



# NFGen: Automatic Non-linear Function Evaluation Code Generator for General-purpose MPC Platforms

Xiaoyu Fan  
fanxy20@mails.tsinghua.edu.cn  
IIIS, Tsinghua University

Kun Chen  
chenkun@tsingj.com  
Tsingjiao Information Technology Co. Ltd.

Guosai Wang  
guosai.wang@tsingj.com  
Tsingjiao Information Technology Co. Ltd.

Mingchun Zhuang  
mczhuang@bupt.edu.cn  
Beijing University of Posts and Telecommunications

Yi Li  
xiaolixiaoyi@tsingj.com  
Tsingjiao Information Technology Co. Ltd.

Wei Xu  
weixu@tsinghua.edu.cn  
IIIS, Tsinghua University

## ABSTRACT

Due to the absence of a library for non-linear function evaluation, so-called *general-purpose* secure multi-party computation (MPC) are not as “general” as MPC programmers expect. Prior arts either naively reuse plaintext methods, resulting in suboptimal performance and even incorrect results, or handcraft *ad hoc* approximations for specific functions or platforms. We propose a general technique, NFGen<sup>1</sup>, that utilizes pre-computed *discrete piecewise polynomials* to accurately approximate generic functions using fixed-point numbers. We implement it using a performance-prediction-based code generator to support different platforms. Conducting extensive evaluations of 23 non-linear functions against six MPC protocols on two platforms, we demonstrate significant performance, accuracy, and generality improvements over existing methods.

## CCS CONCEPTS

• Security and privacy → Security services.

## KEYWORDS

Secure Multi-Party Computation (MPC), Non-linear Function Evaluation, Automatic Code Generation

## ACM Reference Format:

Xiaoyu Fan, Kun Chen, Guosai Wang, Mingchun Zhuang, Yi Li, and Wei Xu. 2022. NFGen: Automatic Non-linear Function Evaluation Code Generator for General-purpose MPC Platforms. In *Proceedings of Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22)*. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3548606.3560565>

## 1 INTRODUCTION

Privacy-preserving computation, especially *secure multi-party computation (MPC)*, has attracted a lot of attention in both academia and industry. They provide a promising trade-off between mining the data and privacy protection. People have proposed many

*general-purpose MPC platforms* [7, 21, 22, 25, 33, 37] that provide high-level abstractions and practical performance, allowing people to develop secure data processing applications without understanding the details of underlying MPC protocols.

Most platforms use a version of *secret sharing* (SS) protocols to build basic secure operations like  $+$ ,  $\times$ , and comparison (e.g.,  $>$ ), and then construct complex functions by composing them, just like writing plaintext expressions. The security of compound operations/functions is guaranteed by the *universal composability* [10] of these protocols. These platforms usually provide built-in support for common non-linear functions such as reciprocal ( $\frac{1}{x}$ , for real number divisions), exponential ( $e^x$ ), logarithm ( $\ln x$ ), and square root ( $\sqrt{x}$ ). They implement these functions either using generic numerical methods (e.g. the *Newton method*) or adopting protocol-specific algorithms like in [15, 35].

It remains a big challenge, however, to support the large variety non-linear functions in scientific computing and machine learning, such as  $\chi^2$  test and *sigmoid*. The naive approach is to compose them with built-in functions. E.g., we can compute  $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$  by composing two  $e^x$ , one  $\frac{1}{x}$  and two  $+$ . We refer to this approach as *direct evaluation*. Unfortunately, there are four pitfalls.

**Pitfall 1: Correctness and precision.** Most practical MPC platforms use *fixed-point (FXP)* numbers instead of the common *floating-point (FLP)* ones for efficiency. Although there are attempts to support FLP in MPC [2], the low performance of  $+$  prevents people from adopting it. FXP still dominates the practical MPC platforms [7, 21, 22, 25, 33, 37]. Unfortunately, ignoring the differences between FXP and FLP leads to two severe issues:

First, FXP supports a *much smaller range and resolution* than FLP, leading to more overflow/underflow as well as precision loss. Even worse, FXP cannot represent *NaN* and *Inf* like FLP. Also, the inputs/outputs on MPC are in ciphertext, so there is no way to detect overflows. Using *tanh* as an example, even if the function has a range of  $[-1, 1]$ , the intermediate results,  $e^x$  and  $e^{-x}$ , can easily overflow when  $|x|$  is large. In plaintext, people use a *scaling conversion* to enlarge the range of FXP, but in MPC, the encrypted  $x$  makes scaling costly. In fact, the built-in *tanh* function in MP-SPDZ [21] gives wrong results if  $|x| > 44$ , even if we increase FXP width to 128 bits. Second, each non-linear function has a precision loss that *accumulates* if we compose them with multiple steps. Section 7.3 shows more examples of both issues.

<sup>1</sup>The source code is released in <https://github.com/Fannxy/NFGen>



This work is licensed under a Creative Commons Attribution International 4.0 License.

**Pitfall 2: Performance.** Even with aggressive optimizations, non-linear function evaluation takes significant time using MPC. Depending on the platform, they can be orders of magnitude slower than the plaintext version. We measure the relative performance between non-linear functions vs. basic operations in 6 MPC protocols and observe dramatic differences (Table 1 in Section 7.1). Composing these functions sequentially makes things even slower.

**Pitfall 3: Generality.** Although we can write many non-linear functions using the built-ins, some functions are hard to implement. For example, functions  $\gamma(x, z)$ ,  $\Gamma(x, z)$  and  $\Phi(x)$  are defined as integrals. It is very tedious, if not impossible, to implement them in MPC using numerical methods and built-in operations.

**Pitfall 4: Portability.** Even if we can afford the engineering effort to build a complete scientific computation library, there are too many performance trade-offs to make it portable to different MPC systems and applications. MPC systems use a variety of protocols (to support different security assumptions), number representations and sizes, as well as custom programming languages. Also, they are deployed in different hardware/software/network environments. We want the computation library to maintain efficiency in all cases with minimal porting efforts.

In this paper, we offer a new scheme, *non-linear function code generator* (NFGen), to evaluate non-linear functions on general-purpose MPC platforms. We approximate each non-linear function using an  $m$ -piece *piecewise polynomial* with max order  $k$  (we automatically determine  $k$  and  $m$ ). Our approach has three advantages. First, it only uses secure  $+$ ,  $\times$  and  $>$  operations supported by all popular MPC platforms, and we can prove that the security properties are the same as the underlying platform with the same adversarial models. Second, the evaluation is *oblivious*, i.e., the operation sequence is not dependent on input data, allowing for predictable running time and avoiding timing-related side-channels. Third, obtaining the approximation is independent of input data and hence can be precomputed on plaintext.

**Key challenges.** Finding a good  $(k, m)$ -*piecewise polynomial* approximation for MPC platforms is a challenging problem. First, fitting polynomials for FXP computation introduces many challenges: 1) the polynomial needs to meet both the *range* and the *resolution* requirements of FXP, making sure all intermediate steps neither overflow nor underflow; 2) an FXP is essentially an integer, making the polynomials discrete. Fitting one polynomial minimizing the approximation error is an NP-complete *integer programming (IP)* problem. Thus we need to find an approximation. Second, there is a trade-off between  $m$  and  $k$ : whether we get more pieces (mainly leading to more  $>$ 's) or use a higher-order polynomial (leading to more  $\times$ 's). The choice depends on the performance of the specific MPC system, as we discussed above.

**Method overview.** We compute the polynomial fitting as a pre-computation step in plaintext. In a nutshell, we first construct a polynomial with order  $k$  in FLP using *Chebyshev Interpolation* [38] (or *Lagrange Interpolation* [30] for corner case) and then discretized it into FXP. We check the accuracy of the FXP polynomial using random data samples. If it does not achieve the user-specified accuracy, we split the input domain into two and recurse on each smaller domain. We design a series of algorithms leveraging the FLP capability to help find a better FXP approximation, like using

two FXPs to expand the *range* and improving the precision through residual functions. Section 4 provides the details of the algorithms.

