Large-Scale Secure Computation: Multi-party Computation for (Parallel) RAM Programs

Elette Boyle[1(✉)], Kai-Min Chung[2], and Rafael Pass[3]

[1] Technion Israel, Haifa, Israel
eboyle@alum.mit.edu
[2] Academica Sinica, Taipei, Taiwan
kmchung@iis.sinica.edu.tw
[3] Cornell University, Ithaca, USA
rafael@cs.cornell.edu

Abstract. We present the first efficient (i.e., polylogarithmic overhead) method for securely and privately processing large data sets over multiple parties with *parallel, distributed algorithms*. More specifically, we demonstrate load-balanced, statistically secure computation protocols for computing Parallel RAM (PRAM) programs, handling $(1/3-\epsilon)$ fraction malicious players, while preserving up to polylogarithmic factors the computation, parallel time, and memory complexities of the PRAM program, aside from a one-time execution of a broadcast protocol per party. Additionally, our protocol has $\mathsf{polylog}$ communication locality—that is, each of the n parties speaks only with $\mathsf{polylog}(n)$ other parties.

1 Introduction

Large data sets, such as medical data, genetic data, transaction data, the web and web access logs, and network traffic data, are now in abundance. Much of the data is stored or made accessible in a distributed fashion, having necessitated the development of efficient distributed protocols that compute over such data. In particular, novel programming models for processing large data sets with *parallel, distributed algorithms*, such as MapReduce (and its implementation Hadoop) are emerging as crucial tools for leveraging this data in important ways.

But these methods require that the data itself is revealed to the participating servers performing the computation—and thus blatantly violate the privacy of

E. Boyle—The research of the first author has received funding from the European Union's Tenth Framework Programme (FP10/ 2010-2016) under grant agreement no. 259426 ERC-CaC, and ISF grant 1709/14.

R. Pass—Pass is supported in part by a Alfred P. Sloan Fellowship, Microsoft New Faculty Fellowship, NSF Award CNS-1217821, NSF CAREER Award CCF-0746990, NSF Award CCF-1214844, AFOSR YIP Award FA9550-10-1-0093, and DARPA and AFRL under contract FA8750-11-2- 0211. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the US Government.

R. Gennaro and M. Robshaw (Eds.): CRYPTO 2015, Part II, LNCS 9216, pp. 742–762, 2015.
DOI: 10.1007/978-3-662-48000-7_36

potentially sensitive data. As a consequence, such methods cannot be used in many critical applications (e.g., discovery of causes or treatments of diseases using genetic or medical data).

In contrast, methods such as secure multi-party computation (MPC), introduced in the seminal works of Yao [Yao86] and Goldreich, Micali and Wigderson [GMW87], enable securely and privately performing any computation on individuals private inputs (assuming some fraction of the parties are honest). However, despite great progress in developing these techniques, there are no MPC protocols whose efficiency and communication requirements scale to the modern regime of large-scale distributed, parallel data processing.

We are concerned with merging these two approaches. In particular,

We seek MPC protocols that efficiently *(technically, with polylogarithmic overhead) enable secure and private processing of large data sets with parallel, distributed algorithms.*

Explicitly, in this large-scale regime, the following properties are paramount:

1. *Exploiting Random Access.* Computations on large data sets are frequently "lightweight": accessing a small number of dynamically chosen data items, relying on conditional branching, and/or maintaining small memory. This means that converting a program first into a circuit to enable its secure computation, which immediately obliterates these gains, will not be a feasible option.
2. *Exploiting Parallelism.* In fact, as mentioned, to effectively solve large-scale problems, modern programming models heavily leverage parallelism. The notion of a Parallel RAM (PRAM) better captures such computing models. In the PRAM model of computation, several (polynomially many) CPUs run simultaneously, potentially communicating with one another, while accessing the same shared external memory. We consider a PRAM model with a variable number of CPUs but with a fixed activation structure (i.e., what processors are activated at which time steps is fixed). Note that such a model simultaneously captures RAMs (a single CPU) and circuits (the circuit topology dictates the CPU activation structure).
3. *Exploiting Plurality of Users.* In the setting of MPC we would like to leverage not only parallelism within a single party (i.e., if a party has multiple CPUs that may run in parallel), but also that we have a large number of parties that can run in parallel. So, if we have n parties, each with k processors, we ideally would like to securely compute PRAMs that use nk CPUs (as opposed to just k CPUs).

Additionally, the following desiderata are often of importance:

4. *Load balancing.* When the data set contains tens or hundreds of thousands of users' data, it is often unreasonable to assume that any single user can provide memory, computation, or communication resources on the order of the data of *all users*. Rather, we would like to *balance* the load across nodes.

5. *Communication Locality.* In many cases, establishing a secure communication channel with a large number of distinct parties may be costly, and thus we would like to minimize the *locality of communication* [BGT13]: that is, the number of total parties that each party must send and receive message to during the course of the protocol.

To date, no existing work addresses secure computation of Parallel RAM programs. Indeed, nearly all results in MPC require a *circuit* model for the function being evaluated (including the line of work on scalable MPC [DI06,DIK+08, DKMS12,ZMS14]), and thus inherit resource requirements that are linear in the circuit size. Even for (sequential) RAM, the only known protocols either only handle two parties [OS97,GKK+11,LO13,GGH+13], or in the context of multi-party computation require all parties to store *all inputs* [DMN11], rendering the protocol useless in a large-scale setting (even forgetting about computation load balancing and locality).

1.1 Our Results

We present a statistically secure MPC for (any sequence of) PRAMs handling $(1/3 - \epsilon)$ fraction static corruptions in a synchronous communication network, with secure point-to-point channels. In addition, our protocol is strongly *load balanced* and *communication local* (i.e., $\mathsf{polylog}(n)$ locality). We state our theorem assuming each party itself is a k-processor PRAM, for parameter k.

Theorem 1 (Informal – Main Theorem). *For any constant $\epsilon > 0$ and polynomial parallelism parameter $k = k(n)$, there exists an n-party statistically secure (with error negligible in n) protocol for computing any adaptively chosen sequence of PRAM programs Π_j with fixed CPU activation structures (and that may have bounded shared state), handling $(1/3 - \epsilon)$ fraction static corruptions with the following complexities, where each party is a k-processor PRAM (and where $|x|, |y|$ denote per-party input and output size,*[1] $\mathsf{space}(\Pi)$*, $\mathsf{comp}(\Pi)$, and $\mathsf{time}(\Pi)$ denote the worst-case space, computation, and (parallel) runtime of Π, and $CPUs(\Pi)$ denotes the number of CPUs of Π):*

- *Computation per party, per Π_j:* $\tilde{O}\big(\mathsf{comp}(\Pi_j)/n + |y|\big)$.
- *Time steps, per Π_j:* $\tilde{O}\left(\mathsf{time}(\Pi_j) \cdot \max\left\{1, \frac{CPUs(\Pi)}{nk}\right\}\right)$.
- *Memory per party:* $\tilde{O}\left(|x| + |y| + \max_{j=1}^{N} \mathsf{space}(\Pi_j)/n\right)$.
- *Communication Locality:* $\tilde{O}(1)$.

given a one-time preprocessing phase with complexity:

- *Computation per party:* $\tilde{O}(|x|)$*, plus single broadcast of* $\tilde{O}(1)$ *bits.*
- *Time steps:* $\tilde{O}\left(\max\left\{1, \frac{|x|}{k}\right\}\right)$.

