(s, \overline{e})$ *to the prover.*

Motivating Remark: *At this point the entire commitment of the verifier is revealed. The verifier now expects to receive, for each* j*, the colors assigned by the* j^{th} *coloring to vertices* u_j *and* v_j (the endpoints of the j^{th} edge in $\overline{e}$).

Prover's partial decommitment step (P2): *The prover checks that the message just received from the verifier is indeed a valid revealing of the commitment* c *made by the verifier at Step* (V1) *(i.e., it checks that* $c = \mathcal{C}_\alpha(\overline{e}, s)$ *indeed holds). Otherwise the prover halts immediately. Let us denote the sequence of* t *edges, just revealed, by* $(u_1, v_1), ..., (u_t, v_t)$*. The prover reveals* (to the verifier)*, for each* j*, the* j^{th} *coloring of vertices* u_j *and* v_j*, along with appropriate decommitment information. Namely, the prover sends to the verifier the sequence of four-tuples*

$$(r_{u_1,1}, \phi_1(u_1), r_{v_1,1}, \phi_1(v_1)), ..., (r_{u_t,t}, \phi_t(u_t), r_{v_t,t}, \phi_t(v_t))$$

Verifier's local decision step (V3): *The verifier checks whether, for each* j*, the values in the* j^{th} *four-tuple constitute a correct revealing of the commitments* $c_{u_j,j}$ *and* $c_{v_j,j}$*, and whether the corresponding values are different. Namely, upon receiving* $(r_1, \sigma_1, r'_1, \tau_1)$ *through* $(r_t, \sigma_t, r'_t, \tau_t)$*, the verifier checks whether for each* j*, it holds that* $c_{u_j,j} = C(\sigma_j, r_j)$*,* $c_{v_j,j} = C(\tau_j, r'_j)$*, and* $\sigma_j \neq \tau_j$ (and both are in $\{1, 2, 3\}$)*. If all conditions hold then the verifier accepts. Otherwise it rejects.*

Goldreich and Kahan proved that Construction 2.2 constitute a (constant-round) zero-knowledge interactive proof for Graph 3-Colorability [15]. (We briefly review their simulator below.) Our first goal is to show that the zero-knowledge property (of Construction 2.2) is preserved under parallel composition. We later extend the result to yield concurrent zero-knowledge under the timing-model.

High level description of the simulator used in [15]. The simulator (using oracle access to the verifier's strategy) proceeds in three main steps:

The Scan Step: The simulator emulates Steps (P0)–(V2), by using commitments to dummy values in Step (P1), and obtained the verifier's decommitment for Step (V2), which may be either proper or not. In case of improper decommitment the simulator *outputs the partial transcript just generated and* `halts`.

The Approximation Step: For technical reasons (discussed in [15]), the simulator next approximates the probability that the first scan ended with a proper decommitment. (This is done by repeated trials, each as in the first scan, until some polynomial number of proper decommitments is found.)

The Generation Step: Using the (proper) decommitment information, obtained in the first scan, the simulator repeatedly tries to generate a full transcript by emulating Steps (P0)–(P2), using commitments to "pseudo-colorings" that do not "violate the coloring conditions imposed by the decommitted edges". The number of trials is inversely proportional to the probability estimated in the approximation step.

3 Simulator for the Parallel Case

Recall that the GK-protocol proceeds in four (abstract) steps:

1. The verifier commits to a challenge.

   (The actual implementation is by two rounds/messages.)
2. The prover commits to a sequence of values.

   (The challenge specifies a subset of these values.)
3. The verifier decommits (either properly or not).
4. Pending on the verifier's proper decommitment, the prover decommits to the corresponding values.

The basic approach towards simulating this protocol (without being able to answer a random challenge) is to first run the first three steps with prover-commitments to arbitrary (dummy) values, obtaining the challenge, and then rewind to Step 2 and make a prover-commitment that passes this specific challenge (alas no other challenge). In case the verifier always decommits properly, this allows to easily simulate a full run of the protocol. In case the verifier always decommits improperly, things are even easier since in this case we only need to simulate Steps 1–3. The general case is when the verifier decommits with some probability. Intuitively, this can be handled by outputting the initial transcript of Steps 1–3 in case it contains an improper decommitment, and repeatedly trying to produce a full passing transcript (as in the first case) otherwise. Difficulties arise in case the probability of proper verifier decommitment is small but not negligible and furthermore when it depends (in a negligible way) on whether the prover commits to dummy or to "passing" values. Indeed, the focus of [15] is on resolving this problem (and their basic approach is to approximate the probability of proper decommitment in case of dummy values, and keep trying to produce a full passing transcript for at most a number of times that is inversely proportional to the latter probability).

The problem we face here is more difficult: several (say n) copies of the protocol are executed in parallel and the verifier may properly decommit in some of them but not in others. Thus, there are 2^n possible configurations of proper/improper decommitment in these n copies, and we certainly cannot insist on rewinding until we obtain again the same configuration (e.g., they may be all equally likely). Instead, referring to the n probabilities corresponding to each of the n copies performing proper decommitment, we add additional rewinding in which we try to obtain a proper decommit from all copies that have at least as high a probability as the copies that actually performed proper-decommitment in the initial simulated run. Once this is obtained, we try to generate a parallel run in which only copies having at least as high a probability (but not necessarily all of them!) properly decommit. Furthermore, in order not to skew the distribution (towards high proper-decommitment probabilities), we insist on having at least one copy with a corresponding probability as low as some copy in the initial run. As one may expect, the problem is to pick thresholds such that the above discussion may be efficiently implemented. We start by clarifying the above discussion.

3.1 A high level description

As is clear from the above discussion, the following basic notions are central to our analysis: An execution of a copy is said to properly decommit if the verifier message in Step 3 is a valid decommitment to its (i.e., the verifier's) commitment in Step 1. In the first part of the simulation, we use prover's commitments to arbitrary values, which are referred to as commitment to dummy values.

Later we use commitments to values that will pass for a certain challenge (which is understood from the context). These are called commitment to passing values.[15] In addition, we also refer to the following more complex notions and notations:

- Let p_i denote the probability that the verifier properly decommits in the i^{th} copy of the parallel run, when Step 2 is played with commitment to dummy values. (When using other commitments (e.g., passing commitments) the probability of proper decommitment may be p'_i such that $|p'_i - p_i|$ is negligible.)

- We use a sequence of thresholds, denoted $t_1, ..., t_n$, that will be determined (probabilistically) on the fly such that

  1. $t_j \in (2^{-(j+1)}, 2^{-j})$,
  2. no p_i lies in the interval $[t_j \pm (1/9n) \cdot 2^{-j}]$.

  Such t_j's exist and t_j can be found when given approximations of all p_i's up-to $(1/9n) \cdot 2^{-j}$ (or so). We also define $t_0 \stackrel{\text{def}}{=} 1$, and so $p_i \leq t_0$ for all i. (We assume for simplicity that every p_i is greater than 2^{-n}, and so every p_i lies in one of the intervals $(t_j, t_{j-1}]$.)

- For such t_j's, define $T_j = \{i : p_i > t_j\}$. (Indeed, $T_0 = \emptyset$ and $T_n = \{1, ..., n\}$.)

  Membership in T_j can be determined in time $poly(n) \cdot 2^j$, since t_j was selected to be sufficiently far-away from all the p_i's (i.e., $|t_j - p_i| = \Omega(2^{-j}/n)$).

- Let E_j denote the event that the verifier properly decommits to some copy in $T_j \setminus T_{j-1}$ but to no copy outside T_j, and let $q_j = \Pr[E_j]$. Note that $q_j \leq n \cdot t_{j-1}$ (since E_j mandates one of $|T_j \setminus T_{j-1}| \leq n$ events, each occuring with probability at most t_{j-1}). However, q_j may be much smaller than $t_j < t_{j-1}$.

  Since $\{1, ..., n\} = T_n \supseteq T_{n-1} \cdots \supseteq \cdots T_1 \supseteq T_0 = \emptyset$, whenever the verifier properly decommits in some copy, one of the events E_j (for $j \geq 1$) must hold. Otherwise (i.e., whenever the verifier decommits improperly in all copies), we say that event E_0 holds.

We now turn to the simulator, which generalizes the one in [15]. All approximations referred to below are quite good w.v.h.p. (i.e., with $1 - 2^{-n}$ each approximation is within a factor of $(1 + (1/\text{poly}(n))$ of the corresponding value). Loosely speaking, after fixing the verifier's coins (at random), the simulator proceeds as follows (while using the residual verifier strategy as a black-box):

Step S0: Obtain the verifier's commitments (of Step 1) in the n parallel copies.

Step S1: The purpose of this step is to generate an index $j \in \{0, 1, ..., n\}$ with distribution corresponding the probability that event E_j holds for a random parallel execution of the protocol, as well as to determine the sets T_j and T_{j-1}. This will allow to determine (in subsequent steps) whether or not event E_j holds.

The above goal is achieved by first simulating Steps 2–3 of the (parallel execution of the) protocol, while using (in Step 2) commitments to dummy values. Based on the verifier's decommitments in Step 3 (of the parallel execution), we determine the set $I \subseteq [n]$ of copies

[15] Recall that in the actual implementation, challenges correspond to sequences of t edges (over the vertex-set $\{1, 2..., n\}$), and the prover commits to a sequence of $t \cdot n$ values in $\{1, 2, 3\}$ (i.e., a block of n values per each edge). For a given edge sequence (i.e., a challenge), a passing sequence of values is one in which (for every i) the values assigned to the i^{th} block are such that the endpoints of the i^{th} edge (in the challenge) are assigned a random pair of distinct elements.

in which the verifier has properly decommitted. Next, we determine an appropriate sequence $t_1, ..., t_j$ of thresholds such that event E_j holds. Finally, using t_{j-1} and t_j, we determine for each i whether or not $p_i > t_j$ (i.e., $i \in T_j$) and whether or not $p_i > t_{j-1}$ (i.e., $i \in T_{j-1}$).

Indeed, the above description (especially of the second part) is way too laconic. We need (and will) describe (below) how to implement it within time $\text{poly}(n) \cdot 2^j$.

Step S2: For each copy $i \in T_j$, we wish to obtain the challenge committed to in Step 1, while working within time $\text{poly}(n) \cdot 2^j$. This is done by rewinding and re-simulating Steps 2–3 for at most $\text{poly}(n) \cdot 2^j$ times, while again using (in Step 2) commitments to dummy values.

Step S3: For technical reasons, analogously to [15], we next obtain a good (i.e., constant factor) approximation of $q_j = \Pr[E_j]$. This approximation, denoted $\tilde{q}_j$, will be obtained within expected time $\text{poly}(n)/q_j$ by repeated rewinding and re-simulating Steps 2–3. (Specifically, we continue till we see some fixed polynomial number (say n^5) of occurrences of the event E_j.)

Step S4: We now try to generate a simulation of Steps 2–3 in which event E_j occurs. However, unlike in previous simulations, here we use (in Step 2) commitments to values that pass the challenges that we have obtained. This will allow us to simulate also Step 4, and complete the entire simulation.

Specifically, we make at most $\text{poly}(n)/\tilde{q}_j$ tries to rewind and re-simulate Steps 2–3, while using (in Step 2 of each copy in T_j) commitments to values that pass the corresponding challenge (which we obtained in Step S2). If the verifier answers (for Step 3) fit event E_j then we proceed to simulate Step 4 in the obvious manner. Otherwise, we rewind and try again (but never try more than $\text{poly}(n)/\tilde{q}_j$ times).