At runtime, we evaluate the  $(k, m)$ -piecewise polynomial in an *oblivious* way, i.e. the execution only depends on  $(k, m)$ , but not input data. We design an *oblivious piecewise polynomial evaluation (OPPE)* algorithm (Section 4.3). To get a good  $(k, m)$  trade-off on different MPC platforms, NFGen uses a profiler to collect performance metrics of a specific deployment of an MPC platform and learns a model to predict the performance with different  $(k, m)$ . NFGen automatically makes the choice using the prediction and generates MPC-platform-specific OPPE code using built-in code templates. NFGen provides templates for both PrivPy [25] and MP-SPDZ [21]. The template is the only platform-dependent part in NFGen, and it only takes a short template to port NFGen to a new MPC platform.

**Evaluation results.** We evaluate NFGen against 6 secret sharing protocols using 15 commonly used non-linear functions on both PrivPy [25] and MP-SPDZ [21]. We observe significant performance improvements over baselines (direct evaluation) in 93% of all cases with an average speedup of  $6.5\times$  and a maximum speedup of  $86.1\times$ . NFGen saves 39.3% network communications on average. We also show that we can avoid the overflow errors and achieve much better accuracy comparing with the baseline, and allow a larger input domain even with small FXP widths. Using logistic regression (LR) as an example, we demonstrate performance and accuracy improvements over direct evaluation and *ad hoc* approximations, with  $3.5\times$  speedup and 0.6% – 14% accuracy improvements. Additionally, we illustrate how NFGen helps users to easily implement otherwise hard-to-implement functions on MPC by using 8 complex functions defined as integrals and the  $\chi^2$  test on real data.

In summary, our contributions include:

- 1) We propose a series of algorithms to fit an effective *piecewise polynomial* approximation on plaintext, fully considering the differences between FLP and FXP representations.
- 2) We design and implement the code generator with automatic profiling to allow portability to different MPC systems with distinct performance characteristics.
- 3) We conduct comprehensive evaluations against six MPC protocols on two platforms over 23 non-linear functions, showing significant improvements in performance, accuracy, portability to different MPC platforms, usability and savings in communications.

## 2 BACKGROUND AND RELATED WORK

In this section, we first introduce the background of general-purpose MPC platforms and FXP numbers for readers unfamiliar with this area and then review related works.

### 2.1 General-purpose MPC Platforms

General-purpose MPC platforms usually provide a high-level programming front-end with a compiler/interpreter to translate a high-level code into a series of cryptographic building blocks [19]. They provide common operators like  $+$ ,  $\times$ ,  $>$  and  $\frac{1}{x}$  in the front-end, and implement these operations using MPC protocols. These platforms guarantee privacy in the end-to-end algorithms based on the *universal composability* of the underlying security protocols. Example

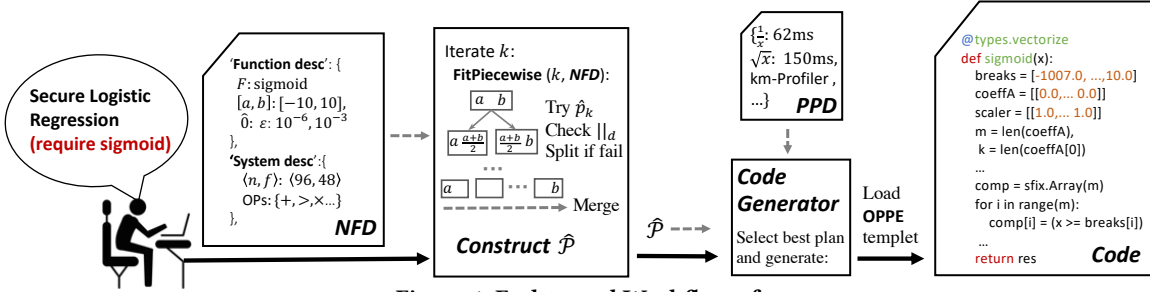


Figure 1: End-to-end Workflow of NFGen

platforms include MP-SPDZ [21], PrivPy [25], CryptTen [22], SecureML [33], ABY [17], and CryptGPU [37] etc. They hugely lower the barriers of developing privacy-preserving applications.

Most of the general-purpose MPC platforms use *secret sharing* (SS) [17, 22, 25, 33, 37] as the underlying protocol. There are a range of security assumptions (e.g., semi-honest vs. malicious) in these protocols, resulting in significantly differing performance. Some platforms allow users to choose the underlying protocols, like [21]. Our goal is to make NFGen protocol-agnostic.

Secret sharing protocols require communications among the participants on each operation, which results in significantly slower performance than plaintext (usually  $10 \times - 1000 \times$  slower, depending on the protocol and network environment). Also, different operators in the same protocol exhibit vast performance difference. For example, in the most common *additive* secret sharing protocol,  $+$  is almost as fast as plaintext as it is communication-free. However,  $\times$  and  $>$  require communication and are slower. Non-linear operations like  $\frac{1}{x}$  and  $\sqrt{x}$  operate more slowly. It is important to make  $+$  as fast as it is the most common operation in many applications.

To make  $+$  fast, MPC systems need to avoid floating point (FLP) numbers because the ciphertext *exponent* in FLP prevents us from adding up FLPs directly. Thus, almost all practical MPC platforms use fixed-point (FXP) numbers [11, 22, 25, 32, 33, 37]. FXP represents each real number as an  $n$ -bit integer, out of which  $f$  least significant bits represent the fraction, and the most significant bit is the sign. We denote it as a  $\langle n, f \rangle$ -FXP number. In the  $\langle n, f \rangle$  format, both the *range* and *resolution* of the FXP are fixed to  $2^{n-f-1}$  and  $2^{-f}$ , respectively. FXP offers a much smaller range and resolution than FLP, and thus programmers need to be more careful with overflows and precision losses. MPC further complicates the problem as we cannot detect overflows on ciphertext. Also, the common scaling method in plaintext FXP requires additional bit operations in ciphertext and thus becomes expensive in MPC.

## 2.2 Non-linear Functions in MPC Platforms

Lacking of a complete general numeric computation library, MPC application developers need to roll their application-specific solutions even for common functions, like logistic regression (LR) [18], decision trees [15, 27], principal component analysis (PCA) [39] and neural networks (NNs) [22, 35, 37]. Mohassel *et al.* [32, 33] use a 3-piece linear function to replace the slow *sigmoid* function in LR and NNs. However, there is no guarantee that the approximation maintains accuracy in all cases. In fact, we find cases with significant LR accuracy loss using this approximation.

The most straightforward method to evaluate non-linear functions in MPC is directly adapting the plaintext code and replacing basic operations to secure ones (e.g.,  $+$ ,  $\times$  and  $>$ ). For example, CryptTen [22] and CryptGPU [37] adopt a series of plaintext algorithms like *Newton-Raphson iterations*, *limit approximations* to implement  $\frac{1}{x}$ ,  $e^x$ ,  $\sqrt{x}$  and  $\ln x$ . Natively adopting plaintext code usually leads to poor performance, as we do not know when we have reached the desirable accuracy and need to iterate more.

There are also protocol-specific approaches. For example, Rathee *et al.* [35] implement 10 efficient cryptographic building blocks to support high-performance non-linear functions like *reciprocal-of-sqrt*, *sigmoid* and *exponent*. However, these functions are optimized for their 2-party-computation platform only. Similarly, Damgård *et al.* [15] propose a series of primitives to support efficient machine learning applications for SPDZ<sub>2k</sub> protocol [13]. More generally, Catrina *et al.* [11] design a collection of general building-blocks suitable for any fixed-point MPC platform and propose an optimized  $\frac{1}{x}$  primitive using *Goldschmidt* algorithm on secret FXP. MP-SPDZ [21] uses the same algorithm.

Another line of methods is to approximate non-linear functions with polynomials. It is a well-studied problem in plaintext of finding an optimal polynomial approximating a given function  $F(x)$ , minimizing the maximum difference over a given input domain  $[a, b]$ . Such polynomial is referred to as *minimax polynomial* [3]. *Chebyshev interpolation* [38] is a well-known solution. It offers a close approximation to the *minimax polynomial* (so-called *Chebyshev (near) minimax polynomial*) [3, 38]. Hesamifard *et al.* [20] adopt the Chebyshev polynomial to evaluate *sigmoid* on homomorphic encryption. However, the range of FXP limits the order  $k$  of the polynomial and prevents it from reaching the desired approximation accuracy either. Another challenge is that the limited resolution of FXP cannot capture sometimes-tiny coefficients in these polynomials. Previous work ignores this problem and leads to large errors [20]. Boura *et al.* [8] propose to use *Fourier series* to approximate *sigmoid* in MPC platform. They work around the range and resolution issues using secure quadruple-precision FLP but brings in expensive computation overhead.