[1] For simplicity of exposition, we assume all parties have the same input size and receive the same output.

Additionally, our protocol achieves a strong "online" load-balancing guarantee: at all times *during the protocol, all parties' communication and computation loads vary by at most a constant multiplicative factor (up to a* $\mathsf{polylog}(n)$ *additive term).*

Remark 1 (Round complexity). As is the case with all general MPC protocols in the information-theoretic setting to date, the round complexity of our protocol corresponds directly with the time complexity (as when restricted to circuits, parallel complexity corresponds to circuit depth). That is, for each evaluated PRAM program Π_j, the protocol runs in $\tilde{O}(\mathsf{time}(\Pi_j))$ sequential communication rounds to securely evaluate Π_j.

Remark 2 (On the achieved parameters). Note that in terms of memory, each party only stores her input, output, and her "fair" share of the required space complexity, up to polylogarithmic factors. In terms of computation (up to polylogarithmic factors), each party does her "fair" share of the computation, receives her outputs, and in addition is required to read her entire input at an initial preprocessing stage (even though the computations may only involve a subset of the input bits; this additional overhead of "touching" the whole input once is necessary to achieve security).[2] Finally, the time complexity corresponds to the parallel complexity of the PRAM being computed, as long as the combined number of available processors nk from all parties matches or exceeds the number of required parallel processes of the program (and degrades with the corresponding deficit).

Remark 3 (Instantiating the single-use broadcast). The broadcast channel can be instantiated either by the $O(\sqrt{n})$-locality broadcast protocol of King *et al.* [KSSV06], or the $\mathsf{polylog}(n)$-*average* locality protocol of [BSGH13] at the expense of a cost of a one-time per-party computational cost of $O(\sqrt{n})$, or average cost of $\mathsf{polylog}(n)$, respectively. We separate the broadcast cost from our protocol complexity measures to emphasize that any (existing or future) broadcast protocol can be directly plugged in, yielding associated desirable properties.[3]

1.2 Construction Overview

Our starting point is an *Oblivious PRAM (OPRAM)* compiler [BCP14b, GO96], a tool that compiles any PRAM program into one whose memory access patterns

[2] For general secure computation, and even if we restrict to functionalities that only access a few parties' inputs, and only a few bits of their data, essentially all parties must perform computation at least $\Omega(|x|)$. To see this, consider secure computation of a "multi-party Private Information Retrieval (PIR)" functionality: each party $i > 1$ has as input some "big data" x_i, and party 1 has as input a party index i and an index j into their data x_i. The functionality returns $x_i[j]$ (i.e., the j'th bit of party i's data) to party 1 and nothing to everyone else. We claim that each party $i > 1$ must access every bit of x_i; if not, it learns that particular bit of its data was not requested, which it cannot learn in an ideal execution of the functionality.

[3] For instance, it remains open to achieve statistically secure broadcast with worst-case $\mathsf{polylog}(n)$ locality.

are independent of the data (i.e., "oblivious"). Such a compiler (with polylogarithmic overhead) was recently attained by [BCP14b].

Indeed, it is no surprise that such a tool will be useful toward our goal. It has been demonstrated in the sequential setting that Oblivious (sequential) RAM (ORAM) compilers can be used to builds secure 2-party protocols for RAM programs [OS97, GKK+11, LO13, GGH+13]. Taking a similar approach, building upon the OPRAM compiler of [BCP14b] directly yields 2-party protocols for PRAMs.

However, OPRAM on its own does not directly provide a solution for *multi*-party computation (when there are many parties). While this approach gives protocols whose complexities scale well with the RAM (or PRAM) complexity of the programs, the complexities grow poorly with the number of parties. Indeed, the only current technique for securely evaluating a RAM program on *multiple* parties' inputs [DMN11] is for all parties to hold secret shares of all parties' inputs, and then jointly execute (using standard MPC for circuits) the trusted CPU instructions of the ORAM-compiled version of the program. This means each party must communicate and maintain information of size equivalent to *all parties'* inputs, and everyone must talk to everyone else for every time step of the RAM program evaluation.

One may attempt to improve the situation by first electing a small $\mathsf{polylog}(n)$-size representative committee of parties, and then only performing the above steps within this committee. This approach drops the total communication and computation of the protocol to reasonable levels. However, this approach does not save the subset of elected parties from carrying the burden of the entire computation. In particular, each elected party must memory storage equal to the size of *all parties' inputs combined*, making the protocol unusable for "large-scale" computation.

In this paper, we provide a new approach for dealing with this issue. We show how to use an OPRAM in a way that achieves balancing of memory, computation, and communication across all parties.

Our MPC construction proceeds in the following steps:

1. **From OPRAM to MPC.** Given an OPRAM, we begin by considering MPC in a "benign" adversarial setting, which we refer to as *oblivious multi-party computation*, where all parties are assumed to be honest, and we only require that an external attacker that views communication and activation (including memory and computation usages) patterns does not learn anything about the inputs. We show:
   (a) OPRAM yields efficient *memory-balanced* oblivious MPC for PRAM.
   (b) Using committee election techniques (à la [KLST11, DKMS12, BGT13]), any oblivious multi-party computation can be compiled into a standard secure MPC with only $\mathsf{polylog}$ overhead (and a one-time use of a broadcast channel per party).
2. **Load Balancing & Communication Locality.** We next show semi-generic compilers for "nice" (formally defined) *oblivious multi-party* protocols, each introducing only $\mathsf{polylog}(n)$ overhead:

(a) From any "nice" protocol to one whose computation and communication are *load-balanced*.
(b) From any "nice" protocol to one that is both *load-balanced* and *communication local* (i.e., $\mathsf{polylog}(n)$ locality).

Our final result is obtained by combining the above steps and observing that Step 1(b) preserves load-balancing and communication locality (and thus can be applied *after* Step 2). Let us mention that *just* Step 1 (together with existing construction of ORAMs) already yields the first MPC protocol for (sequential) RAM programs in which no party must store all parties' inputs. Additionally, *just* Step 1 (together with the OPRAM construction of [BCP14b]) yields the first MPC for PRAMs.

We now expand upon each of these steps.

MPC from OPRAM. Recall that our construction proceeds via an intermediate notion of *oblivious* security, in which we do not require security against corrupted parties, but rather against an external adversary who sees the activation patterns (i.e., accessed memory addresses and computation times) and communication patterns (i.e., sender/receiver ids and message lengths) of parties throughout the protocol.

Oblivious MPC from OPRAM. At a high level, our protocol will emulate a *distributed* OPRAM[4] structure, where the CPUs and memory cells in the OPRAM are *each associated with parties*. (Recall that we need only achieve "oblivious" security, and thus can trust individual parties with these tasks). The "CPU" parties will control the evaluation flow of the (OPRAM-compiled) program, communicating with the parties emulating the role of the appropriate memory cells for each address to be accessed in the (OPRAM-compiled) database.

The distributed OPRAM structure will enable us to evenly spread the memory burden across parties, incurring only $\mathsf{polylog}(n)$ overhead in total memory and computation, and while guaranteeing that the communication patterns between committees (corresponding to data access patterns) do not reveal information on the underlying secret values.

This framework shares a similar flavor to the protocols of [DKMS12, BGJK12], which assign committees to each of the gates of a circuit being evaluated, and to [BGT13], which uses CPU and input committees to direct program execution and distributedly store parties' inputs. The distributed OPRAM idea improves and conceptually simplifies the input storage handling of Boyle *et al.* [BGT13], in which n committees holding the n parties' inputs execute a distributed "oblivious input shuffling" procedure to break the link between which committees are communicating and which inputs are being accessed in the computation.