Pending on the ability to properly implement Step S1, we observe that:

1. The (overall) expected running time is polynomial, because each attempt (in Steps S2, S3, and S4) is repeated for a number of times that is inversely proportional to the probability of entering this repeated-attempts step. Specifically, each of these steps is repeated at most $(\text{poly}(n)/\tilde{q}_j) \approx (\text{poly}(n)/q_j)$ times (use $q_j = O(n \cdot 2^{-j})$ for Step S2), whereas j is selected with probability q_j.

2. The computational-binding property of $\mathcal{C}$ implies that we rarely get into trouble in Step S4; that is, only with negligible probability will it happen that in Step S4 the verifier properly decommit to a value different from the one to which it has properly decommitted in Step S2.

3. Since the probabilities of verifier's proper-decommitment (in Step 3) are almost unaffected by the prover's commitments (of Step 2) and since passing commitments look like commitments to truly valid values, the simulated interaction is computationally indistinguishable (cf. [19, 27]) from the real one.

3.2 Setting the thresholds and implementing Step S1

One naive approach is to try to use fixed thresholds such as $t_j = 2^{-j}$. However, this may not allow to determine (for a given i) with high probability and within time $\text{poly}(n) \cdot 2^j$ whether or not p_i is smaller than t_j. (The reason being that p_i may be very close to 2^{-j}; e.g., $|p_i - 2^{-j}| = 2^{-2n}$.)

Instead, the t_j's will be selected in a more sophisticated way so that they are approximately as above (i.e., $t_j \approx 2^{-j}$) but also far enough (i.e., at distance at least $2^{-j}/9n$) from each p_i. This will

allow to determine (with high probability) and within time $\text{poly}(n) \cdot 2^j$ whether or not p_i is smaller than t_j. The question is how to set the t_j's so that they are appropriately far from all p_i's. Since the p_i's are unknown probabilities (which we can only approximate), it seems infeasible to come-up with a deterministic setting of the t_j's. Indeed, we will settle for a probabilistic setting of the t_j's (provided that this setting is independent of other events).

Recall that Step S1 calls for the setting of $t_1, ..., t_j$ such that event E_j holds, where whether or not event E_j holds depends on t_j and t_{j-1}. Furthermore, it is important that the setting of t_{j-1} in case event E_j holds be the same as the setting of t_{j-1} in case event E_{j-1} holds. Moreover, recalling that the setting of t_j must be performed in time $\text{poly}(n) \cdot 2^j$, we cannot afford to set all t_k's whenever we set a specific t_j. Still, we provide below an adequate threshold-setting process. We start with the following key procedure.

Procedure $T(j,n)$, returns $t_j \in ((3/4) \pm (1/4)) \cdot 2^{-j} \subseteq (2^{-(j+1)}, 2^{-j})$:

1. Approximate each p_i up-to $(1/9n) \cdot 2^{-j}$ (with error probability at most $2^{-n}/n$). Actually, for large p_i's (e.g., $p_i > 2^{-j+2}$) we only approximate p_i up-to a factor of 2 of its true value. This is done by $\text{poly}(n) \cdot 2^j$ rewindings and re-simulations of Steps 2–3 of the protocol.[16] Call the resulting approximations, a_i's.

2. Determine $K \leftarrow \{k \in \{-n/2, ..., +n/2\} : (\forall i)\ a_i \notin ((3/4) + (k/2n) \pm (1/4n)) \cdot 2^{-j}\}$. Note that K is not empty, because each a_i can rule out at most one element of K (whereas $|\{-n/2, ..., +n/2\}| = n+1$ and they are only n values of i).

3. Select an arbitrary (say at random or the first) $k \in K$. Output $t_j = ((3/4) + (k/2n)) \cdot 2^{-j}$.

By construction, $|t_j - a_i| \geq (1/4n) \cdot 2^{-j}$, for all i's. For each i, if $p_i > 2^{-j+1}$ then (using $t_j < 2^{-j}$) definitely $p_i > t_j + 2^{-j}$ holds, and so $|p_i - t_j| > 2^{-j}$ follows. Otherwise (i.e., $p_i \leq 2^{-j+1}$ for this i), with probability at least $1 - 2^{-n-\log_2 n}$, we have $|a_i - p_i| \leq (1/9n) \cdot 2^{-j}$. In this (high probability) case, p_i does not fall in the interval $t_j \pm (1/9n) \cdot 2^{-j}$, because $|p_i - t_j| \geq |a_i - t_j| - |a_i - p_i| \geq ((1/4n) - (1/9n)) \cdot 2^{-j} > (1/9n) \cdot 2^{-j}$. We conclude that, with probability at least $1 - 2^{-n}$, no p_i falls in the interval $t_j \pm (1/9n) \cdot 2^{-j}$.

Implementation of Step S1: Recall that the purpose of Step S1 is to generate an index $j \in \{0, 1, ..., n\}$ with distribution corresponding the probability that event E_j holds (for a random parallel run of the protocol), as well to determine the thresholds $t_1, ..., t_j$, and using these to determine for every $i = 1, ..., n$, whether or not $i \in T_j$ and whether or not $i \in T_{j-1}$. We thus start by generating a random run, and next determine all necessary objects with respect to it.

1. *Generating a reference run*: Simulate Steps 2–3 of the (parallel execution of the) protocol, while using (in Step 2) commitments to dummy values. Based on the verifier's decommitments in Step 3 (of the parallel execution), determine the set $I \subseteq [n]$ of copies in which the verifier has properly decommitted.

2. *Determining the event E_j occuring in the reference run, as well as the sets T_j and T_{j-1}*:
   Case of empty I: Set $j = 0$ and $T_j = T_{j-1} = \emptyset$.

[16] Note that in such a number of attempts, we can first distinguish w.v.h.p between the case $p_i > 2^{-j+2}$ and the case $p_i < 2^{-j+1}$. In the former case we approximate p_i up-to a factor of 2, in the latter case we approximate it up-to an additive deviation of $(1/9n) \cdot 2^{-j}$, and in the intermediate case any of these will do. Recall that approximating a probability p to within factor 2 can be done in a number of trials proportional to $1/p$ (which for $p = \Omega(2^{-j})$ means $O(2^j)$ trials). Similarly, approximating a probability p up-to an additive deviation of q can be done in a number of trials proportional to p/q^2 (which for $p = O(2^{-j})$ and $q = \Omega(2^{-j}/n)$ means $O(2^j n^2)$ trials).

Case of non-empty I: Set $t_0 = 1$ and $T_0 = \emptyset$. For $j = 1, ..., n$ do

(a) $t_j \leftarrow T(j, n)$. (We stress that the value of t_j is set obliviously of I.)

(b) Determine the set T_j by determining, for each i, whether or not $p_i > t_j$. We use approximations to each p_i (as in procedure $T(j, n)$ above), and rely on $|p_i - t_j| > (1/9n) \cdot 2^{-j}$.

(c) Decide whether or not event E_j holds for the reference run, by using T_{j-1} (of the previous iteration) and T_j (just computed). Recall that event E_j holds if $I \subseteq T_j$ but $I \not\subseteq T_{j-1}$.

(d) If event E_j holds then exit the loop with the current value of j as well as with the values of T_j and T_{j-1}. Otherwise, proceed to the next iteration.

Since we have assumed that $(\forall i)\ p_i > 2^{-n}$, some event E_j must hold.[17]

A key point in the analysis is that the values of the T_k's, as determined by Step S1 (i.e., $T_0, ..., T_j$), are independent of the value of j. Of course, which of the T_k's were determined does depend on the value of j. Thus, we may think of Step S1 as of an efficient implementation of the mental experiment in which all T_k's are determined, next j is determined accordingly (analogously to the above), and finally one outputs T_j and T_{j-1} for subsequent use.

3.3 A detailed description of the simulator

For sake of clarity we present a detailed description of the simulator, before turning to its analysis. Recall that we start by selecting and fixing the verifier's coins at random.

Step S0: We simulate the parallel execution of Step 1 (i.e., Steps P0 and V1 of Construction 2.2) as follows. First, acting as the real prover in Step P0, we randomly generate messages $\alpha^1, ..., \alpha^n$ (one per each copy). Invoking the verifier (as per Step V1), while feeding it with $\alpha^1, ..., \alpha^n$, we obtain its n commitments, $c^1, ..., c^n$, for the n copies.

Step S1: As explained in Section 3.2, we determine (for a random reference run) the index j for which E_j holds, as well as the sets T_j and T_{j-1}. Recall that this (and specifically procedure $T(\cdot, \cdot)$) involves $\text{poly}(n) \cdot 2^j$ rewindings and re-simulations of Steps 2–3, while using commitments to dummy values. Each rewinding is performed as in Step S2 below.

In case $j = 0$, we may skip all subsequent steps, and just output the reference run produced in the current step.

Step S2: For each copy $i \in T_j$, we wish to obtain the challenge (edge-sequence) committed to in Step 1, while working within time $\text{poly}(n) \cdot 2^j$. This is done by rewinding and re-simulating Steps 2–3 (i.e., Steps P1 and V2 of Construction 2.2) for $\text{poly}(n) \cdot 2^j$ times, while using commitments to dummy values. (Actually, we may as well do the same for all i's (regardless whether $i \in T_j$ or not), but we are guaranteed to succeed only for i's in T_j. Furthermore, we may work on all i's concurrently.)

Specifically, each rewinding attempt proceeds as follows:

1. Generate n sequences of random commitments to the dummy value 0. That is, for every (copy) $i = 1, ..., n$, select uniformly $r^i_{1,1}, ..., r^i_{n,t} \in \{0,1\}^n$, and compute $\overline{c}^i \stackrel{\text{def}}{=} (c^i_{1,1}, ..., c^i_{n,t})$, where $c^i_{k,\ell} = C(0, r^i_{k,\ell})$.

[17] Removing this assumption enables the situation that no event E_j occurs. This may happen only if $p_i \leq t_n < 2^{-n}$, for every $i \in I$. But the probability that the reference run corresponds to such a set I is at most $\sum_{i:p_i<2^{-n}} p_i < n \cdot 2^{-n}$.

2. Feeding the verifier with (the n prover commitments) $\overline{c}^1, ..., \overline{c}^n$, obtain the verifier's n (Step 3) responses, denoted $(s^1, \overline{e}^1), ..., (s^n, \overline{e}^n)$.
3. For every properly decommitted copy (i.e., i such that $c^i = \mathcal{C}_{\alpha_i}(s^i, \overline{e}^i)$), store the corresponding challenge (i.e., the edge sequence $\overline{e}^i$).

(Note that it is unlikely that we will obtain two conflicting proper decommitments to the same verifier commitment c^i.)

Step S3: For technical reasons, analogously to [15], we next obtain a good approximation of $q_j = \Pr[E_j]$. This approximation, denoted $\tilde{q}_j$, will be obtained within expected time $\text{poly}(n)/\tilde{q}_j$ by repeated rewinding and re-simulating Steps 2–3 (i.e., Steps P1 and V2 of Construction 2.2). Specifically, we repeat the following steps until we obtain n^5 occurrences of event E_j.