Unlike previous efforts, we adopt *piecewise polynomial* and fit a Chebyshev polynomial for each piece, resulting in improved accuracy. We consider the differences between FXP and FLP and handle corner cases. Also, prior works focus on one or a few functions

on a single MPC platform, whereas our goal is to build a general, cross-platform solution for all *Lipschitz-continuous* functions<sup>2</sup>.

### 3 OVERVIEW

**Notations and assumptions.** We first describe the notations and assumptions we use throughout the paper. We assume  $F(x)$  is a Lipschitz continuous function, and we can evaluate  $F(x)$  in plaintext using FLP. NFGGen approximates  $F(x)$  by finding a set of feasible *piecewise polynomials* with different max order  $k$  and the number of pieces  $m$ . We denote the set of all feasible polynomials as  $\hat{\mathcal{P}} = \{\hat{p}_k^m\}$ . Each of the  $\hat{p}_k^m$  contains  $m$  pieces covering the entire input domain of  $[a, b]$  as  $[w_0, w_1, \dots, w_j, \dots, w_m]$  ( $w_0 = a$  and  $w_m = b$ ). In each piece  $j$ , (i.e.,  $[w_{j-1}, w_j]$ ), we have a polynomial  $\hat{p}_k^{(j)}(\hat{x}) = \sum_{i=0}^k \hat{c}_i \hat{x}^i$  to approximate it. When the piece index  $j$  is not important, we use a shorthand of  $\hat{p}_k(\hat{x})$  to denote it. Throughout the paper, symbols with a *hat* ( $\hat{\cdot}$ ) denote  $\langle n, f \rangle$ -FXP representable variables and the definition of functions like  $\hat{p}_k^m$  means all the coefficients and terms are in  $\langle n, f \rangle$ -FXP representation (Section 2.1). We assume 64-bit double-precision FLP is equivalent to  $\mathbb{R}$ . It is a safe assumption in our situation, considering both the resolution and range of FLP are orders of magnitudes larger than FXP of our concern.

**NFGGen input files.** NFGGen generates non-linear function approximation code for different MPC platforms according to two input files. The first is a user-provided *non-linear function definition* (NFD), containing the expression of the target function  $F(x)$ , its domain  $[a, b]$ , FXP format  $\langle n, f \rangle$ , target accuracy  $\epsilon$  and  $\hat{0}$  (in Eq. 1), and a list of the operators supported by the target MPC platform. The users can generate the second input file, *performance profile definition* (PPD), running a NFGGen-provided profiler on the target MPC platform deployment. Note that NFD is MPC-platform-specific but independent of the actual *deployment* (i.e., the CPU and networking configurations), and PPD describes the deployment. Separating them allows better portability across different deployments.

**Workflow.** Figure 1 illustrates NFGGen workflow. NFGGen first reads in the NFD file, and on plaintext runs the algorithms in Section 4 to fit the set of  $\hat{\mathcal{P}}$  with different  $k$  and  $m$  settings. Then using the PPD file, NFGGen chooses one  $\hat{p}_k^m \in \hat{\mathcal{P}}$  with the  $k$  and  $m$  that maximizes performance on the specific deployment (Section 6.1). Finally, NFGGen outputs the generated code that runs just like any user-defined function on the target MPC system, using a set of pre-defined, platform-dependent code templates.

**Requirements for a feasible  $\hat{p}_k^m$ .** Given an MPC platform with  $\langle n, f \rangle$ -FXP, the target function  $F(x)$  on input domain  $[a, b]$ , a feasible  $\hat{p}_k^m$  needs to meet the following three conditions.

- 1) The evaluation of  $\hat{p}_k^m$  should only consist of provided operators in the target MPC platform. This requirement is usually true as all we need are basic operators of  $+$ ,  $\times$  and  $>$ .
- 2) All the intermediate results in  $\hat{p}_k^m$ , including polynomial coefficients  $\hat{c}_i$  and  $\hat{x}^i$  terms, etc., should be representable in  $\langle n, f \rangle$ -FXP without overflow or underflow.

- 3) For all  $\hat{x}$  in the range  $[a, b]$ , the  $\hat{p}_k^m$  should approximate  $F(x)$  with high accuracy, so we need to bound the *max error* of the approximation rather than the mean error. We measure the approximation error using the *soft relative distance* (SRD),

$$|x - y|_d = \begin{cases} |x - y|/|x|, & |x| > \hat{0} \\ |x - y|, & |x| \leq \hat{0} \end{cases}, \hat{0} < \epsilon, \quad (1)$$

and require

$$\max_{\hat{x} \in [a, b]} |F(\hat{x}) - \hat{p}_k^m(\hat{x})|_d \leq \epsilon. \quad (2)$$

The  $\hat{0}$  is the *soft zero*. We use soft zeros because the relative error can be very large when  $|x| \rightarrow 0$ . For example, for  $x = 2^{-50}$  (in FLP), a good representable approximation in a  $\langle 96, 48 \rangle$ -FXP,  $\hat{x} = 2^{-48}$  gives a large relative error of  $2^2 - 1 > \epsilon$ . To avoid ruling out these good-performance approximations, we switch to bound the *absolute error* instead when  $|x| \leq \hat{0}$ . By default, we set  $\epsilon = 10^{-3}$  and  $\hat{0} = 10^{-6}$ . We further relax the accuracy definition to maximum *sample* SRD rather than the true maximum SRD by computing the max error over a sample set of  $\hat{x} \in [a, b]$  for practical performance. Indeed, we prove that for any Lipschitz continuous function  $F(x)$ , the true maximum SRD can be bounded by the sampled version, for details, see [40]. Empirically, we find that a modest sample set (e.g. 1000 per piece) is frequently sufficient (Section 7.3).

## 4 NON-LINEAR APPROXIMATION

The core of NFGGen is to fit a set of piecewise polynomials  $\hat{\mathcal{P}}$  that estimate the function  $F(x)$  in plaintext. All  $\hat{p}_k^m \in \hat{\mathcal{P}}$  can be evaluated obviously in ciphertext. We first introduce the overall algorithm that recursively finds a good  $m$  for a given  $k$  if possible. Then we focus on the algorithm that fits a  $\hat{p}_k$  for a single piece, which is the most challenging part due to FXP limitations. Finally, we introduce the oblivious evaluation algorithm.

### 4.1 Fitting Piecewise Polynomials

The cost for evaluating  $\hat{p}_k^m$  is mostly for computing 1) the  $k$ -th order polynomial and 2) deciding which of the  $m$  pieces that input  $\hat{x}$  belongs to. Each  $\hat{p}_k^m$  has  $(m \times k)$  parameters. We want to determine  $m$  and  $k$  automatically. As different MPC platforms may have different  $>$  and  $\times$  performance, we want to generate multiple plans with different  $(k, m)$  choices and let the latter stages (Section 6.1) to decide which one to use. This step only takes 2-3 seconds on plaintext in most of our experiments.

We iterate through a number of  $k$  values. For each  $k$ , we use Algorithm 1 to find an  $m$  piece polynomial as a candidate if possible. We start with the user-defined domain  $[a, b]$ . We fit a best-effort  $k$ -th order polynomial  $\hat{p}_k$  using `FitOnePiece` subroutine (Algorithm 2) minimizing the maximum absolute error (Line 2). The fitting process is quite involved as we need to deal with the limited range and resolution of FXPs. We leave the details of fitting  $\hat{p}_k$  to Section 4.2.

`FitOnePiece` (Algorithm 2) returns a feasible  $\hat{p}_k$  satisfying both representability and accuracy constraints in domain  $[a, b]$  if it successfully finds it. If it failed, it returns `Null`, which means that the order  $k$  is not enough to fit  $F(x)$  in domain  $[a, b]$ , and thus we split  $[a, b]$  into  $[a, (a+b)/2]$  and  $[(a+b)/2, b]$  and recurse on each smaller ranges (Line 5-10).