[4] We remark that the term "distributed ORAM" was used with a different meaning in [LO13], in regard to an ORAM that was split across two users.

Compiling from "Oblivious" Security to Malicious Security. We next present a general compiler taking an oblivious protocol to one that is secure against $(1/3 - \epsilon)n$ statically corrupted malicious parties. (This step can be viewed as a refinement and generalization of ideas from [KLST11, DKMS12, BGT13].) We ensure the compiler tightly preserves the computation, memory, load-balancing, and communication locality of the original protocol, up to $\mathsf{polylog}(n)$ factors (modulo a one-time broadcast per party). This enables us to apply the transformation to any of the oblivious protocols resulting from the intermediate steps in our progression.

At a high level, the compiler takes the following form: (1) First, the parties collectively elect a large number of "good" committees, each of size $\mathsf{polylog}(n)$, where "good" means each committee is composed of at least 2/3 honest parties, and that parties are spread roughly evenly across committees. (2) Each party will verifiably secret share his input among the corresponding committee C_i. (3) From this point on, the role of each party P_i in the original protocol will be emulated by the corresponding committee C_i. That is, each local P_i computation will be executed via a small-scale MPC among C_i, and each communication from P_i to P_j will be performed via an MPC among committees C_i and C_j.

The primary challenge in this step is how to elect such committees while incurring only $\mathsf{polylog}(n)$ locality and computation per party. To do so, we build atop the "almost-everywhere" scalable committee election protocol of King *et al.* [KSSV06] to elect a single good committee, and then show that one may use a $\mathsf{polylog}(n)$-wise independent function family $\{F_s\}_{s \in S}$ to elect the remaining committees with small description size (in the fashion of [KLST11, BGT13], for the case of combinatorial samplers and computational pseudorandom functions), with committee i defined as $C_i := F_s(i)$ for fixed random seed s.

We remark that, aside from the one-time broadcast, this compiler preserves load balancing and $\mathsf{polylog}(n)$ locality. Indeed, load balancing is maintained since the committee setup procedure is computationally inexpensive, and each party appears in roughly the same number of "worker" committees. The locality of the resulting protocol increases by an additive $\mathsf{polylog}(n)$ for the committee setup, and a multiplicative $\mathsf{polylog}(n)$ term since all communications are now performed among $\mathsf{polylog}(n)$-size *committees* instead of individual parties.

Load Balancing Distributed Protocols

Load-Balancing (Without Locality). We now show how to modify our protocol such that the total computational complexity and memory balancing are preserved, while additionally achieving a strong *computation* load balancing property—with high probability, at all times throughout the protocol execution, every party performs close to $1/n$ fraction of current total work, up to an additive $\mathsf{polylog}(n)$ amount of work. This will hold simultaneously for both computation and communication.[5]

[5] Note that while our current protocol is memory balanced, it is currently rather imbalanced in computation: e.g., the parties emulating OPRAM CPUs are required to perform computation that is proportional to the whole PRAM computation.

We present and analyze our load-balancing solution in the intermediate *oblivious* MPC security setting (recall that one can then apply the compiler from Step 2(b) above to obtain malicious MPC with analogous load-balancing). Let us mention that there is a huge literature on "load-balanced distributed computation" (e.g., [ACMR95,MPS02,MR98,AAK08]): As far as we can tell, our setting differs from the typical studied scenarios in that we must load balance an underlying distributed protocol, as opposed to a collection of independent "non-communicating jobs". Indeed, the main challenge in our setting is to deal with the fact that "jobs" talk to one another, and this communication must remain efficient also be made load balanced. Furthermore, we seek a load-balanced solution with communication locality.

We consider a large class of arbitrary (potentially load-unbalanced and large-locality) distributed protocols Π, where we view each party in this underlying protocol as a "job". Our goal is to load-balance Π by passing "jobs" between "workers" (which will be the actual parties in the new protocols). More precisely, we start off with any protocol Π that satisfies the following (natural) "nice" properties:

- Each "job" has $\mathsf{polylog}(n)$ size state;
- In each round, each "job" performs at most $\mathsf{polylog}(n)$ computation and communication;
- In each round, each "job" communicates (either sending or receiving a message) to at most one other "job".

It can be verified that these properties hold for our oblivious MPC for PRAM protocol.

Our load-balanced version of such a protocol first randomly[6] efficiently assigns "workers" (i.e., parties) to "jobs". Next, whenever a worker W has performed "enough" work for a particular job J, it randomly selects a replacement worker W' and passes the job over to it (that is, it passes over the state of the job J—which is "small" by assumption). The key obstacle in our setting is that the job J may later communicate with many other jobs, and all the workers responsible for those jobs need to be informed of the switch (and in particular, who the new worker responsible for the job J is). Since the number of jobs is $\Omega(n)$, workers cannot afford to store a complete directory of which worker is currently responsible for each job.

We overcome this obstacle by first modifying Π to ensure that it has small locality—this enables each job to only maintain a short list of the workers currently responsible for the *"neighboring"* jobs. We achieve this locality by requiring that parties (i.e., jobs) in the original protocol Π route their messages along the hypercube. Now, whenever a worker W for a job J is being replaced by some worker W', W informs all J's neighboring jobs (i.e., the workers responsible for them) of this change. We use the Valiant-Brebner [VB81] routing procedure to implement the hypercube routing because it ensures a desirable "low-congestion

[6] In the actual analysis, we show that it also suffices to use $\mathsf{polylog}(n)$-wise independent randomness to pick this and subsequent assignments.

property," which in our setting translates to ensuring that the overhead of routing is not too high for any individual worker.

The above description has not yet mentioned what it means for a worker to have done "enough" work for a job J. Each round a job is active (i.e., performing some computation), its "cost" increases by 1—we refer to this as an *emulation cost*. Additionally, each time a worker W is switched out from a job J, then J's and each of J's neighboring jobs' costs are increased by 1—we refer to this as a *switch cost*. Finally, once a job's (total) cost has reached a particular threshold τ, its cost is reset to 1 and the worker responsible for the job is switched out. The threshold τ is set to $2 \log M + 1$ where M is the number of jobs.

We show: (1) This switching does not introduce too much overhead. We, in fact, show that the total induced switching cost is bounded above by the emulation cost. (2) The resulting total work is load balanced across workers—we show this by first demonstrating that the protocol is load-balanced in expectation, and then using concentration to argue our stronger online load-balancing property.

Finally, note that although communication between *jobs* is being routed through the hypercube, and thus the job communication protocol has small locality, the final load-balanced protocol, being run by *workers*, does *not* have small locality. This is because workers are assigned the role of many different jobs over time, and may possibly speak to a new set of neighbors for each position. (Indeed, over time, each worker will eventually need to speak to every other worker). We next show how to modify this protocol to achieve *locality*, while preserving load-balancing.

Achieving Both Load-Balancing and Locality. In our final step, we show how to modify the above-mentioned protocol to also achieve locality. We modify the protocol to also let *workers* route messages through a low-degree network (on top of the routing in the previous step). This immediately ensures locality. But, we must be careful to ensure that the additional message passing does not break load-balancing.

A natural idea is to again simply pass messages between *workers* along a low-degree hypercube network via Valiant-Brebner (VB) routing [VB81]. Indeed, the low-congestion property will ensure (as before) that routing does not incur too large an overhead for each worker.