1. Perform Items 1 and 2 as in Step S2.
2. Let $I' = \{i : \mathcal{C}_{\alpha_i}(s^i, \overline{e}^i) = c^i\}$. If I' fits event E_j (i.e., $I' \subseteq T_j$ and $I' \not\subseteq T_{j-1}$) then increment the "success counter" by one unit. (We proceed to the next iteration only if the "success counter" is still smaller than n^5.)

Suppose we have obtained n^5 successes while making τ trials. Then we set $\tilde{q}_j = n^5/\tau$.

Step S4: We now try to generate a simulation of Steps 2–3 (i.e., Steps P1 and V2 of Construction 2.2) in which event E_j occurs. However, unlike in previous simulations, here we use (in Step 2) commitments to values that pass the challenges that we have obtained. This will allow us to simulate also Step 4, and complete the entire simulation. Specifically, we make at most $\text{poly}(n)/\tilde{q}_j$ tries to rewind and re-simulate Steps 2–3, while using (in Step 2 of each copy in T_j) commitments to values that pass the corresponding challenge (which we obtained in Step S2). Each attempt proceed as follows:

1. Generate n sequences of random commitments to passing values (for copies in T_j and dummy values otherwise). Specifically, suppose that $i \in T_j$ (or more generally that we have obtained (in Step S2) a proper decommitment to c^i), and denote by $((u^i_1, v^i_1), ..., (u^i_t, v^i_t))$ the value of the decommitted challenge (edge sequence). Then, for every $\ell = 1, ..., t$, select uniformly $r^i_{1,\ell}, ..., r^i_{n,\ell} \in \{0,1\}^n$ and $a^i_\ell \neq b^i_\ell \in \{1,2,3\}$, and compute $c^i_{u^i_\ell,\ell} = C(a^i_\ell, r^i_{u^i_\ell,\ell})$, $c^i_{v^i_\ell,\ell} = C(b^i_\ell, r^i_{v^i_\ell,\ell})$, and $c^i_{k,\ell} = C(0, r^i_{k,\ell})$ for $k \notin \{u^i_\ell, v^i_\ell\}$. Let $\overline{c}^i \stackrel{\text{def}}{=} (c^i_{1,1}, ..., c^i_{n,t})$.

   For $i \notin T_j$ (or for i's for which we failed in Step S2), we produce $\overline{c}^i \stackrel{\text{def}}{=} (c^i_{1,1}, ..., c^i_{n,t})$ as in (Item 1 of) Step S2.
2. Feeding the verifier with (the prover's commitments) $\overline{c}^1, ..., \overline{c}^n$, obtain the verifier's n (Step 3) responses, denoted $(s^1, \overline{e}^1), ..., (s^n, \overline{e}^n)$. Let $I' = \{i : \mathcal{C}_{\alpha_i}(s^i, \overline{e}^i) = c^i\}$ denote the set of copies that have properly decommitted (in the current attempt). If I' does not fit event E_j (i.e., $I' \not\subseteq T_j$ or $I' \subseteq T_{j-1}$) then we abort this attempt. That is, we proceed only if I' fits event E_j.
3. For every properly decommitted copy (i.e., $i \in I'$), we provide a proper decommitment (as per Step 4). This complete a full simulation of such a copy, whereas improperly committed copies are simulated by their transcript so far. Specifically, ignoring the rare case of conflicting proper decommitments, a proper decommitment to copy $i \in I' \subseteq T_j$ must use the same challenge (edge sequence) as (found in Step S2 and) used in Item 1 (of the current attempt). Then, for every $i \in I'$ and $\ell = 1, ..., t$, we merely provide the 4-tuple $(r^i_{u^i_\ell,\ell}, a^i_\ell, r^i_{v^i_\ell,\ell}, b^i_\ell)$, where $((u^i_1, v^i_1), ..., (u^i_t, v^i_t))$ is the corresponding challenge.

Indeed, this answer (like the prover's answer in Step 4) passes the verifier's check (since $a^i_\ell \neq b^i_\ell \in \{1,2,3\}$, $c^i_{u^i_\ell,\ell} = C(a^i_\ell, r^i_{u^i_\ell,\ell})$, and $c^i_{v^i_\ell,\ell} = C(b^i_\ell, r^i_{v^i_\ell,\ell})$).

In the rare case in which a conflicting proper decommitment is received, we proceed just as in case event E_j does not occur.

For technical reasons, we modify the above simulation procedure by never allowing it to run more than 2^n steps. (This is easily done by introducing an appropriate step-count (which is implemented in linear or almost-linear time and so does not affect our running-time analysis).)

3.4 A detailed analysis of the simulator

Lemma 3.1 (Simulator's running-time): *The simulator runs in expected polynomial-time.*

Proof: The key observation is that each repeated attempt to produce something is repeated for a number of times that is inversely proportional to the probability that we try this attempt at all. This reasoning is applied with respect to each of the main steps (i.e., Steps S1, S2, S3 and S4). Specifically:

- For Step S1: Recall that event E_j occurs in the reference run (generated at the onset of Step S1), with probability q_j. Letting $Q \stackrel{\text{def}}{=} T_j \setminus T_{j-1}$, we have $q_j \leq |Q| \cdot \max_{i \in Q}\{p_i\} \leq n \cdot t_{j-1} < n \cdot 2^{-(j-1)}$. Also, with probability at least $1 - 2^{-n}$, Step S1 correctly determines j. Pending on the latter (overwhelmingly high probability) event, the expected number of steps conducted in Step S1 is

$$\sum_{j=0}^{n} q_j \cdot (\mathrm{poly}(n) \cdot 2^j) \;<\; \sum_{j=0}^{n} (n \cdot 2^{-(j-1)}) \cdot (\mathrm{poly}(n) \cdot 2^j) \;=\; \mathrm{poly}(n) \tag{1}$$

  Relaying on the fact that the simulator never runs for more than 2^n steps, we cover also the highly unlikely case (in which Step S1 determines a wrong j).

  The same reasoning applies to Step S2. That is, again assuming that Step S1 correctly determines j, the expected number of steps made in Step S2 is as in Eq. (1).

- For Step S3: Assuming that $\tilde{q}_j = \Theta(q_j)$, the expected number of steps made in Step S3 is $\sum_{j=0}^{n} q_j \cdot (\mathrm{poly}(n)/\tilde{q}_j) = \mathrm{poly}(n)$. The above assumption holds with probability at least $1 - 2^{-n}$, and otherwise we relay on the fact that the simulator never runs for more than 2^n steps. The same reasoning applies to Step S4.

Thus, the overall expected running-time is polynomial (and this is proven without relying on any security properties of the commitment schemes). ■

Lemma 3.2 (Simulator's output distribution): *Assuming that the verifier's commitment scheme* (i.e., $\mathcal{C}$) *is computationally-binding and that the prover's commitment scheme* (i.e., C) *is computationally-hiding, the output of the simulator is computationally indistinguishable from the real parallel interaction.*

Recall that the assumption that $\mathcal{C}$ is perfect-hiding and C is perfectly-binding is used in establishing the soundness of the GK-protocol (as a proof system).

Proof: For sake of clarity of the analysis, one may consider an imaginary simulator that goes on to determine all t_j's (rather than determining only part of them as in Item 2 of Step S1). We may assume that all approximations made by the simulator are sufficiently good; that is, in Step S1 the simulator correctly determines j as well as T_j and T_{j-1}, and in Step S3 it obtains $\tilde{q}_j = \Theta(q_j)$. (Indeed, the assumption holds with probability at least $1 - 2^{-n}$.)

Next, we consider three unlikely events in the simulation:

1. In Step S2, the simulator fails to obtain a proper decommitment of some $i \in T_j$. This may happen only with exponentially vanishing probability, because we keep trying for $\mathrm{poly}(n) \cdot 2^j$ times and each time a proper decommitment (for i) occurs with probability $p_i > t_j \geq 2^{-(j+1)}$.

2. In Step S4, the simulator fails to generate a simulation in which event E_j holds. This may happen only with negligible probability, because in order for this to happen event E_j should occur in Step S1 and then we should fail to obtain it in Step S4 although we try $\mathrm{poly}(n)/\tilde{q}_j = O(\mathrm{poly}(n)/q_j)$ times and each time event E_j occurs with probability q'_j that is negligiblly close to q_j (because C is computationally-hiding; cf. [15, Clm. 3]). (Note that q_j refers to the probability that event E_j occurs for a dummy commitment, whereas q'_j refers to its probability for a "passing" commitment.) Thus, the probability of this failure is bounded above by

$$\sum_{j=0}^{n} q_j \cdot (1 - q'_j)^{\mathrm{poly}(n)/q_j} \tag{2}$$

   Letting $\Delta_j \stackrel{\mathrm{def}}{=} q_j - q'_j$, we consider two cases (cf. [15, Clm. 2]): in case $\Delta_j \leq q_j/2$, the corresponding term is exponentially vanishing, whereas in case $\Delta_j \geq q_j/2$ we simply bound the corresponding term by $q_j \leq 2\Delta_j$. Thus, in both cases, we obtain that each term in Eq. (2) is negligible.[18]

3. In Step S4, the simulator obtains a proper decommit to some copy that is different from the proper decommitment obtained for the same copy in Step S2. (In such a case, the simulator's output will definitely look wrong.) However, the hypothesis that $\mathcal{C}$ is computationally-binding implies that this bad event occurs only with negligible probability.

We conclude that, except with negligible probability, the simulator produces an output that looks syntactically fine. Finally, the hypothesis that C is computationally-hiding is used to demonstrate that the simulator's output is computationally indistinguishable from a random transcript of the real interaction. The details are analogous to the proof of [15, Clm. 4]: First we prove that the probabilities of each E_j event is about the same (i.e., differ by a negligible amount) in the simulation's output and in the real interaction. Next we focus on each likely E_j event and prove that the conditional spaces for it are indistinguishable. We capitalize on the fact that a non-negligible difference in the unconditional space must translates to a non-negligible difference on some likely E_j, and that for likely E_j the simulation runs in strict polynomial-time.[19] ■

Combining Lemmas 3.1 and 3.2, we obtain

Theorem 3.3 *The* (constant-round) *GK-protocol is zero-knowledge under parallel composition.*

[18] In the first case, we have $(1 - q'_j)^{n/q_j} \leq (1 - (q_j/2))^{n/q_j} \leq \exp(-n/2)$. Thus, Eq. (2) is upper-bounded by $\sum_{j=0}^{n} \max(2\Delta_j, \exp(-n/2))$, which is a negligible function.

[19] An alternative approach may be to derive, in the contradiction argument, an expected polynomial-time algorithm that violates the hiding property of C, and to derive from it (via truncating long runs) a strict polynomial-time algorithm that violates the hypothesis that C is computationally-hiding.

Recall that the GK-protocol is a proof system for $\mathcal{NP}$ (with exponentially vanishing soundness error) [15].

3.5 An Extension

We relax the parallel execution condition to concurrent execution of polynomially-many copies that satisfy the following two conditions:

C1: No copy enters Step 2 before all copies complete Step 1.

C2: No copy enters Step 4 before all copies complete Step 3.

In other words, the concurrent execution proceeds in three phases:

Phase 1: All copies perform Step 1 (in arbitrary order).

Phase 2: All copies perform Steps 2 and 3 (in arbitrary order except for the obvious local timing condition (i.e., each copy performs Step 3 after it has completed Step 2)).