<sup>2</sup>Even if the functions are Lipschitz continuous, there is no theoretical guarantee that we will find a feasible approximation for *all* functions on any accuracy requirements. However, empirically, NFGGen works well on all the functions we tested.

**Algorithm 1: FitPiecewise Algorithm**


---

**Global config:** FXP format  $\langle n, f \rangle$ , max sampling numbers  $MS$  and max pieces  $m_{max}$ .  
**Global state** : The fitted piecewise polynomial  $\hat{p}_k^m$ .  
**Input** : Target function  $F(x)$ , input domain  $[a, b]$  and order  $k$ .

---

```

1 Initialize piece counter  $m_c \leftarrow 0$ ;
2 if  $\hat{p}_k \leftarrow \text{FitOnePiece}(F, [a, b], k)$  is NOT Null then
3   | Add  $\hat{p}_k$  to global state  $\hat{p}_k^m$ ;
4 end
5 else
6   |  $\text{FitPiecewise}(F, [a, \frac{a+b}{2}], k)$ ;
7   |  $\text{FitPiecewise}(F, [\frac{a+b}{2}, b], k)$ ;
8   |  $m_c + 1$ ;
9   | If  $m_c > m_{max}$ : Exit;
10 end
11  $i \leftarrow 0$ ;
12 while  $i$  not reach the tail of  $\hat{p}_k^m$  do
13   |  $a, b \leftarrow w_i, w_{i+2}$  in  $\hat{p}_k^m$ ;
14   | if  $\hat{p}_k \leftarrow \text{FitOnePiece}(F, [a, b], k)$  is NOT Null then
15     | Replace  $\hat{p}_k^{(i)}$  and  $\hat{p}_k^{(i+1)}$  with single  $\hat{p}_k$ ;
16   | else
17     |  $i + 1$ ;
18   | end
19 end
20 Exit

```

---

As a final step, we try to merge adjacent pieces because splits may result in unnecessary pieces. For each adjacent pair of pieces, we try `FitOnePiece` (Algorithm 2) again to fit a single polynomial in the combined range with the accuracy requirement satisfied (Line 11-19). Finally, we get the set of all  $m$   $k$ th-order polynomials  $\hat{p}_k$ 's, constructing candidate  $\hat{p}_k^m$ .

The algorithm eventually terminates, either when  $m$  exceeds the limit  $m_{max}$  (Line 9), or finds a  $\hat{p}_k^m$  that passes the accuracy test. There is no theoretical guarantee that the algorithm will find a feasible solution or guarantee for the optimality. However, it works well empirically on all functions in our tests (Section 7.3).

## 4.2 Fitting Polynomial for One Piece

We introduce the core part of Algorithm 1, the `FitOnePiece` function in Algorithm 2, fitting a single  $\hat{p}_k$  in the domain of  $[a, b]$ .

**Problem Definition.** We want to find a  $k$ th-order polynomial  $\hat{p}_k = \sum_{i=0}^k (\hat{c}_i \hat{x}^i)$  that approximates  $F(x)$  over the domain of  $\hat{x} \in [a, b]$ , minimizing the max error. It is important that we limit the max error instead of the mean error to avoid occasional wrong results. Formally, we define the following optimization problem.

$$\begin{aligned}
 & \text{Minimize } \max_{\hat{x} \in [a, b]} |F(\hat{x}) - \hat{p}_k(\hat{x})| \\
 & \text{s.t., } \hat{p}_k = \sum_{i=0}^k (\hat{c}_i \hat{x}^i)
 \end{aligned} \tag{3}$$

**Algorithm 2: FitOnePiece Algorithm**


---

**Input** : Target function  $F(x)$ , domain  $[a, b]$  and order  $k$ .  
**Return** : Feasible discrete polynomial  $\hat{p}_k$  or Null.

---

```

1  $\bar{k} \leftarrow \text{ConstrainK}([a, b], \langle n, f \rangle) /* 1) Constrain k */
/* 2) Fit best polynomial in FLP space */
2 Maximum representable points  $N \leftarrow \frac{b-a}{2^{-f}}$ ;
3 if  $N > \bar{k} + 1$  then
4   |  $p_{\bar{k}} \leftarrow \text{Cheby-Interpolation}(F, [a, b], \bar{k})$ 
5 else
6   |  $\bar{k} = N - 1$ ;
7   |  $p_{\bar{k}} \leftarrow \text{Lagrange-Interpolation}(F, N \text{ feasible points and } \bar{k})$ 
8 end
/* 3) & 4) Convert to FXP space */
9  $\hat{p}_{\bar{k}} \leftarrow \text{ScalePoly}(p_{\bar{k}}, [a, b])$  (Algo 5);
10  $\hat{p}_{\bar{k}} \leftarrow \text{ResidualBoosting}(\hat{p}_{\bar{k}}, F, [a, b])$  (Algo 6);
11  $\hat{p}_k$ : Expand coefficients and scaling factors of  $\hat{p}_{\bar{k}}$  to  $k$ , filling 0;
/* 5) Check accuracy, return valid  $\hat{p}_k$  or Null */
12 Sampled number  $N_s \leftarrow \min(MS, N)$ ;
13  $\hat{X} \leftarrow \text{FLPsimFXP}(\text{Linspace}([a, b], N_s))$ ;
14 if  $\max_{\hat{x} \in \hat{X}} |\hat{p}_k(\hat{x}) - F(\hat{x})|_d < \epsilon$  then
15   | Return:  $\hat{p}_k$ ;
16 Return: Null;$ 
```

---

It is easy to show that Eq. 3 is an NP-Complete *integer programming (IP)* problem, because

$$\hat{p}_k(\hat{x}) = \sum_{i=0}^k \hat{c}_i \hat{x}^i = \sum_{i=0}^k ((\hat{c}_i \cdot 2^f) \cdot \frac{\hat{x}^i}{2^f}) = \sum_{i=0}^k (\bar{c}_i \cdot \frac{\hat{x}^i}{2^f}), \tag{4}$$

where coefficients  $\bar{c}_i$ 's are  $n$ -bit integers. We present an effective approximation by firstly solve the optimal polynomial  $p_k$  in continuous space and then discretize it to  $\hat{p}_k$  and optimize it in FXP.

**Issues in FXP approximation.** FLP offers a much larger range compared with FXP. For 64-bit double precision FLP, the representable range is from  $-2^{1024}$  to  $2^{1024}$  (IEEE754 standard [1]), while even for  $\langle 128, 48 \rangle$ -bit FXP, the range is only  $-2^{79}$  to  $2^{79}$ . Thus, overflows are more common in FXP. For precision, double precision FLP can represent the smallest number of  $2^{-1023}$ , while FXP only has a fixed  $f$ -bit resolution. Any number smaller than  $2^{-f}$  is rounded off to zero. Unfortunately, prior MPC algorithms do not handle FXP correctly, leading to wrong results even if both the domain and range are representable.

Specifically, We need to find a  $\hat{p}_k = \sum_{i=0}^k \hat{c}_i \hat{x}^i$  in FXP to approximate  $p_k = \sum_{i=0}^k c_i x^i$  in FLP and avoid the following three issues.

**Issue 1)**  $\hat{x}^k$  can overflow if  $|\hat{x}|$  is too large, or underflow if  $\hat{x}$  is close to zero, especially with a large  $k$ .