However, when analyzing the overall load balance (for workers), we see an inherent distinction between this case and the previous. Previously, the nodes of the hypercube corresponded to *jobs*, each emulated by workers who swap in and out over time. When the underlying jobs protocol required job s to send a message to job t, the resulting message routing induced a cost along a path of neighboring jobs (that is, the workers emulating them), *independent of which workers are currently emulating them*. This independence, together with the fact that a worker passes his job after performing "enough" work for it, enabled us to obtain concentration bounds on overall load balancing over the random assignment of workers to jobs.

Now, the nodes correspond directly to *workers*. When the underlying jobs protocol requires a message transferred from job s to job t, routing along the

workers' graph must traverse a path from the *worker currently emulating job* s to the *worker currently emulating job* t, removing the crucial independence property from above. Even worse, workers along the routing path can now incur costs *even if they are not assigned to any job*. In this case, it is not even clear that job passing in of itself will be sufficient to ensure balancing.

To get around these issues, we add an extra step in the VB routing procedure (itself inspired by [VB81]) to break potential bad correlations. The idea is as follows: To route from the worker W_s emulating job s to the worker W_t emulating job t, we first route (as usual) from W_s to a *random* worker W_u, and then from W_u to W_t; i.e., travel from W_s to W_t by "walking into the woods" and back. We may now partition the cost of routing into these two sub-parts, each associated with a single active job (s or t). Now, although workers along the worker-routing path will still incur costs from this routing (even though their jobs may be completely unrelated), the *distribution* of these costs on workers depends only on the identity of the initiating worker (W_s or W_t). We may thus generalize the previous analysis to argue that if the expectation of work is load-balanced, then it still has concentration in this case.

For a modular analysis, we formalize the required properties of the underlying communication network and routing algorithm (to be used for the s-to-u and u-to-t routing) as a *local load-balanced routing network*, and show that the hypercube network together with VB routing satisfies these conditions.

1.3 Discussion and Future Work

With the explosive growth of data made available in a distributed fashion, and the growth of efficient parallel, distributed algorithms (such as those enabled by MapReduce) to compute on this data, ensuring privacy and security in such large-scale parallel settings is of fundamental importance. We have taken the first steps in addressing this problem by presenting the first protocols for secure multi-party computation, that with only polylogarithmic overhead, enable evaluating PRAM programs on a (large) number of parties' inputs. Our work leaves open several interesting open problems:

Honest Majority. We have assumed that 2/3 of the players are honest. In the absence of a broadcast channel,[7] it is known that this is optimal. But if we assume the existence of a broadcast channel, it may suffice to assume 1/2 fraction honest players.

Asynchrony. Our protocol assumes a synchronous communication network. We leave open the handling of asynchronous communication.

Trading efficiency for security. An interesting avenue to pursue are various tradeoffs between boosted efficiency and partial sacrifices in security. For example, in some settings, it is not detrimental to leak which parties' inputs were used within the computation; in such scenarios, one could then hope

[7] While the statement of our result makes use of a broadcast channel, as we mention, this channel can also be instantiated with known protocols.

to remove the one-time $\Theta(n|x|)$ input preprocessing cost. Similarly, it may be acceptable to reveal the input-specific resources (runtime, space) required by the program on parties inputs; in such cases, we may modify the protocol to take only input-specific runtime and use input-specific memory.

In this work we focus only on achieving standard "full" security. However, we remark that our protocol can serve as a solid basis for achieving such tradeoffs (e.g., a straightforward tweak to our protocol results in input-specific resource use).

Communication complexity. As with all existing generic multi-party computation protocols in the information-theoretic setting, the communication complexity of our protocol is equal to its computation complexity. In contrast, in the computational setting (based on cryptographic assumptions), protocols with communication complexity below the complexity of the evaluated function have been constructed by relying on *fully homomorphic encryption (FHE)* [Gen09] (e.g., [Gen09, AJLA+12, MSS13]). We leave as an interesting open question whether FHE-style techniques can be applied also to our protocol to improve the communication complexity, based on computational assumptions.

1.4 Overview of the Paper

Section 2 contains preliminaries. In Sect. 3 we provide our ultimate theorem, and the sequence of intermediate notions and theorems which combine to yield this final result. We refer the reader to the full version of this work [BCP14a] for a complete descriptions and proofs.

2 Preliminaries

2.1 Multi-party Computation (MPC)

Protocol Syntax. We model parties as (parallel) RAM machines. An n-party protocol Φ is described as a collection of n (parallel) RAM programs $(P_i)_{i\in[n]}$, to be executed by the respective parties, containing additional special communication instructions $\mathsf{Comm}(i, \mathsf{msg})$, indicating for the executing party to send message msg to party i.

The per-party space, computation, and time complexities of the protocol $\Phi = (P_i)_{i\in[n]}$ are defined directly with respect to the corresponding party's PRAM program P_i, where each Comm is charged as a single computation time step. (See Sect. 2.2 for a definition of $CPUs(P)$, $\mathsf{space}(P)$, $\mathsf{comp}(P)$, $\mathsf{time}(P)$ for PRAM P). The analogous total protocol complexities are defined as expected: Namely, $\mathsf{space}(\Phi)$ and $\mathsf{comp}(\Phi)$ are the *sums*, $\mathsf{space}(\Phi) = \sum_{i\in[n]} \mathsf{space}(P_i)$, $\mathsf{comp}(\Phi) = \sum_{i\in[n]} \mathsf{comp}(P_i)$, and $\mathsf{time}(\Phi)$ is the *maximum*, $\mathsf{time}(\Phi) = \max_{i\in[n]} \mathsf{time}(P_i)$.

MPC Security. We consider the standard notion of (statistical) MPC security. We refer the reader to e.g. [BGW88] for more a more complete description of MPC security within this setting.

2.2 Parallel RAM (PRAM) Programs

A Concurrent Read Concurrent Write (CRCW) m-processor *parallel random-access machine (PRAM)* with memory size n consists of numbered processors $CPU_1, \ldots, CPU_m$, each with local memory registers of size $\log n$, which operate synchronously in parallel and can make access to shared "external" memory of size n.

A PRAM program Π (given m, n, and some input x stored in shared memory) provides CPU-specific execution instructions, which can access the shared data via commands $\mathsf{Access}(r, v)$, where $r \in [n]$ is an index to a memory location, and v is a word (of size $\log n$) or $\perp$. Each $\mathsf{Access}(r, v)$ instruction is executed as:

1. **Read** from shared memory cell address r; denote value by v_{old}.
2. **Write** value $v \neq \perp$ to address r (if $v = \perp$, then take no action).
3. **Return** v_{old}.

In the case that two or more processors simultaneously initiate $\mathsf{Access}(r, v_i)$ with the same address r, then all requesting processors receive the previously existing memory value v_{old}, and the memory is rewritten with the value v_i corresponding to the lowest-numbered CPU i for which $v_i \neq \perp$.

We more generally support PRAM programs with a dynamic number of processors (i.e., m_i processors required for each time step i of the computation), as long as this sequence of processor numbers $m_1, m_2, \ldots$ is fixed, public information. The complexity of our OPRAM solution will scale with the number of required processors in each round, instead of the maximum number of required processors.