Phase 3: All copies perform Step 4 (in arbitrary order).

Our treatment of parallel executions extends to the above (concurrent) case. The reason being that the simulator treats Steps 2–3 as one unit, and so the fact that these steps may be interleaving among copies is of no importance. Specifically, Step S0 of the extended simulator refers to Phase 1 (rather than to Step 1), its Steps S1–S3 refer to Phase 2 (rather than to Steps 2–3), and its Step S4 refers to Phases 2–3 (rather than to Steps 2–4).

4 Simulator for the case of Bounded-Simultaneity

Recall that the GK-protocol proceeds in four steps:

1. The verifier commits to a challenge.
2. The prover commits to a sequence of values.
3. The verifier decommits (either properly or not).
4. Pending on the verifier's proper decommitment, the prover decommits to the corresponding values.

Here we consider concurrent executions in which up-to w copies of the GK-protocol run simultaneously at any given time, where w may be any fixed constant.

4.1 Motivation

The case of $w = 1$ corresponds to sequential composition, and it is well-known that any zero-knowledge protocol maintains its security in this case. So let us turn (as a warm-up) to the case of $w = 2$. Trying to use the single-session simulator of [15] in this case, we encounter the following problem: when we try to deal with the simulation of one copy, the verifier may invoke another copy. A natural thing to do is to suspend our dealing of the first copy, and apply the single-session simulator to the second copy. The good news are that the verifier cannot initiate yet another copy (before it terminates either the first or second ones, since this would have violated the bounded-simultaneity condition (for $w = 2$)). Instead, one of two things will happen (eventually[20]):

[20] Actually, in a few steps.

1. The verifier may execute Step 3 in the second copy, in which case we make progress on treating the second copy (and will eventually complete a simulation of it, which would put us back in the one-session case).
2. The verifier may execute Step 3 in the first copy, in which case we make progress on treating the first copy. For example, if we were trying to get the decommitment value for the first copy and we just got it, we may abandon the treatment of the second copy and proceed by rewinding the first copy.

Thus, in each of these cases, we make progress. Intuitively, the cost of dealing with two simultaneous copies is that we have to invoke the single-session simulator (for the second copy) per each operation of the single-session simulator (for the first copy). It is not surprising that we can similarly deal with any constant number of simultaneous copies, and do so within time $T(n)^w$, where $T(n)$ is the time complexity of the single-session simulator and w is the bound on simultaneity.

4.2 The actual simulation

Indeed, although appealing, the above suggestion requires careful implementation.

In correspondence to the three main steps of the single-copy simulator (cf. Section 2.3), we introduce three recursive procedures: `Scan`, `Approx` and `Generate`. Each of these procedures tries to handle a single copy (just as done by the corresponding step of the single-copy simulator), while making recursive calls when encountering a Step 2 message of some other copy.[21] The recursive call will take place before execution this Step 2, and execution of this Step 2 will be the first thing that the invoked procedure will do. The three procedure maintain (and pass along) the state of currently handled copies as well as related auxiliary information. In particular, $\overline{h}$ will denote a partial transcript of the (concurrent) execution, and $\overline{a}$ will denote a list of currently active copies together with auxiliary information regarding each of them (e.g., decommitment information obtained in previous related runs). For sake of clarity, although the following is implicit in $\overline{h}$, we will also pass explicitly the identity of the copy that is responsible for the current procedure call (i.e., the copy that encountered Step 2). The (simulator's) main program merely consists of a special invocation of `Generate` with empty history (i.e., $\overline{h} = \overline{a} = \lambda$) and a fictitious copy index. Throughout the rest of the description we fix the (deterministic) adversarial verifier (although we only use black-box access to it).

4.2.1 The specification of the procedures

Let us first elaborate on the structure of the auxiliary information $\overline{a}$, which consists of *records*, each corresponding to some encountered copy of the protocol. Each record consists of three fields:

1. *The verifier decommitment field* indicates whether (among related runs) the first encounter of Step 3 (i.e., the verifier's decommitment) of this copy was proper or improper (i.e., the type of decommitment), and in the former case the field includes also the value of the decommitment. That is, if non-empty, the field stores a pair (X, v), where $X \in \{\texttt{proper}, \texttt{improper}\}$ is a decommitment *type* and v is a decommitment value (which is meaningful only in case $X = \texttt{proper}$). This field (of the record of the i^{th} copy) is filled-up according to the answer returned by some invocation of $\texttt{Scan}(\overline{h}, \cdot, i)$.

2. *The decommitment probability field* holds an approximation of the probability that an invocation with parameters as the one that filled-up the first field, actually turns out returning

[21] This is no typo; we do mean Step 2, not Step 1.

exactly this information. That is, suppose that the first field of record i (i.e., the record of the i^{th} copy) was filled-up according to the answer returned by $\texttt{Scan}(\overline{h}, \overline{a}, i)$, which resulted with a decommitment of type $X \in \{\texttt{proper}, \texttt{improper}\}$. Then the second field of record i should hold an approximation of the probability that $\texttt{Scan}(\overline{h}, \overline{a}, i)$ returns with an answer that encodes the same type of decommitment of copy i. (Jumping ahead, we hint that $\texttt{Scan}(\overline{h}, \overline{a}, i)$ may return with a decommitment to some other copy, and so the sum of the two probabilities corresponding to the two types is not necessarily 1.)

3. *The prover decommitment field* encodes the decommitment information corresponding to the prover's commitment in Step 2. This field (of the record of the i^{th} copy) is filled-up at the up-front of the execution of $\texttt{Generate}(\overline{h}, \overline{a}', i)$, which follows the invocation of $\texttt{Scan}(\overline{h}, \overline{a}, i)$, where $\overline{a}'$ is $\overline{a}$ augmented by the verifier decommitment information of copy i and the prover's commitment is performed so to passed the latter.

As hinted above, the fields are filled-up in the order they appear above (i.e., the verifier decommitment field is filled-up first). In reading the following specifications, it may be instructive to consider the special case of a single copy (in which case failure never occurs and $j = i$ always holds).

Specification of $\texttt{Scan}(\overline{h}, \overline{a}, i)$: This call produces a prefix of a "pseudorandom" execution transcript that extends the prefix $\overline{h}$, and returns some related information. The transcript is *pseudorandom* in the sense that it is computationally indistinguishable from a (prefix of a random) real continuation of $\overline{h}$ (by the adversary interacting with copies of the prover).[22] The extended transcript is truncated (i.e., the extended prefix ends) at the first point where one of the following holds:

1. *Progress*: This is a case where the (extended) execution reaches Step 3 of some copy j (possibly but not necessarily $j = i$) so that the first field of record j is empty. In this case, the procedure should return the index j as well as the decommitment information (provided in the current execution of Step 3 of copy j). That is, the answer is a pair (j, y), where j is a index of a copy and y is a decommitment information (which may be either proper or improper).

2. *Failure*: This is a case where the (extended) execution reaches Step 3 of some copy j so that the first field of record j encodes a decommitment type different than the one occuring in the current extension. That is, the first field of record j encodes decommitment type $X \in \{\texttt{proper}, \texttt{improper}\}$, whereas in the current execution Step 3 of copy j has a decommitment type different from X (i.e., opposite to X). (In fact $j \neq i$ will always hold.) In this case, the procedure should return a special `failure` symbol.

Specification of $\texttt{Approx}(\overline{h}, \overline{a}, X, i)$: Always returns an approximation of the probability that $\texttt{Scan}(\overline{h}, \overline{a}, i)$ answers with a pair (i, y) such that y has type $X \in \{\texttt{proper}, \texttt{improper}\}$. The approximation is required to be correct to within a factor of 2 with probability at least $1 - 2^{-n}$.

Specification of $\texttt{Generate}(\overline{h}, \overline{a}, i)$: This call produces a prefix of a pseudorandom execution transcript that extends the prefix $\overline{h}$, and returns either this extension or related information. The notion of pseudorandom is the same as in case of `Scan`, and the extended transcript is truncated at the first point where one of the following holds:

[22] The reader may wonder as to what will happen in case $\overline{h}$ itself is not consistent with any prefix of such a real interaction. The answer is that the extended execution will always be truncated before this fact becomes evident (i.e., we never perform Step 4 of a copy unless Step 2 of that copy was performed in a passing manner).

1. *Failure*: Exactly as in the specification of `Scan`, except that here $j = i$ is possible too.
2. *Progress*: Here there are a few sub-cases:
   (a) This is a case where the (extended) execution reaches Step 3 of some copy j so that the first field of record j is empty. This sub-case is handled exactly as the Progress Case of `Scan`. (Unlike in `Scan`, here $j = i$ cannot not possibly hold.)
   (b) This is a case where the (extended) execution reaches Step 4 of copy i. In this case, the procedure returns the currently extended transcript (including the execution of Step 4 of copy i), along with a corresponding update to the auxiliary information $\overline{a}$.
   (We comment that this sub-case is not really necessary; we choose to add it in order to make the description and the analysis slightly easier.)
   (c) The execution reaches its end (i.e., the adversary terminates). In this sub-case, we act exactly as in the previous one.

Note that we did not specify (above) with what probability `Generate` (or `Scan`) should make progress. In fact, unlike in the presentation of the basic simulator, here `Generate` does not make progress almost always (not even in the case of a single copy), but rather makes progress with probability that is close to the one approximated by the corresponding `Approx`-call. Thus, `Generate` is actually a generation-attempt, and the repetition of this attempt is made by the higher level invocator (rather than in the procedure itself).

4.2.2 The implementation of the procedures

We refer to the notion of *passing*, as defined and used in Section 3. Recall that a *passing commitment* is a sequence of (prover's) commitments to values that pass for the corresponding challenge (encoded in the first field of the corresponding copy): see Footnote 15.

We start with the description of `Generate` (although `Generate`$(\cdot,\cdot,i)$ is invoked after `Scan`$(\cdot,\cdot,i)$). We note that `Generate`$(\overline{h},\overline{a},i)$ is always invoked when the first field in the i^{th} record in $\overline{a}$ is not empty (but rather encodes some decommitment, of arbitrary proper/improper type), and the third field is empty (and will be filled-up at the very beginning of the execution).

Procedure `Generate`$(\overline{h},\overline{a},i)$: Initializes $\overline{h}' = \overline{h}$ and $\overline{a}' = \overline{a}$, generates a passing commitment for (Step 2 of) copy i, and augments $\overline{h}'$ and $\overline{a}'$ accordingly. Specifically:

1. The procedure generates a random sequence of values, denoted $\overline{v}$, that pass the challenge described in the first field of the i^{th} record of $\overline{a}$. That is, $\overline{v}$ may be arbitrary if the said field encodes an improper decommitment; but in case of proper decommitment, $\overline{v}$ must pass with respect to the challenge value encoded in that field.
2. The procedure generates a random sequence of (prover's) commitments, denoted $\overline{c}$, to $\overline{v}$, augments $\overline{h}'$ by $\overline{c}$, and augments $\overline{a}'$ by placing the corresponding decommitment information in the third field of the i^{th} record.

Next, the procedure proceeds in iterations according to the following cases that refer to the next step taken in the concurrent execution.

Step 1 by some (new) copy: Just augment $\overline{h}'$ accordingly (and proceed to the next iteration).