**Issue 2)** When a coefficient  $\hat{c}_i$  gets small, we need to use many of the  $f$  bits to represent the leading 0's, losing significant bits, and even causing an underflow if  $|\hat{c}_i| < 2^{-f}$ . However,  $\hat{x}^i$  may be still large and we need an accurate  $\hat{c}_i$  for  $\hat{c}_i \hat{x}^i$  to approximate  $c_i x^i$  in the continuous space. In fact, we observe that  $c_i$  tends to be small when  $i$  is close to  $k$ . Intuitively, there is a relationship between the smoothness of target functions and their polynomial approximations. The smoother the target function is, the faster its polynomial

**Algorithm 3: ConstrainK**


---

```

1 Function ConstrainK(domain  $[a, b]$ , FXP format  $\langle n, f \rangle$ ):
2    $|\hat{x}|_{\max} \leftarrow \max(|a|, |b|)$  and  $|\hat{x}|_{\min} \leftarrow \min(|a|, |b|)$ ;
3    $k_O \leftarrow k$  if  $(|\hat{x}|_{\max} < 1)$  else  $\frac{n-f-1}{\log_2(|\hat{x}|_{\max})}$ ;
4   If  $a \cdot b < 0$  then  $k_U \leftarrow 3$ ;
5   Else  $k_U \leftarrow k$  if  $(|\hat{x}|_{\min} > 1)$  else  $\frac{f}{-\log_2(|\hat{x}|_{\min})}$ ;
   Return: Maximum feasible  $\bar{k} \leftarrow \min(k, k_O, k_U)$ 

```

---

**Algorithm 4: FLPsimFXP :  $x \rightarrow \hat{x}$** 


---

```

Input :  $x \in \mathbb{R}$  and FXP format  $\langle n, f \rangle$ .
Return:  $\hat{x}$ 
1 If  $(|x| > 2^{n-f-1})$ : Return:  $2^{n-f-1}$ ;
2 If  $(|x| < 2^{-f})$ : Return: 0;
3 Return:  $\text{round}_2(x, f)$ 

```

---

approximation converges (coefficient  $|c_n| \rightarrow 0$  with  $n \rightarrow \infty$ ). Theoretically, [38] shows that for Chebyshev polynomials, the absolute value of  $k$ th-order coefficient  $|c_k|$  is inversely proportional to the exponent of its order  $k$ , presenting a quick descending rate.

**Issue 3)** Converting all parameters into FXP involves many roundings to evaluate the polynomial (rounding the fractional parts beyond the  $f$  bits to 0, in both computing  $x$  to the  $k$ th power and adding-up all terms), hurting the precision.

**Our solution.** Algorithm 2 outlines our solution. The algorithm firstly uses *Chebyshev interpolation* or *Lagrange interpolation* to find the optimal polynomial in the continuous space (represented by double-precision FLPs) and transfer the polynomial to discrete space (represented by FXPs) while avoiding the above issues.

**Step 1: Constraining  $k$  to avoid overflow (Issue 1).** First, to avoid  $\hat{x}^k$  overflow or underflow (issue 1), we find the max feasible  $\bar{k} \leq k$ , guaranteeing that  $\hat{x}^{\bar{k}}$  does not overflow or underflow  $\forall \hat{x} \in [a, b]$  (Line 1 in Algorithm 2). In ConstrainK (Algorithm 3), overflow is easy to constrain, as we only need  $(|\hat{x}|_{\max})^{k_O} \leq 2^{n-f-1}$ , or  $k_O \leq \frac{n-f-1}{\log_2 |\hat{x}|_{\max}}$  if  $|\hat{x}|_{\max} > 1$  (Line 3). Underflow is more involved, as if  $0 \in [a, b]$ ,  $|\hat{x}|$  can be arbitrarily close to zero. In such case, we heuristically limit  $k_U$  to 3 (Line 4). Otherwise, we need  $(|\hat{x}|_{\min})^{k_U} \geq 2^{-f}$ , or  $k_U \leq \frac{f}{(-\log_2 |\hat{x}|_{\min})}$  if  $|\hat{x}|_{\min} < 1$  (Line 5). The max feasible  $\bar{k}$  is the smallest number among  $\{k, k_O \text{ and } k_U\}$ .

**Step 2: Fitting a polynomial in FLP.** We use *Chebyshev interpolation* [38] that interpolates on  $\bar{k} + 1$  Chebyshev roots to construct the Chebyshev polynomial (Line 4), which is a close approximation to the real *minimax polynomial*, i.e., minimizing the max approximation error. The method is widely adopted in practice (plaintext) [3, 38]. There is a corner case when  $[a, b]$  is so small that  $\frac{b-a}{2^{-f}} \leq \bar{k} + 1$ , i.e., the domain in such case does not contain enough points to fit the  $\bar{k}$ th-order polynomial. We construct the  $(\frac{b-a}{2^{-f}} - 1)$ th-order polynomial trying to cover all the discrete points. Specifically, we use *Lagrange interpolation* [30] to solve the polynomial using all discrete points (Line 7).

Note that both Chebyshev and Lagrange methods fit the polynomial in continuous space using FLP. Then we need to convert them back into the FXP space. Issue 2-3 arise on this conversion.

**Algorithm 5: ScalePoly :  $p_k \rightarrow \hat{p}_k$** 


---

```

Input :  $p_k = \sum_{i=0}^{k-1} (c_i \cdot x^i)$  in continuous space, domain  $[a, b]$ .
Return:  $\hat{p}_k = \sum_{i=0}^{k-1} (\hat{c}_i \cdot x^i \cdot \hat{s}_i)$  in discrete space.
1 Character  $\hat{x} \leftarrow \max(|a|, |b|)$ , most likely to overflow;
2 for  $i \leftarrow 0$  to  $k$  do
3    $\hat{c}_i, \hat{s}_i \leftarrow \text{ScaleC}(c_i, \langle n, f \rangle, i, \hat{x})$ ;
4 end
5 Return  $\hat{p}_k$ .
6 Function ScaleC( $c, \langle n, f \rangle$ , order  $k$  and  $\hat{x}$ ):
7    $\hat{s}_{CUF} \leftarrow 2^{-f}$ ;
8    $\hat{s}_{COF} \leftarrow \text{FLPsimFXP}(\frac{c\hat{x}^k}{2^{n-f-1}}, n, f)$ ;
9    $\hat{s} \leftarrow \min\{\max(\hat{s}_{CUF}, \hat{s}_{COF}), 1\}$ ;
10   $\hat{c} \leftarrow \text{FLPsimFXP}(\frac{c}{\hat{s}}, n, f)$ ;
11  Return  $\hat{c}, \hat{s}$ ;

```

---

**Step 3: Converting to FXP space with a scaling factor to enlarge the representation range (Issue 2).** As we mentioned, when we round the fitted  $c_i$ 's to  $\hat{c}_i$ 's in FXP,  $\hat{c}_i$  may be too small to represent precisely. As  $\hat{c}_i$ 's are in plaintext, we can use the typical scaling factors to enlarge its representation range. We translate each FLP  $c_i$  into two FXP numbers,  $(\hat{c}_i, \hat{s}_i)$ , letting  $\hat{c}_i \geq c_i$  to be large to preserve sufficient significant bits and  $\hat{s}_i \leq 1$  is a scaling factor such that  $c_i \approx \hat{c}_i \hat{s}_i$ . We want to let  $\hat{c}_i$  contain more significant bits to maintain precision, but we need to avoid a too large  $\hat{c}_i$  that causes  $\hat{c}_i \hat{x}^k$  to overflow, especially when  $\hat{x}^k$  is large. More precisely, we require: 1)  $\hat{c}_i \hat{x}^k \leq 2^{n-f-1}$ , i.e., it does not overflow; 2)  $\hat{s}_i$  itself is a valid FXP, and 3)  $0 < \hat{s}_i \leq 1$ , i.e., it indeed scales up the coefficient, not making it even smaller. Algorithm 5 converts all FLP coefficients  $c_i$  into  $(\hat{c}_i, \hat{s}_i)$  FXP pairs, satisfying all the three requirements.

**Step 4: Further reducing the rounding precision loss using residual boosting (Issue 2 - 3).** After step 3, we get an FXP approximation  $\hat{p}(\hat{x}) = \sum_{i=0}^k (\hat{c}_i \hat{x}^i \hat{s}_i)$ . The rounding errors in  $\hat{c}_i$  and  $\hat{s}_i$  exacerbate the difference between the approximation and the real  $F(x)$ , and we use a *residual function*  $R(x) = F(x) - \hat{p}_k(x)$  in FLP to capture the difference. If we can estimate  $R(x)$  with another discrete polynomial  $\hat{r}_{k'}(x)$  with  $k' < k$ , we may get a better precision if we use  $\hat{p}_k(x) + \hat{r}_{k'}(x)$  to approximate  $F(x)$ .