We consider the following *worst-case* metrics of a PRAM (over all inputs):

- $CPUs(\Pi)$: number of parallel processors required by Π.
- $\mathsf{space}(\Pi)$: largest database address accessed by Π.
- $\mathsf{time}(\Pi)$: maximum number of time steps taken by any processor to evaluate Π (where each Access is charged as a single step).[8]
- $\mathsf{comp}(\Pi)$: the total sum of all computation steps of active CPUs evaluating Π (which, for programs with fixed activation schedules as we consider, is a fixed value).

3 Local, Load-Balanced MPC for PRAM

Ultimately, we construct a protocol that securely realizes the ideal functionality $\mathcal{F}_{\mathsf{PRAMs}}$ (Fig. 1) for evaluating a sequence of PRAM programs (with bounded state maintained between program) on parties' fixed inputs. For simplicity of exposition, we assume each party has equal input size and receives the same

[8] We remark that the PRAM time complexity of any function f is bounded above by its circuit depth complexity (where the PRAM complexity of f is defined as the minimal value of $\mathsf{time}(\Pi)$ of any PRAM Π which evaluates f).

output. We further assume the total remnant state from one program execution to the next is bounded in size by the combined input size of all parties.[9]

Theorem 2 (Main Theorem). *For any constant $\epsilon > 0$ and polynomial parallelism parameter $k = k(n)$, there exists an n-party statistically secure (with error negligible in n) protocol realizing the functionality $\mathcal{F}_{\mathsf{PRAMs}}$, handling $(1/3 - \epsilon)$ fraction static corruptions with the following complexities, where each party is a k-processor PRAM (and where $|x|, |y|$ denote per-party input and output size, $\mathsf{space}(\Pi)$, $\mathsf{comp}(\Pi)$, and $\mathsf{time}(\Pi)$ denote the worst-case space, computation, and (parallel) runtime of Π, and $CPUs(\Pi)$ denotes the number of CPUs of Π):*

- *Computation per party, per Π_j: $\tilde{O}\big(\mathsf{comp}(\Pi_j)/n + |y|\big)$.*
- *Time steps, per Π_j: $\tilde{O}\left(\mathsf{time}(\Pi_j) \cdot \max\left\{1, \frac{CPUs(\Pi)}{nk}\right\}\right)$.*
- *Memory per party: $\tilde{O}\left(|x| + |y| + \max_{j=1}^{N} \mathsf{space}(\Pi_j)/n\right)$.*
- *Communication Locality: $\tilde{O}(1)$.*

given a one-time preprocessing phase with complexity:

- *Computation per party: $\tilde{O}(|x|)$, plus single broadcast of $\tilde{O}(1)$ bits.*
- *Time steps: $\tilde{O}\left(\max\left\{1, \frac{|x|}{k}\right\}\right)$.*

Additionally, the protocol achieves $\mathsf{polylog}(n)$ communication locality, and a strong "online" load-balancing guarantee:

Online Load Balancing: *For every constant $\delta > 0$, with all but negligible probability in n, the following holds at* all times *during the protocol: Let cc and $\mathsf{cc}(W_j)$ denote the total communication complexity and communication complexity of party P_j, comp and $\mathsf{comp}(P_j)$ denote the total computation complexity and computation complexity of party P_j, we have*

$$\frac{(1-\delta)}{n}\mathsf{cc} - \mathsf{polylog}(n) \leq \mathsf{cc}(P_j) \leq \frac{(1+\delta)}{n}\mathsf{cc} + \mathsf{polylog}(n)$$
$$\frac{(1-\delta)}{n}\mathsf{comp} - \mathsf{polylog}(n) \leq \mathsf{comp}(P_j) \leq \frac{(1+\delta)}{n}\mathsf{comp} + \mathsf{polylog}(n).$$

3.1 Proof of Main Theorem

At a very high level, the proof takes three steps: We first obtain MPC realizing $\mathcal{F}_{\mathsf{PRAMs}}$ with a weaker notion of *oblivious* security. We then show how to attain communication locality and load balancing, while preserving oblivious security. (This combines two steps described within the introduction). Finally, we convert the obliviously secure protocol to one secure in the malicious setting. We now proceed to describe these steps in greater technical detail.

[9] To support larger shared state size $\mathsf{space}^{\mathsf{Remnant}}$, the memory requirements of the protocol must grow with an extra additive $\tilde{O}(\mathsf{space}^{\mathsf{Remnant}})$.

Ideal Functionality $\mathcal{F}_{\mathsf{PRAMs}}$:
$\mathcal{F}_{\mathsf{PRAMs}}$ running with parties $P_1, \ldots, P_n$ and an adversary proceeds as follows. The functionality maintains longterm storage of parties' inputs $\{x_i\}_{i\in[n]}$ (each of equal size $|x|$), per-CPU state information state_i, and remnant memory $\mathsf{data}^{\mathsf{Remnant}}$ of total size $\mathsf{space}^{\mathsf{Remnant}} \in O(n \cdot |x|)$ transferred from computation to computation.

- Initialize $\mathsf{data}^{\mathsf{Remnant}} \leftarrow \emptyset$ and $\mathsf{state}_i \leftarrow \emptyset$ for each processor $i \in [m]$.
- Input Submission: Upon receiving an input $(commit, sid, input, x_i)$ from party P_i, record the value x_i as the input of P_i.
- Computation: Upon receiving a tuple $(compute, sid, \Pi, \mathsf{space}, \mathsf{time})$ consisting of an m-processor PRAM program Π, a space bound space, and a time bound time, execute Π as $(\mathsf{output}, \mathsf{state}_1, \ldots, \mathsf{state}_m, \mathsf{data}^{\mathsf{Remnant}}) \leftarrow \Pi(x_1, \ldots, x_n, \mathsf{state}_1, \ldots, \mathsf{state}_m, \mathsf{data}^{\mathsf{Remnant}})$ with the current value of state_i for each CPU $i \in [m]$. Send output to all parties.

Fig. 1. The ideal functionality $\mathcal{F}_{\mathsf{PRAMs}}$, corresponding to secure computation of a sequence of adaptively chosen PRAMs on parties' inputs.

Step 1: Oblivious-Secure MPC for PRAM. Intuitively, an adversary in the oblivious model is not allowed to corrupt any parties, and instead is restricted to seeing the "externally measurable" properties of the protocol (e.g., party response times, communication patterns, etc.).

Definition 1 Oblivious Secure MPC). *Secure realization of a functionality F by a protocol in the* oblivious model *is defined by the following real-ideal world scenario:*

Ideal World: Same as standard MPC without corrupted parties. That is, the adversary learns only public outputs of the functionality F evaluated on honest-party inputs.

Real World: Instead of corrupting parties, viewing their states, and controlling their actions (as in the standard malicious adversarial setting), the adversary is now limited as an external observer, and is given access only to the following information:

- *Activation Patterns: Complete list of tuples of the form*
  - $(\mathsf{timestep}, \mathsf{party\text{-}id}, \mathsf{compute\text{-}time})$*: Specifying all local computation times of parties.*
  - $(\mathsf{timestep}, \mathsf{party\text{-}id}, \mathsf{local\text{-}mem\text{-}addr})$*: Specifying all memory access patterns of parties.*
- *Communication Patterns: Complete list of tuples of the form*
  - $(\mathsf{timestep}, \mathsf{sndr\text{-}id}, \mathsf{rcvr\text{-}id}, \mathsf{msg\text{-}len})$*: Specifying all sender-receiver pairs, in addition to the corresponding communicated message bit-length.*

The output of the real-world experiment consists of the outputs of the (honest) parties, in addition to an arbitrary PPT function of the adversary's view at the conclusion of the protocol.