Step 2 by some copy j (certainly $j \neq i$)**:** We consider two cases depending on whether or not $\overline{a}'$ contains the verifier's decommitment information for copy j (i.e., whether or not the first field of the j^{th} record is non-empty).

1. In case $\overline{a}'$ does contain such information, we generate a corresponding passing commitment (i.e., a prover commitment to values that pass w.r.t challenge encoded in the first field of the j^{th} record), augment $\overline{h}'$ and $\overline{a}'$ accordingly, and proceed to the next iteration. (Specifically, analogously to the up-front activity for (Step 2 of) the i^{th} copy, the third field in the j^{th} record of $\overline{a}'$ is augmented by the decommitment information corresponding to this prover commitment, and $\overline{h}'$ is augmented by the commitment itself.)
2. In case $\overline{a}'$ does not contain such information (i.e., the first field of the j^{th} record is empty), we proceed as follows:
   (a) Invoke $\texttt{Scan}(\overline{h}', \overline{a}', j)$, and consider its answer, which is either `failure` or a progress pair (k, y). In case of progress, determine the type $X \in \{\texttt{proper}, \texttt{improper}\}$ of the decommitment information y (with respect to the corresponding Step 1 commitment in $\overline{h}'$).
   (b) If either the answer is `failure` or (it is a progress pair (k, y) with) $k \neq j$ then `return` with the very same answer (k, y).
   (Here, in case of progress, $k \neq i$ must hold.)
   (c) If the above answer is (a progress pair (k, y)) with $k = j$ (and y is a decommitment of type X), then we let $\tilde{q} \leftarrow \texttt{Approx}(\overline{h}', \overline{a}', X, j)$, and update the j^{th} record of $\overline{a}'$ placing (X, y) in the first field and $\tilde{q}$ in the second field. (Actually, it suffices to place (X, v) in the first field, where v is the decommitment value included in the decommitment information y.)
   (We comment that in case $X = \texttt{improper}$, we could have skipped all subsequent substeps, and used instead the extended transcript generated by the above invocation of `Scan`, provided that `Scan` were modified to return this information as well. However, avoiding this natural modification makes the extension in Section 5 more smooth.)
   (d) Next, repeatedly invoke $\texttt{Generate}(\overline{h}', \overline{a}', j)$ until getting a progress, but not more than $\text{poly}(n)/\tilde{q}$ times. (We will show that only with negligible probability can it happen that all calls return `failure`.) If all attempts have returned `failure` then `return failure`, otherwise act according to the sub-cases of the (first) progress answer:
      i. If the "progress answer" provides a pair (k', y') (certainly $k' \neq j$ as well as $k' \neq i$), then (analogously to sub-step 2b) `return` with the very same answer (k', y').
      ii. If the "progress answer" provides an updated history $\overline{h}''$ (together with updated auxiliary information $\overline{a}''$) such that $\overline{h}''$ is not terminating (i.e., but rather ends with execution of Step 4 of copy j), then update $\overline{h}'$ and $\overline{a}'$ (i.e., $\overline{h}' \leftarrow \overline{h}''$ and $\overline{a}' \leftarrow \overline{a}''$), and proceed to the next iteration.
      iii. If the "progress answer" provides an updated history $\overline{h}''$ (together with updated auxiliary information $\overline{a}''$) such that $\overline{h}''$ is terminating (i.e., $\overline{h}''$ ends with the verifier halting), then `return` with the very same answer.

   The handling of this case (i.e., Step 2 for a copy for which the first field is empty) is the the most involved part of the procedure.

Step 3 by copy i: Just as the first sub-case in the next case (i.e., Step 3 by some copy $j \neq i$ with a non-empty first field).

Step 3 by some copy $j \neq i$: We consider two cases depending on whether or not $\overline{a}'$ contains the verifier's decommitment information for copy j (i.e., the first field of the j^{th} copy is not empty).

1. In case $\overline{a}'$ does contain such information, we consider sub-cases according to the relation of the contents of the the first field of the j^{th} copy, denoted $(X, \cdot)$, and the current answer of the verifier.
   (a) If the decommitment type of the current Step 3 (of the j^{th} copy) fits X then we just augment $\overline{h}'$ accordingly (and proceed to the next iteration).
   (b) Otherwise (i.e., the decommitment type of the current Step 3 does not fit X), `return failure`.
2. In case $\overline{a}'$ does not contain such information (i.e., the first field of the j^{th} copy is empty), obtain the relevant decommitment information from the adversary (it may be either an improper or proper decommitment), and `return` (as progress) with this information only. That is, `return` with (j, y), where y encodes the decommitment information just obtained from the adversary.

Step 4 by some copy j (possibly $j = i$): We will show that this case may happen only in case the corresponding (Step 2) prover commitment is passing and $\overline{a}'$ contains the corresponding decommitment (in the third field of the j^{th} record). Using the latter prover's decommitment information, we emulate Step 4 in the straightforward manner (and augment $\overline{h}'$ accordingly). In case $j = i$, `return` with the current $\overline{h}'$ and $\overline{a}'$ (otherwise proceed to the next iteration).

Note that Step 2 of copy i is handled up-front. In case of a single copy i, the above procedure degenerates to the basic handling of Steps 2–4 of copy i. In the fictitious invocation of `Generate` (i.e., with empty $\overline{h}$ and a fictitious i), only the handlings of Steps 2–4 for copies $j \neq i$ are activated (whereas, in handling Step 2, sub-steps 2b and 2(d)i are never activated). We now turn to procedure `Scan`, which is similar to `Generate`, except for its handling of the steps of copy i.

Procedure `Scan`$(\overline{h}, \overline{a}, i)$: Initializes $\overline{h}' = \overline{h}$ and $\overline{a}' = \overline{a}$, generates a dummy commitment for (Step 2 of) copy i, and augments $\overline{h}'$ accordingly. (Specifically, the procedure generates a random sequence of commitments, $\overline{c}$, to dummy values, and augments $\overline{h}'$ by $\overline{c}$.) Next, the procedure proceeds in iterations according to the following cases that refer to the next step taken in the concurrent execution.

Step 1 by some (new) copy: As in `Generate`.

Step 2 by some copy j (certainly $j \neq i$): As in `Generate`.

(We comment that unlike in sub-step 2b of `Generate`, here $k = i$ is possible. Also, here sub-case 2(d)iii cannot occur in sub-step 2d.)

Step 3 by copy i: Obtain the relevant decommitment information from the adversary (it may be either an improper or proper decommitment), and `return` (as progress) with this information. That is, `return` with (i, y), where y encodes the decommitment information just obtained from the adversary.

Step 3 by some copy $j \neq i$: As in `Generate`.

Step 4 by some copy $j \neq i$: As in `Generate`.

Note that we never reach Step 4 of copy i (and that Step 2 of copy i is handled up-front).

Procedure $\texttt{Approx}(\overline{h}, \overline{a}, X, i)$**:** This procedure merely invokes $\texttt{Scan}(\overline{h}, \overline{a}, i)$ until it obtains $m = \text{poly}(n)$ invocations that return a pair that is a decommitment of type X for copy i, and returns the fraction of m over the number of tries. Specifically, the procedure proceeds as follows:

Set $\texttt{cnt}_{\text{total}} = \texttt{cnt}_{\text{succ}} = 0$.
Until $\texttt{cnt}_{\text{succ}} = m$ do
  increment $\texttt{cnt}_{\text{total}}$ (unconditionally),
  $(j, y) \leftarrow \texttt{Scan}(\overline{h}, \overline{a}, i)$,
  increment $\texttt{cnt}_{\text{succ}}$ if and only if $j = i$ and y is of type X.
Output: $m/\texttt{cnt}_{\text{total}}$.

4.2.3 Comments about the analysis

It is quite straightforward to show that the procedure `Approx` satisfies its specification. Ignoring the exponentially vanishing probability that any single approximation (by the procedure `Approx`) is off by more than a factor of 2, we may bound the total expected running-time by using the recursive structure of the simulation. (We start with bounding the running-time, because we will have to use this bound in analyzing the output of the simulator.) Towards the running-time analysis, it is useful to pass among the procedures also the corresponding path in the tree of recursive calls. For example, instead of saying that $\texttt{Scan}(\overline{h}, \overline{a}, i)$ invokes $\texttt{Generate}(\overline{h}', \overline{a}', j)$, we may say that $\texttt{Scan}(\overline{h}, \overline{a}, i; \overline{p})$ invokes $\texttt{Generate}(\overline{h}', \overline{a}', j; (\overline{p}, i))$. Bounded-simultaneity implies that the depth of the recursive tree is at most w, which is the key to the entire analysis of the running-time.

Running-time analysis. Considering oracle calls to the adversary's strategy as atomic steps, the expected running-time of $\texttt{Scan}(\overline{h}, \overline{a}, i; \overline{p})$ (resp., $\texttt{Generate}(\overline{h}, \overline{a}, i; \overline{p})$) is dominated by the time spent by the recursive calls invoked by $\texttt{Scan}(\overline{h}, \overline{a}, i; \overline{p})$ (resp., $\texttt{Generate}(\overline{h}, \overline{a}, i; \overline{p})$). Such calls are made only when handling Step 2 of a copy with no verifier decommitment information. Each of these handlings consists of first invoking $\texttt{Scan}(\overline{h}', \overline{a}', j; (\overline{p}, i))$ and, pending on its not returning `failure`, invoking `Approx` and `Generate` on the same input. Specifically, the latter are invoked only if $\texttt{Scan}(\overline{h}', \overline{a}', j; (\overline{p}, i)) = (j, \cdot)$. In particular, $\texttt{Approx}(\overline{h}', \overline{a}', X, j; (\overline{p}, i))$ invokes $\texttt{Scan}(\overline{h}', \overline{a}', j; (\overline{p}, i))$ for an expected number of times that is inversely proportional to the probability that $\texttt{Scan}(\overline{h}', \overline{a}', j; (\overline{p}, i))$ answers with a type X decommitment to copy j, and $\texttt{Generate}(\overline{h}', \overline{a}', j; (\overline{p}, i))$ is invoked for the same (absolute) number of times. That is, letting $\texttt{Scan}'(\overline{h}', \overline{a}', j) \stackrel{\text{def}}{=} (k, X)$ if $\texttt{Scan}(\overline{h}', \overline{a}', j)$ answers with a type X decommitment to copy k, we conclude that the expected number of recursive calls made by $\texttt{Scan}(\overline{h}, \overline{a}, i; \overline{p})$ (resp., $\texttt{Generate}(\overline{h}, \overline{a}, i; \overline{p})$) when handling a Step 2 message of Copy j is

$$\sum_{X \in \{\texttt{proper}, \texttt{improper}\}} \Pr[\texttt{Scan}'(\overline{h}', \overline{a}', j) = (j, X)] \cdot \frac{\text{poly}(n)}{\Pr[\texttt{Scan}'(\overline{h}', \overline{a}', j) = (j, X)]} \;=\; \text{poly}(n) \qquad (3)$$

The key point is that all these recursive calls (of, say, $\texttt{Scan}(\overline{h}, \overline{a}, i; \overline{p})$) have the longer path $(\overline{p}, i)$. Thus, each node in the (depth w) tree of recursive-calls has an expected polynomial number of children, and so the expected size of the tree is upper-bounded by $\text{poly}(n)^w$. It follows that, for probabilistic polynomial-time adversaries and constant w, the simulation terminates in expected polynomial-time.