We observe that we may approximate  $R(x)$  using a series of *lower-order* FXP polynomials because the lower-order coefficients tend to be larger and preserve more significant bits. Algorithm 6 illustrates the residual boosting procedure. In a nutshell, the algorithm iterates  $k'$  through  $(k-1)$  to 1, trying to fit a  $k'$ th-order polynomial  $\hat{r}_{k'}$  approximating  $R(x)$  using the same Chebyshev interpolation as in Algorithm 2 and ScalePoly Algorithm 5. We add  $\hat{r}_{k'}$  to  $\hat{p}_k$  through function Boost in Algorithm 6 if we are able to obtain smaller max error on sample set  $\hat{X}$ . The residual boosting algorithm is best-effort and opportunistic, but empirically, it performs well (Section 7.5).

**Step 5: Checking if the polynomial is actually feasible and returning it if so.** As the last step, we get *sample set*  $\hat{X}$  from  $[a, b]$  in FXP and check the accuracy of  $\hat{p}_k(\hat{x})$  obtained from the previous steps by computing  $|\hat{p}_k(\hat{x}) - F(\hat{x})|_d, \forall \hat{x} \in \hat{X}$  and find the max error. If the check passes (max SRD less than  $\epsilon$ ), we return the



**Algorithm 6: ResidualBoosting**


---

**Input** :  $\hat{p}_k$  in discrete space, target  $F(x)$  and domain  $[a, b]$ .  
**Return** :  $\hat{p}_k^*$  in discrete space.

```

1  $\hat{p}_k^* \leftarrow \hat{p}_k, R \leftarrow F - \hat{p}_k^*$ ;
2 Sample number  $N_s \leftarrow \min(\text{Max samples } MS, \text{All points } \frac{b-a}{2^{-f}})$ ;
3  $\hat{X} \leftarrow \text{FLPsimFXP}(\text{Linspace}([a, b], N_s))$ ;
4 for  $k' \leftarrow k - 1$  to 0 do
5    $r_{k'} \leftarrow \text{Cheby-Interpolation}(R, [a, b], k')$ ;
6    $\hat{p}_k^{tmp} \leftarrow \text{Boost}(\hat{p}_k^*, r_{k'}, [a, b])$ ;
7   /* Boost when benefit exist. */
8   if  $(\max_{\hat{x} \in \hat{X}} |F(\hat{x}) - \hat{p}_k^{tmp}(\hat{x})|_d < \max_{\hat{x} \in \hat{X}} |F(\hat{x}) - \hat{p}_k^*(\hat{x})|_d)$  then
9      $\hat{p}_k^* = \hat{p}_k^{tmp}, R = F - \hat{p}_k^*$ ;
10  end
11 Return :  $\hat{p}_k^*$ 

```

---

**Function**  $\text{Boost}(p_k, r_{k'}, k \geq k'$  with domain  $[a, b])$ :

```

12  $p_{k'}(x) = \sum_{i=0}^{k'} (\hat{c}_i^{(\hat{p}_k)} \cdot \hat{s}_i^{(\hat{p}_k)} + c_i^{(r_{k'})}) \cdot x^i$ ;
13  $\hat{p}_{k'} \leftarrow \text{ScalePoly}(p_{k'}, [a, b])$ ;
14  $\hat{p}_k(x) = \sum_{i=0}^{k'} (\hat{c}_i^{(\hat{p}_{k'})} \cdot x^i \cdot \hat{s}_i^{(\hat{p}_{k'})}) + \sum_{i=k'+1}^k (\hat{c}_i^{(\hat{p}_k)} \cdot x^i \cdot \hat{s}_i^{(\hat{p}_k)})$ ;
15 Return :  $\hat{p}_k$ 

```

---

**Algorithm 7: OPPE Algorithm (OPPE)**


---

**Config** : Three parts plaintext parameters of  $\hat{p}_k^m$ :  $\hat{W}$  (without endpoint  $\hat{w}_m$ ),  $C = \{\hat{c}_{j,i}\}$  and  $S = \{\hat{s}_{j,i}\}$ .  
**Input** : Secret input  $[\hat{x}]$ .  
**Return** : The secret evaluation result of  $[\hat{p}_k^m(\hat{x})]$ .

```

1  $[\text{comp}] \leftarrow \text{GT}([\hat{x}], \hat{W})$  # compare  $x$  with each break point.;
2  $[\text{mask}] \leftarrow \text{ADD}([\text{comp}], \text{-leftshift}([\text{comp}], 1))$ ;
3 for  $i \leftarrow 0$  to  $k - 1$  do
4    $[\text{coeff}]_i \leftarrow \sum_{j=0}^{m-1} \text{MUL}([\text{mask}]_j, \hat{c}_{j,i})$ ;
5    $[\text{scaler}]_i \leftarrow \sum_{j=0}^{m-1} \text{MUL}([\text{mask}]_j, \hat{s}_{j,i})$ ;
6 end
7  $[\text{xterm}] \leftarrow \text{CalculateKx}([\hat{x}], k)$ ;
8 Return  $\sum_{i=0}^{k-1} (\text{MUL}(\text{MUL}([\text{coeff}]_i, [\text{xterm}]_i), [\text{scaler}]_i)$ ;
9 Function  $\text{CalculateKx}([\hat{x}], k)$ :
10   /* Calculate  $[1, [\hat{x}], [\hat{x}]^2, \dots, [\hat{x}]^k]$  */
11   shift  $\leftarrow 1$ ,  $[\text{res}] \leftarrow [1, \text{tail}([\hat{x}], k)]$  /* repeat  $[\hat{x}]$   $k$  times */
12   while shift  $< k$  do
13      $[\text{res}]_{\text{shift}} = \text{MUL}([\text{res}]_{\text{shift}}, [\text{res}]_{-\text{shift}})$ ;
14     shift  $\times = 2$ ;
15   end
16   Return  $[\text{res}]$ ;

```

---

polynomial to Algorithm 1, otherwise the function returns a Null, causing Algorithm 1 to recurse on smaller ranges.

**4.3 The Runtime Evaluation Algorithm OPPE**

At runtime, we take the output of Algorithm 1,  $\hat{p}_k^m$ , as plaintext config, and take the secret-shared value  $[\hat{x}]$  as ciphertext input ( $[\hat{x}]$  indicate the secret shares of value  $\hat{x}$  among each party), to compute the result  $[\hat{p}_k^m(\hat{x})]$  in the *oblivious piece-wise polynomial evaluation (OPPE)* Algorithm 7. The piecewise polynomial  $\hat{p}_k^m$  is described

by three parameters:  $W = [a = \hat{w}_0, \hat{w}_1, \hat{w}_2, \dots, \hat{w}_m = b]$  (without endpoint  $\hat{w}_m$ ) are the boundaries for the  $m$  pieces, the coefficients  $\hat{c}_{j,i}$  and scaling factors  $\hat{s}_{j,i}$  for all  $j = 0 \dots m - 1$  and  $i = 0 \dots k$ . All plaintext and ciphertext inputs are FXP numbers. We use ADD, MUL and GT in Algorithm 7 to denote the subroutines evaluating *secure* addition, multiplication and greater-than, respectively ( $\sum$  means summation through ADD).

**OPPE Design.** OPPE Algorithm 7 treats each subroutine as an *arithmetic black box* and organizes them *obliviously*, i.e., the execution path is independent of the inputs. OPPE first determines which piece  $[\hat{x}]$  belongs to, using one vectorized GT and one ADD (Lines 1-2). The comparison result is a ciphertext vector  $[\text{mask}]$ , containing a single *one*, and all other elements are *zeros*. Line 3-6 select the coefficients and scaling factors obliviously, using the  $[\text{mask}]$ . Line 7 computes all  $[\hat{x}]^i, \forall i = 0 \dots k$  using  $(\lfloor \log k \rfloor + 1)$  vectorized MUL's (each on two ciphertext vectors with size  $< k$ ) using the subroutine CalculateKx. Line 8 computes every term using two MUL's:  $\text{MUL}(\text{MUL}(\hat{c}_{j,i}, \hat{x}^i), \hat{s}_{j,i})$ , and adds up the products to compute the result. Note that we must execute the two MULs in this specific order to take advantage of the scaling factor.