(Statistical) Security: For every PPT adversary $\mathcal{A}$ in the real-world execution, there exists a PPT ideal-world adversary $\S$ for which for every environment $\mathcal{Z}$, we have $\mathsf{output}_{\mathsf{Real}}(1^k, \mathcal{A}, \mathcal{Z}) \overset{s}{\cong} \mathsf{output}_{\mathsf{Ideal}}(1^k, \S, \mathcal{Z})$.

Toward our result, it will be advantageous to think of computations as composed of several sub-parts, or *"jobs,"* that each maintain and compute on small polylogarithmic-size state (Note that this is natural in the PRAM setting, where each CPU has polylogarithmic-size local memory). Later, to achieve load balancing, jobs will be assigned to and passed around between *"workers,"* so that each worker roughly performs the same amount of work. (The small state requirement per job will guarantee that "job passing" is not too expensive). Then, to obtain malicious security, each worker will ultimately be emulated by a committee of parties via small-scale MPCs; because of the polynomial overhead in the underlying MPC protocol, it will be important that this is only done for computations of $\mathsf{polylog}(n)$ size on $\mathsf{polylog}(n)$-size memory.

We now define the notion of a protocol in the *jobs model*.

Definition 2 (Jobs Model). *Let n be a security parameter. A* jobs protocol *consists of a $\mathsf{poly}(n)$-size set* Jobs *of agents (called* jobs*), and a distributed protocol description $\Pi_{\mathcal{J}}$, instructing each job to perform local computations and to communicate over a synchronized network (via point-to-point communication), with the following properties:*

- *Bounded memory: each job's space complexity is $w \in \mathsf{polylog}(n)$.*
- *Bounded per-round computation and communication: the computation and communication complexity of each job at each round is upper bounded by $w \in \mathsf{polylog}(n)$.*

A job is active *in a round if it performs computation within this round.*

A jobs protocol is further said to have injective communication *if the following property is satisfied:*

- *Injective communication: each round, a set of jobs are activated, and each sends a single $\mathsf{polylog}(n)$-sized message to a distinct job.*

By convention, we assume the first m_{in} jobs of a jobs protocol are *input jobs*, the last m_{out} are *output jobs*, and the remaining jobs are *helper jobs*. Each input job J_i holds a single-word input $x_i \in \{0,1\}^w$ (for $w \in \mathsf{polylog}(n)$); output and helper jobs have no input. We then have a canonical correspondence between functionalities in the standard n-party setting and the equivalent functionalities in the Worker-Jobs Model:

- Functionality $\mathcal{F}$: In the n-party setting. Accepts inputs x_i from each party P_i, evaluates $y \leftarrow F(x_1 || \cdots || x_n)$, outputs the resulting value y to all parties P_i.
- Functionality $\mathcal{F}^{\mathsf{Jobs}}$: In the Jobs Model. Accepts (short) inputs x_u^i from each Input Job, evaluates $y \leftarrow F(x_1 || \cdots || x_\ell)$, and distributes the resulting value y (in short pieces) to the Output Jobs.

We may analogously define *oblivious security* of a jobs protocol (where jobs are honest and the adversary sees only "externally measurable" properties of the protocol, as in Definition 1). Within the jobs model, we thus wish to securely realize the functionality $\mathcal{F}_{\mathsf{PRAMs}}^{\mathsf{Jobs}}$, equivalent to $\mathcal{F}_{\mathsf{PRAMs}}$ with the above syntactic change. Note that in the regime of oblivious security, a jobs protocol yields a *memory*-balanced protocol in the standard n-party model, by simply assigning jobs to the n parties evenly.

Theorem 3. *There exists an oblivious-secure protocol in the Jobs Model realizing the functionality $\mathcal{F}_{\mathsf{PRAMs}}^{\mathsf{Jobs}}$ for securely computing a sequence of N adaptively chosen PRAM programs Π_j, with the following complexities (where $n \cdot |x|, |y|$ denote the total input and output size, and $\mathsf{space}(\Pi)$, comp, and $\mathsf{time}(\Pi)$ denote the worst-case space, computation, and (parallel) runtime of Π over all inputs):*

- *Number of jobs:* $\tilde{O}\left(n \cdot |x| + |y| + \max_{j\in[N]} \mathsf{space}(\Pi_j)\right)$.
- *Computation complexity, per Π_j:* $\tilde{O}\big(\mathsf{comp}(\Pi_j)\big)$.
- *Time steps, per Π_j:* $\tilde{O}\left(\mathsf{time}(\Pi_j)\right)$.
- *The number of* active *jobs in each round is* $O(\max_{j\in[N]} CPUs(\Pi_j))$.

given a one-time preprocessing phase with complexity

- *Computation complexity:* $\tilde{O}(n \cdot |x|)$.
- *Time steps:* $\tilde{O}(1)$.

Further, the protocol has injective communication*: in each round, each activated job sends a single* $\mathsf{polylog}(n)$*-size message to a* distinct *job.*

Recall within the Jobs Model each job is limited to maintaining state of size $\mathsf{polylog}(n)$; thus the *memory requirement* of the above protocol is

$$\tilde{O}\Big(n \cdot |x| + |y| + \max_{j\in[N]} \mathsf{space}(\Pi_j)\Big),$$

based on the number of required jobs.

Idea of proof. The result builds upon the existence of an Oblivious PRAM compiler with $\mathsf{polylog}(n)$ time and space overhead that is *collision-free* (i.e., where no two CPUs must access the same memory address in the same timestep), which is guaranteed to exist unconditionally based on [BCP14b]. In addition to the standard Input and Output jobs, our protocol will have one Helper job for each of the CPUs and each memory cell in the database of the OPRAM-compiled program. The CPU jobs store the local state and perform the computations of their corresponding CPU. In each round that the ith CPU's instructions dictate a memory access at location $\mathsf{addr}^{(i)}$, the CPU job i will communicate with the Memory job $\mathsf{addr}^{(i)}$ to perform the access. (Thus, in each round, at most $2 \cdot CPUs(\mathsf{OPRAM}(\Pi))$ jobs are active, where $\mathsf{OPRAM}(\Pi)$ denotes the OPRAM-compilation of Π). Activation and communication patterns in the resulting protocol are simulatable directly by the OPRAM security. The preprocessing phase of the protocol corresponds to inserting all inputs into the OPRAM-protected database in parallel (i.e., emulating the OPRAM-compiled input insertion program that simply inserts each input x_i into address i of the database).