Output distribution analysis. We start the analysis (of the output) by justifying the discarding of the (remote) possibility that during the (polynomial-time) simulation we ever get two conflicting

proper decommitments to the same verifier commitment. (In fact, the above functional description suggests this assumption, although formally it is not needed in the functional description.) Here the polynomial bound on the expected running-time is used to derive a contradiction to the computational-binding property of the verifier's commitment.

Next, we establish that in sub-step 2d of the handling of a Step 2 message, it rarely happens that all invocations of `Generate` return `failure` (i.e., this bad event occurs with negligible probability). Here the polynomial bound on the expected running-time is used to bound the number of bad events in a union bound, where each event (i.e., failure in sub-step 2d) occurs with negligible probability.

At this point, we get to the straightforward but tedious task of establishing that the main procedures (i.e, `Scan` and `Generate`) satisfy their corresponding specifications. This is proven by induction on the recursive execution. Once this is established, we look at the initial (fictitious) invocation of `Generate`, which cannot possibly return with `failure`, and conclude that the simulator's output is computational indistinguishable from a real interaction of the cheating verifier with copies of the prover.

5 Simulation under the Timing Model

Recall that the timing assumptions refer to two constants, Δ and ρ, such that Δ is an upper bound on the message handling-and-delivery time, and $\rho \geq 1$ is a bound on the relative rates of the local clocks. Specifically, each real-time period of Δ units elapses Δ' units of time on the local clock, where $\Delta/\rho \leq \Delta' \leq \rho\Delta$. For simplicity, we may assume without loss of generality that $\Delta/\rho \leq \Delta' \leq \Delta$ (i.e., that all clocks are at least as slow as the real time).[23]

5.1 The Time-Augmented GK-protocol

Recall that the GK-protocol proceeds in four abstract steps, but the actual implementation of the first step consists of the prover sending a preliminary message that is used as basis of the verifier actual commitment. Thus, the GK-protocol is actually a 5-round protocol starting with a prover message. We augment this protocol with the following time-driven instructions, where all times are measured according to the prover's clock starting at the time of the invocation of the prover's program:

1. The prover time-outs Step 1 after $\Delta_1 \stackrel{\text{def}}{=} 2\Delta$ units of time (as measured on its clock).

   (By the timing assumption, this does not disrupt honest operation, because 2Δ real units of time suffice for the delivery of a message from the prover to the verifier and back.)

2. The prover delays its execution of Step 2 to time $\Delta_2 \stackrel{\text{def}}{=} \rho \cdot \Delta_1 + \Delta$. That is, it sends its message exactly when its clock shows that Δ_2 units of time have elapsed.

3. The prover time-outs Step 3 after $\Delta_3 \stackrel{\text{def}}{=} \Delta_2 + \Delta$ units of time.

   (Note that $\Delta_3 = (2\rho + 2) \cdot \Delta$.)

4. The prover delays its execution of Step 4 to time $\Delta_4 \stackrel{\text{def}}{=} \rho \cdot \Delta_3 + \Delta$.

We comment that, compared to Dwork *et. al.* [11], we are making a slightly more extensive use of the time-out and delay mechanisms: Specifically, they only used the last two items and did so while setting $\Delta_3 = 3\Delta$ and $\Delta_4 = \rho\Delta_3$. On the other hand, our use of the time-out and delay mechanisms is less extensive than the one suggested by Section 1.4: We only guarantee that for two copies that

[23] We comment that although our formulation looks different than the one in [11], it is in fact equivalent to it.

start at the same time, Step 2 (resp., Step 4) in each copy starts after Step 1 (resp., Step 3) is completed in the other copy, but we do not guarantee anything about the relative timing of Steps 2 and 3. Relying on special properties of the GK-protocol (as analyzed in Section 3.5), we can afford doing so, whereas the description in Section 1.4 is generic and refers to any c-round protocol.[24]

5.2 The Simulation

As mentioned in the introduction, the simulation relies on a decomposition of any schedule that satisfies the timing model into sub-schedules such that each sub-schedule resembles parallel composition, whereas the relations among the sub-schedules resembles bounded-simultaneity concurrent composition. In fact, we can prove something stronger:

Claim 5.1 *Consider an arbitrary scheduling of concurrent sessions of the time-augmented GK-protocol that satisfy the timing assumption. Place a session in* block *i if it is invoked within the real-time interval $((i-1)\cdot\Delta, i\cdot\Delta]$. Then, for every i:*

1. *Each session in block i terminates Step 1 by real-time $i\cdot\Delta+\rho\Delta_1$, starts Step 2 after real-time $i\cdot\Delta+\rho\Delta_1$, terminates Step 3 by real-time $i\cdot\Delta+\rho\Delta_3$, and starts Step 4 after real-time $i\cdot\Delta+\rho\Delta_3$.*

2. *The number of blocks that have a session that overlaps with some session in block i is at most $14\rho^3$. That is, the number of $j\neq i$ such that there exists a time t, a session s in block i, and a session s' in block j such that s and s' are both active at time t is at most $14\rho^3$.*

The first item corresponds to Conditions C1 and C2 in Section 3.5, and the second item corresponds to bounded-simultaneity.[25]

Proof: The latest and slowest possible session in block i is invoked by real-time $i\cdot\Delta$, and takes $\rho\Delta$ units of real-time to measure Δ local-time units. It follows that such a session terminates Step 1 (resp., Step 3) by real-time $i\cdot\Delta+\rho\cdot\Delta_1$ (resp., $i\cdot\Delta+\rho\cdot\Delta_3$). On the other hand, the earliest and fastest possible session in block i is invoked after real-time $(i-1)\cdot\Delta$, and takes Δ units of real-time to measure Δ local-time units. It follows that such a session starts Step 2 (resp., Step 4) after real-time $(i-1)\cdot\Delta+\Delta_2 = i\cdot\Delta+\rho\Delta_1$ (resp., $(i-1)\cdot\Delta+\Delta_4 = i\cdot\Delta+\rho\Delta_3$). The first item follows.

For the second item, note that the earliest possible session in block i is invoked after real-time $(i-1)\cdot\Delta$, whereas the latest and slowest possible session in block i terminates by real-time $i\cdot\Delta+\rho\Delta_4+\Delta = (i+1)\cdot\Delta+\rho\cdot(2\rho^2+2\rho+1)\cdot\Delta$. Thus, each block resides in a time interval of length $(2\rho^3+2\rho^2+\rho+2)\cdot\Delta$, and therefore may overlap at most $2\cdot(2\rho^3+2\rho^2+\rho+2)\leq 14\rho^3$ other blocks. ■

5.2.1 Combining the simulation techniques – the perfect case

Given the above claim, we extend the simulation of Section 4 by showing how that simulator can handle blocks of "practically parallel" sessions rather than single copies (which may be viewed as

[24] However, in the typical case where $\rho\approx 1$, the difference between the various time-augmentations of the GK-protocol is quite small.

[25] The second item is actually stronger than bounded-simultaneity, because it upper-bounds the total number of blocks that overlap with a given block (rather than upper-bounding the number of blocks that are (simultaneously) active at any given time).

"singleton blocks"). To motivate the final construction, we consider first the special case in which each block is a perfect parallel composition of some sessions.

The key to the extension is to realize that all that changes is the types of verifier decommitment events (corresponding to Step 3). Recall that in case of a single session, there were two possible events (i.e., proper and improper decommitment), and these were the two decommitment types we have considered. Here, for m parallel copies (of some block), we may have 2^m possible events corresponding to whether each of the m copies is proper or improper. However, the decommitment types we consider here are (not these 2^m events but rather) the $n+1$ events considered in Section 3: the events $E_0, E_1, ..., E_n$, where event E_j holds if all the properly decommitting sessions (in the current run) have proper-decommitment probability above the threshold $t_j \approx 2^{-j}$ but not all these sessions have proper-decommitment probability above the threshold $t_{j-1} \approx 2^{-(j-1)}$. Indeed, E_0 is the event that all sessions have improperly decommitted in the current run. (It is important that the number of decommitment types is bounded by a polynomial; this will be reflected when trying to extend the analysis captured in Eq. (3).)

Given the new notion of decommitment types, the three procedures of Section 4 (`Scan`, `Approx` and `Generate`) are extended by using the corresponding operations in Section 3. We stress that, in case of progress, `Scan` (as well as the first progress case in `Generate`) returns the decommitment information, which includes the indication of whether each session has properly decommitted, but not the decommitment type. The latter will be determined as in Section 3 (which is far more complex than the trivial case handled in Section 4, where decommitment type equals the decommitment indicator bit). The decommitment type (rather than the sequence of decommitment indicators) is what matters in much of the rest of the activities of the modified procedures.

We focus on the most interesting modifications to the main procedures (`Scan` and `Generate`), and ignore straightforward extensions (which apply also to other steps):

1. *The handling of Step 2 messages by a block j with a non-empty first information field* is analogous to the treatment in the original procedure, and we merely wish to clarify what this means here. The point is that the first field of block j encodes a decommitment type E_k as well as decommitment information for all sessions that properly decommit with probability at least $t_k \approx 2^{-k}$. The prover commitment produced here is designed to pass with respect to these decommitment values. (The same applies to the initial actions in `Generate`.)

2. *The handling of Step 2 messages by a block j with an empty first information field* (i.e., the only case that invokes recursive calls). The following sub-steps correspond to the sub-steps in the original procedures (`Scan` and `Generate`):

   (a) We invoke `Scan` with a block index j (rather than with a copy index), and consider its answer which is either `failure` or a progress pair (k, y), where k is a block index, and y is a list of decommitments corresponding to the various copies of block k. We refer to the above invocation of `Scan` as to the initial one, and note that many additional invocations (with the same parameters) will take place in handling the current step.

   If (the initial invocation of) `Scan` returned with a progress pair (k, y) such that $k = j$, then we turn to the complex task of determining the decommitment type E_ℓ (which holds with respect to y) as well as the corresponding sets T_ℓ and $T_{\ell-1}$. This is done analogously to the main part of Step S1 (of Section 3), which needs to be implemented in the current context. In particular, the implementation of Step S1 calls for the approximation of the probabilities (denoted p_i's in Section 3) each of the sessions properly decommits. This, in turn, amounts to multiple executions of Steps 2–3 of these sessions, which in

our case should be handled by multiple invocation of $\texttt{Scan}(\cdot,\cdot,j)$. (If $k \neq j$ then the following activity will not be conducted here, but rather be conducted by the instance that invoked $\texttt{Scan}(\cdot,\cdot,k)$.)

Let $I \subseteq [n]$ denote the set of sessions in which the verifier has properly decommitted in y. (Recall we are in the case where the initial invocation of $\texttt{Scan}(\overline{h}', \overline{a}', j)$ has returned the progress pair (j, y).) Our objective is to determine the corresponding event index ℓ as well as the sets T_ℓ and $T_{\ell-1}$. We consider the following cases (w.r.t I):

Case of empty I: Set $\ell = 0$ and $T_\ell = T_{\ell-1} = \emptyset$.