**Complexity.** Algorithm 7 uses  $(2km)$  plaintext-with-ciphertext MUL's,  $m$  GT's and  $O(k \log k)$  MUL's. We can also leverage the vector (a.k.a., SIMD) optimization in many MPC platforms. If so, we only need  $(\lfloor \log k \rfloor + 1)$  rounds of ciphertext MUL, 2 rounds of plaintext-with-ciphertext MUL, and 1 round of GT.

**Independent operations and parallelism.** We observe that the  $[\text{mask}]$  computation and coefficient selection step (Line 1-6) is independent of the CalculateKx routine (Line 7). Thus if an MPC platform supports concurrency, we can run both independently, further reducing the running time. Also, when the input vector  $\hat{x}$  is long, we automatically break it up into multiple pieces to utilize the underlying platform's threading support to evaluate each piece.

**5 SECURITY ANALYSIS**

**Security definitions.** NFGen uses the same security definitions as the *secure multi-party protocol* of underlying MPC platform,  $\pi_f$  that aims to let  $n$  parties evaluate function  $f$  without a trusted third party. The security is defined as the *security properties* achieved in the presence of some adversary  $\mathcal{A}$  who can control a set of at most  $t$  corrupted parties according to some *adversarial model*. Different protocols have their own choices of both the adversarial model and security properties to achieve, usually for trade-off of performance.

The common *adversarial model* trade-offs include 4 dimensions [26]: 1) *corruption strategy*: adaptive vs. non-adaptive; 2) *corruption proportion*: dishonest vs. honest-majority; 3) *behavior*: malicious vs. semi-honest; 4) *power*: informational vs. computational-secure.

Under these assumptions, protocols usually achieve the following two essential security properties: *privacy* and *correctness*. Optionally, there are other security properties a protocol may consider [28]. E.g., *fairness*, i.e., if one party receives the result, all parties receive it; and *guaranteed delivery*: whether the joint parties can always receive the results.

**Security assumptions of NFGen.** NFGen builds on top of general-purpose MPC platforms with each party carrying out the computation connected through *secure channels*. NFGen assumes that three

secure subroutines ADD, MUL and GT evaluating secret addition, multiplication and greater-than are provided and all the inputs are secret-shared among computation parties before the evaluation of the generated protocol. These assumptions are easily to meet as various implementations of secret addition, multiplication and greater-than have been proposed and some of them are widely adopted (e.g., [4, 16]). Specifically, NFGen introduces no different assumptions about the *adversary model* for the underlying protocols that implement the three required subroutines.

**Security analysis of NFGen.** We illustrate the security of NFGen's code generation approach by showing that it guarantees the *same security properties* as the three subroutines in the presence of the same adversary.

1) Obviously, the pre-computation steps (Algorithm 1 - 6) are independent of secret inputs and performed offline, thus cannot affect any security property;

2) To prove the privacy and security properties, we directly follow the *real-ideal* paradigm introduced in [9]. It defines security by requiring that the distribution of the protocol evaluation in the *real* world is indistinguishable from the *ideal* world with a trusted third party. Under this paradigm, we show that NFGen generates the  $\hat{p}_k^m$  evaluation protocol by *composing* the three subroutines as so-called arithmetic black boxes in the standard *modular composition* way [9] without revealing any information nor introducing any interaction. Thus, NFGen naturally inherits the same security property of the subroutines from the *Canetti's composition theorem* [9]. The security preserving property is the direct result of the composition theorem.

3) We show that NFGen provides the same optional security properties as the underlying protocols. Using *guaranteed delivery* as an example, if the provided subroutines offer this property, meaning that there is no *abort* within these subroutines, the  $\hat{p}_k^m$  evaluation protocol (Algorithm 7) will not abort either, as there is no breakpoint in the routine. On the other hand, if the subroutines are *secure with abort*, Algorithm 7 does not try to handle these abortions at all and lets the protocol abort.

## 6 IMPLEMENTATION

In this section, we briefly introduce how we integrate the fitted  $\hat{\mathcal{P}}$  into the target MPC system.

### 6.1 Profiler and Performance Prediction

In order to select the best  $\hat{p}_k^m \in \hat{\mathcal{P}}$ , we need to model the performance of specific deployed MPC systems. Such modeling is not straightforward as the performance not only depends on the MPC protocols but also on the implementation and deployment. We use a profiler that runs on the target system to automatically build the performance prediction model.

The execution time of Algorithm 7 depends on  $(k, m)$  only, making the prediction possible. We use the profiler to measure the evaluation time  $t$  of 2,000 configurations of piecewise polynomial samples with different  $(k, m)$  combinations ( $k \in [3, 10]$  and  $m \in [2, 50]$ ). Then we fit a *multivariate polynomial regression model* [24] on these samples. Note that the performance model is independent of  $F(x)$ , and thus we only need to profile once per system.

Some MPC systems provide built-in functions that are highly optimized for their settings, such as  $\frac{1}{x}$ ,  $e^x$ , or  $\ln x$ . For example, MP-SPDZ provides a very efficient  $\frac{1}{x}$  [11]. If users list these functions in the NFD, the profiler also measures the performance of such functions. If  $F(x)$  is simple, it may be better off taking the direct evaluation approach. Direct evaluation is only viable with all the three conditions: 1)  $F(x)$  does *not* contain  $e^x$  as an intermediate step, as it is highly likely to overflow; 2)  $F(x)$  contains less than three steps with non-linear functions to avoid unpredictable error accumulations and 3) the predicted running time of direct evaluation is shorter than all  $\hat{p}_k^m \in \hat{\mathcal{P}}$ . We rarely find suitable cases to use direct evaluation, but on functions like *isru*, which is  $1/(1 + |x|)$ , effectively just a single  $\frac{1}{x}$ , direct evaluation is 1.6× faster than the best  $\hat{p}_k^m$  (Section 7.2).

We can either run the profiler in pre-computation or run it *just-in-time* right before running a large MPC task. In this paper, we do all profiling and plan selections in the pre-computation.

### 6.2 OPPE Code Generation

Different MPC platforms offer not only different high-level languages, but also different support for vector operations and multi-threading. To best utilize these platform-specific optimizations while still remain portable, we use a template-based code generation approach to implement OPPE (Algorithm 7).

NFGen provides MPC-platform-specific code templates implementing OPPE. Each template is highly optimized for a specific platform. For example, the code template uses multi-threading in PrivPy to compute all independent segments concurrently [25], and leverages the *probability truncation* optimizations in MP-SPDZ [14]. Note that we only need to customize the OPPE template for each platform, and all the other procedures in NFGen are reusable across platforms. Currently, we support both PrivPy and MP-SPDZ.

Using the templates, it is straightforward to generate  $\hat{p}_k^m$  evaluation code, as we only need to insert  $\hat{p}_k^m$  parameters in to the code template as literals. The generated code runs the same as normal functions in the target MPC system.

We pass  $\hat{p}_k^m$  as literals in the generated code instead of arguments to OPPE function, because compilers in platforms like MP-SPDZ significantly improves performance if all input lengths (in our case,  $m$  and  $k$ ) are statically known.

## 7 EVALUATION

In our evaluation, we show that NFGen is able to 1) offer better performance and lower communication costs across algorithms, protocols and systems; 2) avoid the overflow/underflow errors in the traditional approaches, and provide better accuracy in complex functions; 3) calculate sophisticated non-linear functions otherwise requiring extensive calculus knowledge to implement; 4) support a large domain with reasonable accuracy even with a very limited number of bits; and 5) benefit real applications with both performance and accuracy improvements. We also evaluate the different design choices in NFGen, such as the effectiveness of profiling, scaling and residual boosting.

### 7.1 Experiment Setup



**Table 1: MPC Settings with Varied Operation Performance**

Sec model	No.	MPC sys $S$	$\times$ (ms)	$\times$ : $>$ : $\frac{1}{x}$ : $\sqrt{x}$ : $\ln x$ : $e^x$
Semi-honest	A	<i>PrivPy Rep2k</i>	1	1 : 11 : 67 : 55 : 118 : 44
	B	<i>Rep2k</i>	2	1 : 4 : 31 : 75 : 68 : 107
	C	<i>RepF</i>	32	1 : 1 : 11 : 28 : 26 : 47
	D	<i>Shamir</i>	81	1 : 1 : 8 : 16 : 15 : 29
Malicious	E	<i>Ps-Rep2k</i>	851	1 : 1 : 16 : 35 : 26 : 97
	F	<i>Ps-RepF</i>	84	1 : 2 : 24 : 56 : 44 : 137

B-E use  $\langle 96, 48 \rangle$ -FXP and A uses  $\langle 128, 48 \rangle$ . Column 3 is the absolute time (in ms) to compute  $\times$  on 100-dimensional vector and Column 4 is the relative performance to  $\times$ .