Step 2: Locality and Load Balancing. This step attains $\mathsf{polylog}(n)$ communication *locality*,[10] and computation *load balancing* from any jobs protocol $\Pi_{\mathcal{J}}$ with injective communication. We do so by emulating $\Pi_{\mathcal{J}}$ by a fixed set of parties (which we sometimes refer to as "workers"), where each worker is assigned several jobs, and will pass jobs to other workers once he has performed a certain amount of work. This yields a standard N-party protocol with a special *decomposable state* structure: i.e., parties' memory can be decomposed into separate $\mathsf{polylog}(n)$-size memory blocks, which are only ever computed on independently or in pairs, in steps of $\mathsf{polylog}(n)$ computation per round. This is because parties' computation is limited to individual jobs to which it was assigned.[11]

Definition 3 (Decomposable State). *An N-party protocol Π is said to have* decomposable state *if for every party P, the local memory mem of P can be decomposed into $\mathsf{polylog}(n)$-size blocks $\mathsf{mem} = (\mathsf{mem}_1, \mathsf{mem}_2, \ldots, \mathsf{mem}_m)$ such that: In each round of Π, the (parallel) local computation performed by party P is described as a list $\{(i, j, f_{i,j})\}_{(i,j)\in I}$ for some $I \subseteq [m] \times [m]$, such that each $f_{i,j}$ has complexity $\mathsf{polylog}(n)$. For each $(i, j) \in I$, party P executes $(\mathsf{mem}_i, \mathsf{mem}_j) \leftarrow f_{i,j}(\mathsf{mem}_i, \mathsf{mem}_j)$.*[12] *By convention, received communication messages are stored in local memory.*

We achieve the following "fully load-balanced" properties. Note that the first two properties correspond directly to our final load-balancing goal. The final property will be used to ensure that no individual worker is ever assigned drastically more than the expected number of simultaneous parallel computation tasks; this is important since workers will eventually be emulated by (technically, committees of) parties, who themselves may have bounded parallelism capability (i.e., small number of CPUs).

Definition 4 (Fully Load Balanced). *An N-party protocol Π is said to be* fully load balanced *with respect to security parameter n if the following properties hold:*

- *Memory load balancing: Let $\mathsf{space}(\Pi)$ denote the total space complexity of protocol Π. For every constant $\delta > 0$, with all but negligible probability in n, every party P_j has space complexity*

$$\mathsf{space}(P_j) \leq \frac{(1+\delta)}{N}\mathsf{space}(\Pi) + \mathsf{polylog}(n).$$

- *Online computation/communication load balancing: For every constant $\delta > 0$, with all but negligible probability in n, the following holds at* all times *during*

[10] Recall a protocol has (communication) locality $\ell(n)$ if during the course of the protocol every party communicates with at most $\ell(n)$ other parties.

[11] Looking ahead, pairwise computation will be used when emulating job-to-job communication, and will be sufficient when the original jobs protocol has *injective communication*, so that each job communicates with at most one other job per round.

[12] With some canonical resolution for write conflicts. (In our constructions, the sets (i, j) will be disjoint).

the protocol: Let cc *and* $\mathsf{cc}(P_j)$ *denote the total communication complexity and communication complexity of party* P_j, comp *and* $\mathsf{comp}(P_j)$ *denote the total computation complexity and computation complexity of party* P_j, *we have*

$$\frac{(1-\delta)}{N}\mathsf{cc} - \mathsf{polylog}(n) \leq \mathsf{cc}(P_j) \leq \frac{(1+\delta)}{N}\mathsf{cc} + \mathsf{polylog}(n)$$
$$\frac{(1-\delta)}{N}\mathsf{comp} - \mathsf{polylog}(n) \leq \mathsf{comp}(P_j) \leq \frac{(1+\delta)}{N}\mathsf{comp} + \mathsf{polylog}(n).$$

– *Per-round per-party efficiency:*[13] *Let* A *be an upper bound on the number of active jobs at each round in* $\Pi_{\mathcal{J}}$. *With all but negligible probability in* n, *the per-round per-party computation complexity is upper bounded by* $\tilde{O}(1+(A/N))$.

Theorem 4. *Let* $\Pi_{\mathcal{J}}$ *be an* M*-job protocol with computation complexity* comp *and* injective communication, *realizing functionality* $\mathcal{F}^{\mathsf{Jobs}}$. *Then there exists a* fully load-balanced *(Definition 4)* $\tilde{O}(n)$*-party protocol* $\Pi_{\mathcal{W}}$ *with* decomposable states *(Definition 3) that realizes* $\mathcal{F}$ *with total computation* $\tilde{O}(\mathsf{comp})$, *space complexity* $\tilde{O}(M)$, *and* $\mathsf{polylog}(n)$ *locality. If* $\Pi_{\mathcal{J}}$ *satisfies oblivious security, so does* $\Pi_{\mathcal{W}}$.

Idea of proof. Recall that in our construction of $\Pi_{\mathcal{W}}$ (in the introduction), at any point of the protocol execution, each job is assigned to a random worker[14] and is stored in at most 2 workers. This is sufficient to imply memory load balancing by standard concentration and union bounds. Online computation/communication load balancing follows by observing that (i) the job-passing pattern is independent of the worker-job assignment, and (ii) jobs are passed frequently enough before accumulating large cost. This allows us to think of the execution as partitioned into "job chunks" each of which is assigned to a random worker, thus amenable to concentration bounds. The last load-balanced property follows again by the fact that each job is independently assigned to a random worker and that each job only performs $\mathsf{polylog}(n)$ amount of work per round. To obtain locality, we consider a fixed low-degree communication network between workers, and pass messages using a load-balanced routing algorithm. Load balancing of this modified scheme follows by similar, but more delicate analysis.

The resulting protocol has *decomposable state*, since parties' memory and computation are completely local to individual jobs, or pairs of jobs in the case of emulating job-to-job communication (since the starting jobs protocol has injective communication).

Step 3: From Oblivious to Malicious Security. Finally, we present a general transformation that produces an n-party MPC protocol securely realizing a functionality $\mathcal{F}$ against $(1/3-\epsilon)n$ static corruptions, given any $\tilde{\Theta}(n)$-party protocol

[13] We note that the last two properties are related but incomparable. The online load balancing property focuses on accumulated work, whereas the per-round per-party efficiency concerns upper bounds on per-round work, which is used to bound the required amount of parallelism to execute the protocol with efficient parallel time.

[14] Technically, the initial job-worker assignment is only K-wise independent for $K = \log^3 n$. Nevertheless, this is sufficient for concentration bounds to go through.

with decomposable states (see Definition 3) realizing the corresponding jobs-model functionality $\mathcal{F}^{\mathsf{jobs}}$ with only *oblivious* security. This step can be viewed as a refinement and generalization of ideas from [KLST11, DKMS12, BGT13].

Theorem 5 (From Oblivious Security to Malicious Security). *Suppose there exists an $N \in \Theta(n \cdot \mathsf{polylog}(n))$-party* oblivious *protocol with decomposable state, realizing functionality $\mathcal{F}^{\mathsf{jobs}}$ in space, computation, and (parallel) time complexity* $\mathsf{space}, \mathsf{comp}, \mathsf{time}$. *Then for any constant $\epsilon > 0$ there exists an n-party MPC protocol (with error negligible in n) securely realizing the corresponding functionality $\mathcal{F}$ against $(1/3 - \epsilon)n$ static corruptions, with the following complexities (where each party is a PRAM with possibly many processors), given a one-time preprocessing phase with a single broadcast of $\tilde{O}(1)$ bits per party:*

- *Per-party memory: $\tilde{O}(\mathsf{space}/n)$.*
- *Total computation: $\tilde{O}(\mathsf{comp})$.*
- *Time complexity: $\tilde{O}(\mathsf{time})$.*

In addition, if the original protocol has $\tilde{O}(1)$ locality and is fully load-balanced *(i.e., satisfying all properties of Definition 4), then the resulting protocol additionally possesses the following properties:*

- *Communication locality $\tilde{O}(1)$.*
- *Online computation load balancing, as in Definition 4(c).*
- *Time complexity $\tilde{O}\left(\mathsf{time} \cdot \max\left\{1, \frac{A}{nk}\right\}\right)$ when each party is limited to being a k-processor PRAM, where A denotes the maximum per-round per-party computation complexity of any party in the original oblivious-secure protocol.*[15]