Case of non-empty I: Set $t_0 = 1$ and $T_0 = \emptyset$. We determine $\ell \geq 1$ (as well as T_ℓ), by iteratively considering $\ell = 1, ..., n$ (as in Section 3.2). That is, for $\ell = 1, ..., n$ do

i. We obtain t_ℓ by invoking a procedure analogous to $T(\ell, n)$ (of Section 3.2). Specifically, we approximate each of the p_s's by $\text{poly}(n) \cdot 2^\ell$ invocations of $\texttt{Scan}(\overline{h}', \overline{a}', j)$. Recall that each call of $\texttt{Scan}(\overline{h}', \overline{a}', j)$ specifies whether each session in Block j has properly decommitted, and approximations to the p_s's, denoted a_s's, are determined accordingly. We stress that p_s is the probability that $\texttt{Scan}(\overline{h}', \overline{a}', j)$ returns a progress pair (j, y') such that Session s properly decommits in y' (e.g., p_s is upper bounded by the probability that $\texttt{Scan}(\overline{h}', \overline{a}', j)$ returns a progress pair $(j, \cdot)$). Once all a_s's are determined, we determine t_ℓ just as in the second step of $T(\ell, n)$.

ii. Determine the set T_ℓ by determining, for each s, whether or not $p_s > t_\ell$. We use the above approximations to each p_s and rely on $|p_s - t_\ell| > (1/9n)2^{-\ell}$.

iii. Decide if event E_ℓ holds for y by using $T_{\ell-1}$ (of the previous iteration) and T_ℓ (just computed). Recall that event E_ℓ holds for y if $I \subseteq T_\ell$ but $I \not\subseteq T_{\ell-1}$.

iv. If event E_ℓ holds then exit the loop with the current value of ℓ as well as with the values of T_ℓ and $T_{\ell-1}$. Otherwise, proceed to the next iteration (i.e., the next value of ℓ).

In both cases (of I), we have determined the commitment type $X = E_\ell$ with respect to y (as obtained in the initial invocation of $\texttt{Scan}$) as well as the corresponding sets T_ℓ and $T_{\ell-1}$.

(This corresponds to Step S1 of the simulator of Section 3.)

(b) Exactly as in the original sub-step 2b. (That is, if either the initial answer is `failure` or (is a progress pair (k, y) with) $k \neq j$ then `return` with the very same answer (k, y).)

(c) Recall that we reach this sub-step only if the answer of the initial invocation of $\texttt{Scan}$ is a progress pair (j, y), and that we have already determined the event E_ℓ that holds (for y). By $\text{poly}(n) \cdot 2^\ell$ additional invocations of $\texttt{Scan}$ (with the same parameters as above), we may obtain progress pairs of the form $(j, \cdot)$ several times. In all cases the second component consists of a list of proper decommitment values. With overwhelmingly high probability, for each $s \in T_\ell$, we will obtain (from at least one of these lists) a proper decommitment for Session s (because $p_s > 2^\ell$). Ignoring the question of what decommitment types hold in these lists,[26] we combine all these lists to a list v of all proper decommitment values (obtained in any of these lists). This list v together with T_ℓ and $T_{\ell-1}$ (as obtained in sub-step 2a) forms a new information string $z = (v, T_\ell, T_{\ell-1})$,

[26] In particular, we do not care if the decommitment event happens to be of type E_ℓ or not. Furthermore, we may ignore y itself and not use it below (although we may also use y if we please).

which will be used below (i.e., recorded in $\overline{a}'$ for future use). (This corresponds to Step S2 of the simulator of Section 3.)

Next, analogously to the original sub-step 2c, we obtain an approximation to the probability that $\texttt{Scan}(\overline{h}', \overline{a}', j) = (j, y)$ such that E_ℓ holds in y. Specifically, we let $\tilde{q} \leftarrow \texttt{Approx}(\overline{h}', \overline{a}', (E_\ell, T_\ell, T_{\ell-1}), j)$, where procedure `Approx` uses T_ℓ and $T_{\ell-1}$ in order to determine whether the event E_ℓ holds in each of invocations of $\texttt{Scan}(\overline{h}', \overline{a}', j)$. We update the j^{th} record of $\overline{a}'$ by placing (E_ℓ, z) in the first field and $\tilde{q}$ in the second field. (This corresponds to Step S3 of the simulator of Section 3.)

(d) Finally, analogously to the original sub-step 2d, we invoke $\texttt{Generate}(\overline{h}', \overline{a}', j)$ up-to $\text{poly}(n)/\tilde{q}$ times and deal with the outcomes as in the original sub-step 2d. (This corresponds to Step S4 of the simulator of Section 3.)

3. *The handling of Step 3 messages by a block j* (possibly $j = i$) is analogous to the treatment in the original procedure, and we merely wish to spell out what this means: We consider two cases depending on whether or not $\overline{a}'$ contains the verifier's decommitment information for block j (i.e., the first field of the j^{th} block is not empty).

   (a) In case $\overline{a}'$ does contain such information, we consider sub-cases according to the relation of the contents of the the first field of the j^{th} block, denoted (E_ℓ, z), and the current answer of the verifier. Specifically, we check whether the verifier's current answer is of type E_ℓ. We note that the type of the current verifier decommitment is determined using the sets T_ℓ and $T_{\ell-1}$ provided in z (i.e., $z = (v, T_\ell, T_{\ell-1})$, where v is a sequence of decommitment values not used here). The sub-cases (fit versus non-fit) are handled as in the original procedure.

   (b) In case $\overline{a}'$ does not contain such information (i.e., the first field of the j^{th} block is empty), we obtain the relevant decommitment information (i.e, a sequence of decommitments) from the adversary, and `return` (as progress) with this information only.

This completes the description of the modification to the main procedures for the current setting (of bounded-simultaneity of blocks of parallel sessions). We stress that here (unlike in Section 3) the events E_ℓ regarding the decommitment to block j are not the only things that may happen when we invoke `Scan` with block index j (which corresponds to Step S1 in Section 3). As in Section 4, the answer may be `failure` or progress with respect to a different block. Indeed, the latter may not occur in case there is only one block, in which case the above treatment reduces to the treatment in Section 3. It is also instructive to note that when each block consists of a single copy, the above modified procedures degenerate to the original one (as in Section 4).

To analyze the current setting (of bounded-simultaneity of blocks of parallel sessions), we plug the analysis of Section 3 into the analysis of Section 4. The only point of concern is that we have introduced additional recursive calls (i.e., in the handling of Step 2, specifically in the handling sub-step 2a). However, as shown in Section 3, the expected number of these calls is bounded above by a polynomial (i.e., it is $\sum_{\ell=0}^{n} \Pr[E_\ell] \cdot 2^\ell \text{poly}(n)$, whereas $\Pr[E_\ell] = O(n \cdot 2^{-\ell})$). Thus, again, the tree of recursive calls has expected $\text{poly}(n)$ branching and depth at most w. Consequently, again, the expected running-time is bounded by $\text{poly}(n)^w$.

5.2.2 Combining the simulation techniques – the real case

In the real case the execution decomposes into blocks of almost parallel sessions (rather than perfectly parallel ones) such that (again) bounded-simultaneity holds with respect to the blocks.

In view of the extension in Section 3.5, the non-perfect parallelism within each block does not raise any problems (as far as a single block is concerned). What becomes problematic is the relation between the (non-perfectly parallel) blocks, and in particular our references to the ordering of steps taken by the different blocks. That is, our treatment of the perfect-parallelism case treats the parallel steps of each block as an atom. Consequently we have related to an ordering of these steps such that if one "block step" comes before another then all sessions in the the first block take the said step before any session of the other block takes the other step. However, in general, we cannot treat the parallel steps of each block as an atom, and the following problem arises: what if one session of block i takes Step A, next one session of block $j \neq i$ takes Step B, and then a different session of block i takes Step A. This problem seems particularly annoying if handling the relevant steps requires passing control between recursive calls. In general, the problem is resolved by treating differently the first (resp., last) session and other sessions of each block that reach a certain step. Loosely speaking, the first (or last) such session will be handled similarly to the atomic case, whereas in some cases other sessions (of the block) will be handled differently (in a much simpler manner). In particular, recursive calls are made only by the first session, and control is returned only by either the first or last such sessions. For sake of clarity, we present below the modification to the procedure `Generate`$(\overline{h}, \overline{a}, i)$. Note that this procedure is invoked when the immediate extension of $\overline{h}$ calls for execution of Step 2 by the *first* session in block i (i.e., $\overline{h}$ contains no Step 2 by any session that belongs to block i).

Initialization (upon invocation) step: Initializes $\overline{h}' = \overline{h}$ and $\overline{a}' = \overline{a}$, generates a passing commitment for (Step 2 of) the current (i.e., first) session of block i, and augments $\overline{h}'$ and $\overline{a}'$ accordingly. Specifically, the commitment is generated so that it passes the challenge corresponding to the current session (as recorded in the first field of record i), and only the corresponding part of the third field of the i^{th} record (in $\overline{a}'$) is updated.

In all the following cases, $\overline{h}'$ and $\overline{a}'$ denote the current history prefix and auxiliary information, respectively. (The following cases refer to the next message to be handled by the procedure, which handles such messages until it returns.)

Step 1 by some (new) session: Exactly as in the atomic case (i.e., augment $\overline{h}'$ and proceed to the next iteration).

Step 2 by the first session in block j (certainly $j \neq i$)**:** Analogous to the atomic case (see Section 5.2.1). Specifically, the handling depends on whether or not $\overline{a}'$ contains the verifier's decommitment information for copy j (i.e., whether or not the first field of the j^{th} record is non-empty).

1. In case $\overline{a}'$ does contain such information, we just generate a corresponding passing commitment (i.e., passing w.r.t the first field of the j^{th} record), augment $\overline{h}'$ and $\overline{a}'$ accordingly, and proceed to the next iteration.
2. In case $\overline{a}'$ does not contain such information (i.e., the first field of the j^{th} record is empty), we try to obtain such information. This is done analogously to the atomic case (see Section 5.2.1). We stress that this activity will yield the necessary information for all sessions in the j^{th} block, and not merely for the current (first) session in the block. Recall that the handling of this sub-case involves making recursive calls to the three procedures (with parameters $(\overline{h}', \overline{a}', j)$).

Step 2 by a non-first session in block j (here $j = i$ may hold)**:** We consider two cases depending on whether or not $\overline{a}'$ contains the verifier's decommitment information for copy j (i.e., whether or not the first field of the j^{th} record is non-empty).

1. In case $\overline{a}'$ does contain such information, we just generate a corresponding passing commitment, augment $\overline{h}'$ and $\overline{a}'$ accordingly, and proceed to the next iteration.

   (This is exactly as in the corresponding treatment of the first session of block j to reach Step 2.)

2. In case $\overline{a}'$ does not contain such information (i.e., the first field of the j^{th} record is empty), we generate a dummy commitment, augment $\overline{h}'$ accordingly, and proceed to the next iteration. (Recall that we count on the first session in the j^{th} block to find out the necessary information (for all sessions in the block).)

   (This is very different from the treatment of the first session of block j to reach Step 2.)

Step 3 by a non-last session of block j (possibly $j = i$)**:** Just augment $\overline{h}'$ accordingly (and proceed to the next iteration).

(This is very different from the treatment of the last session of block j to reach Step 3.)

Step 3 by the last session of block j (possibly $j = i$)**:** Analogous to the atomic case. We consider two cases depending on whether or not $\overline{a}'$ contains the verifier's decommitment information for block j (i.e., the first field of the j^{th} block is not empty).