Idea of Proof. The compiler takes the following form: First, parties collectively elect a large number of "good" committees, each of size $\mathsf{polylog}(n)$, where "good" means each committee is composed of at least 2/3 honest parties, and that parties are spread roughly evenly across committees. The one-time broadcast is used to reach full agreement on the first committee. These committees will then emulate each of the *decomposable sub-computations* of the original protocol Π (see Definition 3), via small-scale MPCs. That is, committees are initialized with inputs by having the parties in Π' split their inputs into $\mathsf{polylog}(n)$-size pieces and verifiably secret share them to the appropriate committee(s). Each local computation (and communication) in Π decomposes as a collection of $f_{i,j}$, each affecting only two committees (emulating mem_i and mem_j). Since committees are only size $\mathsf{polylog}(n)$, and *each small-scale MPC has only* $\mathsf{polylog}(n)$ *memory and computation* (because of decomposability), the memory, computation, and time complexity overhead is small. Since parties are spread across committees, the protocol remains load balanced. Finally, by using a perfectly secure underlying MPC protocol (such as [BGW88]), the only information revealed corresponds directly to the "observable" properties (communication patterns, etc.), thus reducing directly to oblivious security (as per Definition 1).

[15] In particular, for our MPC for PRAMs protocol formed by combining Steps 1 and 2, the parameter A will correspond to the number of CPUs required in the evaluated PRAM Π, with $\mathsf{polylog}$ overhead.

References

[AAK08] Awerbuch, B., Azar, Y., Khandekar, R.: Fast load balancing via bounded best response. In: SODA 2008, pp. 314–322 (2008)

[ACMR95] Adler, M., Chakrabarti, S., Mitzenmacher, M., Rasmussen, L.E.: Parallel randomized load balancing (preliminary version). In: STOC 1995, pp. 238–247 (1995)

[AJLA+12] Asharov, G., Jain, A., López-Alt, A., Tromer, E., Vaikuntanathan, V., Wichs, D.: Multiparty computation with low communication, computation and interaction via threshold FHE. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 483–501. Springer, Heidelberg (2012)

[BCP14a] Boyle, E., Chung, K.-M., Pass, R.: Large-scale secure computation. Cryptology ePrint Archive, Report 2014/404 (2014)

[BCP14b] Boyle, E., Chung, K.-M., Pass, R.: Oblivious parallel ram. Cryptology ePrint Archive, Report 2014/594 (2014)

[BGJK12] Boyle, E., Goldwasser, S., Jain, A., Kalai, Y.T.: Multiparty computation secure against continual memory leakage. In: STOC, pp. 1235–1254 (2012)

[BGT13] Boyle, E., Goldwasser, S., Tessaro, S.: Communication locality in secure multi-party computation. In: Sahai, A. (ed.) TCC 2013. LNCS, vol. 7785, pp. 356–376. Springer, Heidelberg (2013)

[BGW88] Ben-Or, M., Goldwasser, S., Wigderson, A.: Completeness theorems for non-cryptographic fault-tolerant distributed computation (extended abstract). In: STOC, pp. 1–10 (1988)

[BSGH13] Braud-Santoni, N., Guerraoui, R., Huc, F.: Fast byzantine agreement. In: PODC, pp. 57–64 (2013)

[DI06] Damgård, I.B., Ishai, Y.: Scalable secure multiparty computation. In: Dwork, C. (ed.) CRYPTO 2006. LNCS, vol. 4117, pp. 501–520. Springer, Heidelberg (2006)

[DIK+08] Damgård, I., Ishai, Y., Krøigaard, M., Nielsen, J.B., Smith, A.: Scalable multiparty computation with nearly optimal work and resilience. In: Wagner, D. (ed.) CRYPTO 2008. LNCS, vol. 5157, pp. 241–261. Springer, Heidelberg (2008)

[DKMS12] Dani, V., King, V., Movahedi, M., Saia, J.: Breaking the o(nm) bit barrier: Secure multiparty computation with a static adversary. CoRR, abs/1203.0289 (2012)

[DMN11] Damgård, I., Meldgaard, S., Nielsen, J.B.: Perfectly secure oblivious RAM without random oracles. In: Ishai, Y. (ed.) TCC 2011. LNCS, vol. 6597, pp. 144–163. Springer, Heidelberg (2011)

[Gen09] Gentry, C.: Fully homomorphic encryption using ideal lattices. In: STOC, pp. 169–178 (2009)

[GGH+13] Gentry, C., Goldman, K.A., Halevi, S., Julta, C., Raykova, M., Wichs, D.: Optimizing ORAM and using it efficiently for secure computation. In: De Cristofaro, E., Wright, M. (eds.) PETS 2013. LNCS, vol. 7981, pp. 1–18. Springer, Heidelberg (2013)

[GKK+11] Gordon, S.D., Katz, J., Kolesnikov, V., Malkin, T., Raykova, M., Vahlis, Y.: Secure computation with sublinear amortized work. IACR Cryptology ePrint Archive, 2011:482 (2011)

[GMW87] Goldreich, O., Micali, S., Wigderson, A.: How to play any mental game or a completeness theorem for protocols with honest majority. In: STOC, pp. 218–229 (1987)

[GO96] Goldreich, O., Ostrovsky, R.: Software protection and simulation on oblivious rams. J. ACM **43**(3), 431–473 (1996)

[KLST11] King, V., Lonargan, S., Saia, J., Trehan, A.: Load balanced scalable byzantine agreement through quorum building, with full information. In: Aguilera, M.K., Yu, H., Vaidya, N.H., Srinivasan, V., Choudhury, R.R. (eds.) ICDCN 2011. LNCS, vol. 6522, pp. 203–214. Springer, Heidelberg (2011)

[KSSV06] King, V., Saia, J., Sanwalani, V., Vee, E.: Scalable leader election. In: SODA, pp. 990–999 (2006)

[LO13] Lu, S., Ostrovsky, R.: Distributed oblivious RAM for secure two-party computation. In: Sahai, A. (ed.) TCC 2013. LNCS, vol. 7785, pp. 377–396. Springer, Heidelberg (2013)

[MPS02] Mitzenmacher, M., Prabhakar, B., Shah, D.: Load balancing with memory. In: Proceedings, (FOCS), 16–19 November 2002, Vancouver, BC, Canada, pp. 799–808 (2002)

[MR98] Muthukrishnan, S., Rajaraman, R.: An adversarial model for distributed dynamic load balancing. In: SPAA 1998, pp. 47–54 (1998)

[MSS13] Myers, S., Sergi, M., Shelat, A.: Black-box proof of knowledge of plaintext and multiparty computation with low communication overhead. In: Sahai, A. (ed.) TCC 2013. LNCS, vol. 7785, pp. 397–417. Springer, Heidelberg (2013)

[OS97] Ostrovsky, R., Shoup, V.: Private information storage (extended abstract).In: STOC, pp. 294–303 (1997)

[VB81] Valiant, L.G., Brebner, G.J.: Universal schemes for parallel communication. In: STOC, pp. 263–277 (1981)

[Yao86] Yao, A.C.-C.: How to generate and exchange secrets (extended abstract). In: FOCS, pp. 162–167 (1986)

[ZMS14] Zamani, M., Movahedi, M., Saia, J.: Millions of millionaires: Multiparty computation in large networks. Cryptology ePrint Archive, Report 2014/149 (2014)