1. In case $\overline{a}'$ does contain such information, we consider sub-cases according to the relation of the contents of the the first field of the j^{th} block, denoted (E_ℓ, z), and the Step 3 answer of the verifier (for all sessions in the j^{th} block). Specifically, we should consider the answers to previous sessions in the j^{th} block as recorded in $\overline{h}'$ and the answer to the last session in the block as just obtained. Recall that the type of the verifier decommitments (for the sessions in the j^{th} block) is determined using the sets T_ℓ and $T_{\ell-1}$ provided in the first field of the j^{th} block. The sub-cases (fit versus non-fit) are handled as in the original procedure. That is:

   (a) If the decommitment type of the Step 3 answers (of the j^{th} block) fits E_ℓ then we just augment $\overline{h}'$ accordingly (and proceed to the next iteration).

   (b) Otherwise (i.e., the decommitment type of the current Step 3 does not fit E_ℓ), `return failure`.

   (As in the atomic setting this case must hold if $j = i$.)

2. In case $\overline{a}'$ does not contain such information (i.e., the first field of the j^{th} block is empty), we obtain the relevant decommitment information as in the previous case, and `return` (as progress) with this information only. Specifically, the decommitment information for the previous sessions of the j^{th} block is recorded in $\overline{h}'$, whereas the the decommitment information for the last session has just been obtained (from the adversary).

Step 4 by a session of block j (possibly $j = i$)**:** Using the prover's decommitment information (as recorded in the third field of the j^{th} record), we emulate Step 4 in the straightforward manner (and augment $\overline{h}'$ accordingly). If this is the last session of block j and $j = i$, then `return` with the current $\overline{h}'$ and $\overline{a}'$ (otherwise proceed to the next iteration).

The modifications to procedure `Scan` are analogous. We stress that although the above description treats the schedule as if it is fixed, the treatment actually extends to a dynamic schedule where the membership of sessions in blocks is determined on-the-fly (i.e., upon their execution of Step 1).[27] The analysis of the perfect case can now be applied to the real case, and Theorem 1.1 follows.

Acknowledgments

We are grateful to Uri Feige and Alon Rosen for useful discussions.

[27] Recall that by our assumption that the verifier never violates the time-out condition (cf. Sec. 2.2), the "last session in a block to reach a certain step" can be determined as well.

References

[1] B. Barak. How to Go Beyond the Black-Box Simulation Barrier. In *42nd FOCS*, pages 106–115, 2001.

[2] B. Barak and Y. Lindell. Non-Black-Box Proofs of Knowledge (tentative title). In preparation, 2001.

[3] M. Bellare, R. Impagliazzo and M. Naor. Does Parallel Repetition Lower the Error in Computationally Sound Protocols? In *38th FOCS*, pages 374–383, 1997.

[4] M. Bellare, M. Jakobsson and M. Yung. Round-Optimal Zero-Knowledge Arguments based on any One-Way Function. In *EuroCrypt'97*, Springer-Verlag LNCS Vol. 1233, pages 280–305.

[5] G. Brassard, D. Chaum and C. Crépeau. Minimum Disclosure Proofs of Knowledge. *JCSS*, Vol. 37, No. 2, pages 156–189, 1988. Preliminary version by Brassard and Crépeau in *27th FOCS*, 1986.

[6] G. Brassard, C. Crépeau and M. Yung. Constant-Round Perfect Zero-Knowledge Computationally Convincing Protocols. *Theoretical Computer Science*, Vol. 84, pages 23–52, 1991.

[7] R. Canetti, O. Goldreich, S. Goldwasser, and S. Micali. Resettable Zero-Knowledge. In *32nd STOC*, pages 235–244, 2000.

[8] R. Canetti, J. Kilian, E. Petrank and A. Rosen. Black-Box Concurrent Zero-Knowledge Requires $\tilde{\Omega}(\log n)$ Rounds. In *33rd STOC*, pages 570–579, 2001.

[9] I. Damgård. Efficient Concurrent Zero-Knowledge in the Auxiliary String Model. In *Eurocrypt'00*, 2000.

[10] D. Dolev, C. Dwork, and M. Naor. Non-Malleable Cryptography. In *23rd STOC*, pages 542–552, 1991. Full version available from authors.

[11] C. Dwork, M. Naor, and A. Sahai. Concurrent Zero-Knowledge. In *30th STOC*, pages 409–418, 1998.

[12] C. Dwork, and A. Sahai. Concurrent Zero-Knowledge: Reducing the Need for Timing Constraints. In *Crypto98*, Springer LNCS 1462.

[13] U. Feige and A. Shamir. Zero-Knowledge Proofs of Knowledge in Two Rounds. In *Crypto'89*, Springer-Verlag LNCS Vol. 435, pages 526–544, 1990.

[14] O. Goldreich. *Foundation of Cryptography – Basic Tools*. Cambridge University Press, 2001.

[15] O. Goldreich and A. Kahan. How to Construct Constant-Round Zero-Knowledge Proof Systems for NP. *J. of Crypto.*, Vol. 9, No. 2, pages 167–189, 1996. Preliminary versions date to 1988.

[16] O. Goldreich and H. Krawczyk. On the Composition of Zero-Knowledge Proof Systems. *SICOMP*, Vol. 25, No. 1, February 1996, pages 169–192. Preliminary version in *17th ICALP*, 1990.

[17] O. Goldreich, S. Micali and A. Wigderson. Proofs that Yield Nothing but their Validity or All Languages in NP Have Zero-Knowledge Proof Systems. *JACM*, Vol. 38, No. 1, pages 691–729, 1991. Preliminary version in *27th FOCS*, 1986.

[18] O. Goldreich and Y. Oren. Definitions and Properties of Zero-Knowledge Proof Systems. *J. of Crypto.*, Vol. 7, No. 1, pages 1–32, 1994.

[19] S. Goldwasser and S. Micali. Probabilistic Encryption. *JCSS*, Vol. 28, No. 2, pages 270–299, 1984. Preliminary version in *14th STOC*, 1982.

[20] S. Goldwasser, S. Micali and C. Rackoff. Knowledge Complexity of Interactive Proofs. In *17th STOC*, pages 291–304, 1985. This is a preliminary version of [21].

[21] S. Goldwasser, S. Micali and C. Rackoff. The Knowledge Complexity of Interactive Proof Systems. *SICOMP*, Vol. 18, pages 186–208, 1989. Preliminary version in [20].

[22] J. Håstad, R. Impagliazzo, L.A. Levin and M. Luby. A Pseudorandom Generator from any One-way Function. *SICOMP*, Vol. 28, No. 4, pages 1364–1396, 1999. Preliminary versions by Impagliazzo *et. al.* in *21st STOC* (1989) and Håstad in *22nd STOC* (1990).

[23] J. Kilian and E. Petrank. Concurrent and resettable zero-knowledge in poly-logarithmic rounds. In *33rd STOC*, pages 560–569, 2001.

[24] J. Kilian, E. Petrank, and C. Rackoff. Lower Bounds for Zero-Knowledge on the Internet. In *39th FOCS*, pages 484–492, 1998.

[25] M. Naor. Bit Commitment using Pseudorandom Generators. *J. of Crypto.*, Vol. 4, pages 151–158, 1991.

[26] R. Richardson and J. Kilian. On the Concurrent Composition of Zero-Knowledge Proofs. In *EuroCrypt99*, Springer LNCS 1592, pages 415–413.

[27] A.C. Yao. Theory and Application of Trapdoor Functions. In *23rd FOCS*, pages 80–91, 1982.

Appendix: Application to the BJY-protocol

We start by briefly recalling the BJY-protocol (due to Bellare, Jakobsson and Yung [4], which in turn builds upon the work of Feige and Shamir [13]). Their protocol uses an adequate three-round witness indistinguishable proof system (e.g., parallel repetition of the basic zero-knowledge proof of [17].). Specifically, we consider a three-round witness indistinguishable proof systems (e.g., for $G3C$) of the form:

Step WI1: The prover commits to a sequence of values (e.g., the colors of each vertex under several 3-colorings of the graph).

Step WI2: The verifier send a random challenge (e.g., a random sequence of edges).

Step WI3: The prover decommits to the corresponding values.

(The implementation details are as in Construction 2.2.)[28] For technical reasons, it is actually preferable to use protocols for which demonstrating a "proof of knowledge" property is easier (e.g., parallel execution of Blum's basic protocol; cf. [14, Sec. 4.7.6.3] and [14, Chap. 4, Exer. 28]). The commitment scheme used above is perfectly-binding (and non-interactive; see Footnote 14). Given the above, the (four-round) BJY-protocol (for any language $L \in \mathcal{NP}$) proceeds as follows:

1. The verifier sends many hard "puzzles", which are unrelated to the common input x. These puzzles are random images of a one-way function f, and their solutions are corresponding preimages. In fact, the verifier selects these puzzles by uniformly selecting preimages of f, and applying f to obtain the corresponding images. Thus, the verifier knows solutions to all puzzles he has sent.

   In the rest of the protocol, the prover will prove (in a witness indistinguishable manner) that either it knows a solution to one of (a random subset of) these puzzles or $x \in L$. The latter proof is by reduction to some instance of an NP-complete language.
2. The prover performs Step WI1 in parallel to asking to see a random subset of the solutions to the above puzzles. Specifically, the puzzles are paired, and the prover asks to see a solution to one (randomly *selected*) puzzle in each pair. Furthermore, in executing Step WI1, the prover refers to a statement derived from the reduction of the assertion $x \in L$ or some of the *non-selected* puzzles has a solution.
3. The verifier performs Step WI2 in parallel to sending the required solutions (to the selected puzzles).
4. The prover verifies the correctness of the solutions provided by the verifier, and in case all solutions are correct it performs Step WI3.

As shown in [4], the BJY-protocol is a four-round zero-knowledge argument system for L. The simulator is similar to the one presented for the GK-protocol. Specifically, it starts by executing Steps 1–3, while using dummy commitments (in Step 2). Such a partial execution is called *proper* if the adversary has revealed all solutions to the selected puzzles (and is called *improper* otherwise). In case the partial execution is improper, the simulator halts while outputting it. Otherwise, the simulator moves to generating a full execution transcript by repeatedly rewinding to Step 2 and trying to emulate Steps 2–4 using the fact that (unless it selects the same set of puzzles again (which is highly unlikely)) it already knows a solution to one of the puzzles not selected (by it) in

[28] Specifically, the prover commits to the colors of each vertex under t random relabelings of a 3-colorings of the graph, the challenge is a sequence of t edges, and the prover decommits to the values corresponding to the end-points of the i^{th} edge with respect to the i^{th} committed coloring.

the current execution (but rather selected in the initial execution of Steps 1–3). As in the simulation of the GK-protocol (cf. [15]), the number of repetitions must be bounded by the reciprocal of the probability of a proper (initial) execution (as approximated by an auxiliary intermediate step).[29]

Given the similarity of the two simulators (i.e., the one here and the one for the GK-protocol), it is evident that our treatment of concurrent composition of the GK-protocol applies also to the BJY-protocol.

[29] Unfortunately, this technical issue is avoided by Bellare *et. al.* [4], but it arises here (i.e., in [4]) similarly to the way it arises in [15], and it can be resolved in exactly the same manner. (The issue is that the prover commitments in the initial scan are distributed differently (but computational-indistinguishablly) than its commitments in the generation process.)