**MPC platforms.** We evaluate NFGen on two MPC platforms, MP-SPDZ [21] that implements over 30 secret sharing protocols (and we choose 5 for evaluation), and PrivPy [25] that only supports a single protocol. MP-SPDZ first compiles the high-level code into bytecode to execute with underlying protocols, while PrivPy executes programs by interpreting Python code at runtime. Both platforms support FXP number with different width.

**Secret sharing protocols.** We adopt six different secret-sharing protocols, covering different security assumptions over adversarial behaviors (*semi-honest* or *malicious*); *computation domains* (over a ring of  $\mathbb{Z}_{2k}$  or finite field  $\mathbb{F}_p$  by modulo a prime  $p$ ) and *sharing methods* (using replicated secret sharing or *shamir* secret sharing). We briefly introduce each protocol in the following.

First we introduce the four *semi-honest* protocols: A. *PrivPy-Rep2k* is an 2-out-of-4 replicated secret sharing protocol, proposed by Li *et al.* [25]. It splits each value  $x$  into four shares over a ring of  $\mathbb{Z}_{2k}$  and let each party  $P_i$  ( $i = 1, 2, 3, 4$ ) holds two shares, satisfying that any two parties can reconstruct  $x$  while each one sees two random integers. B. *Rep2k* and C. *RepF* two protocols split a value  $x$  into three shares, satisfying that  $x \equiv x_1 + x_2 + x_3 \pmod{M}$  and let each party  $P_i$  ( $i = 1, 2, 3$ ) holds  $(x_i, x_{i+1})$  (indexes wrap around 3).  $M = 2^k$  for *Rep2k* protocol and  $M = p$  for *RepF* protocol where  $p$  is a prime. D. *Shamir* shares a value  $x \in \mathbb{F}$  through a random chosen 2-degree polynomial  $f_s$ , such that  $f_s(0) = x$ . Each party  $P_i$ , ( $i = 1, 2, 3$ ) holds a distinct point over polynomial  $f_s$ . They together can reconstruct  $f_s$  and obtain  $x = f_s(0)$  while any set less than three parties contain no information about  $x$ .

Then we introduce the two *malicious* protocols: E. *Ps-Rep2k* and F. *Ps-RepF*: *Ps* refers to *Post-Sacrifice* strategy proposed by Lindell *et al.* [26]. It compiles a semi-honest protocol into a malicious secure version by adding a verification step. The verification step let the honest parties detect cheating behavior with high probability. The initial work [26] only considers finite field (*Ps-RepF*) and the follow-up work [13] extends it to ring (*Ps-Rep2k*).

We choose these six protocols not only because they are common practical protocols with different assumptions, but also because they exhibit various performance characteristics on the basic operations, which is helpful to show the generality of NFGen. Table 1 summarizes the selected MPC settings ( $S$ ) with the absolute performance of  $\times$  and the performance of the other basic operations relative to  $\times$ .

**Evaluation environment.** We perform all the evaluations on a cluster of four servers with two 20-core 2 GHz Intel Xeon CPUs and 180 GB RAM each, connected through 10 Gbps Ethernet. MP-SPDZ uses only three servers while PrivPy uses all four.

## 7.2 Performance

We use 15 widely-used non-linear functions for performance evaluation, including 8 activation functions used in deep learning and 7 probability distribution functions. The input domain of each function is set to the interval without a close-to-zero derivative, as these intervals are hard to approximate while others can be simply approximated with some constants. We run these functions on all six protocols in the two MPC platforms, and compare the performance with *direct evaluation* as the baseline (except *sigmoids* which is the built-in functions in both platforms). In all cases, we set the accuracy requirement to  $\epsilon = 10^{-3}$  and  $\hat{0} = 10^{-6}$  (defined in Eq. 1), and run the experiments on 10,000 evenly spaced  $\hat{x}$  samples. We measure the computation time on 100-dimensional vectors. Table 2 shows five examples and the full evaluation is listed in [40]. We have the following important observations:

**Performance.** 1) NFGen achieves significant performance gain in 93% of these cases, with an average speedup of 6.5 $\times$  and a max speedup of 86.1 $\times$  ( $Bs\_dis$  on *Rep2k*).

2) NFGen significantly reduces communication, with an average reduction of 39.3% and a max of 93%, but it is not proportional to the speedups. This is because the OPPE algorithm uses a fixed number of communication rounds, and each round involves vectorized comparisons and multiplications (depending on  $k$  and  $m$ ), while the baseline evaluates the function step-by-step, and thus may involve more rounds, leading to longer computation time.

3) The more complicated a function is, the more likely NFGen achieves a better speedup, for the same reason above.

4) Smaller  $m$  values perform better for most functions. As both MP-SPDZ and PrivPy support vectorized  $>$  operations, the latency is largely dependent of  $m$ , even on the 100-dimensional vectors.

5) In *malicious* protocols (settings E and F), NFGen is more likely to achieve smaller speedup or even slowdowns on a few functions like *sigmoid*. There are two reasons: a) the  $\times$  and  $>$  are very slow, and the intensive  $\times$ 's introduce more *validation checks* (the *Ps*-protocols use extra *multiplication triplets* to detect cheating); b) the validation checks prevent efficient batch (vector) operations in MP-SPDZ, resulting in less efficient OPPE execution.

**Effectiveness of the profiler.** 1) For the *soft\_sign* function, on all protocols in MP-SPDZ, NFGen automatically falls back to direct evaluation (Section 6.1), while on PrivPy, it uses  $\hat{p}_8^8$  polynomial and achieves a 6.1 $\times$  speedup. This is because the function essentially computes an absolute value (equivalent to a  $>$ , plus a  $\frac{1}{x}$ ). Both  $>$  and  $\frac{1}{x}$  are significantly slower than  $\times$  in PrivPy, but it is not the case for MP-SPDZ, as Table 7.1 shows. Thus, based on the profiler results, NFGen chooses different evaluation strategies. In fact, we manually test the polynomial approach for MP-SPDZ, and it is 1.6 $\times$  slower than direct evaluation (with the most efficient polynomial ( $\hat{p}_6^{10}$ )), showing the effectiveness of the profiler.

2) The profiler can also select different  $(k, m)$  settings for different protocols in MP-SPDZ. For example, it chooses different  $(k, m)$ s for different protocols in computing *sigmoid*. Specifically, it uses a larger  $k$  for *Rep2k* (more  $\times$  and fewer  $>$ ), as  $>$  is 4 $\times$  slower than  $\times$  in *Rep2k*, according to Table 7.1.

3) For *RepF*, *Ps-Rep2k*, *Ps-RepF* and *Shamir*, the profiler tends to choose larger  $m$  for a smaller  $k$ , because the relative performance

**Table 2: Examples in Performance Evaluations**

$F(x)$	$S$	$\checkmark$	$(k, m)$	$T_{\text{Fit}}$	Communication (MB)			Computation time (ms)		
					Base	NFGen	Save	Base	NFGen	SpeedUp
$\text{sigmoid}(x) = \frac{1}{1+e^{-x}}$ $x \in [-50, +50], F(x) \in [0.0, 1.0]$ Non-linear building-blocks: 2	A	$\times$	(10, 8)	4.3	618	263	<b>60%</b>	147	23	<b>6.3×</b>
	B	$\checkmark$	(7, 10)	3.5	1	1	-5%	137	124	<b>1.1×</b>
	C	$\checkmark$	(5, 14)	3.5	4	4	-5%	1155	802	<b>1.4×</b>
	D	$\checkmark$	(5, 14)	3.5	18	19	-8%	1863	1525	<b>1.2×</b>
	E	$\checkmark$	(5, 14)	3.5	212	308	-45%	75949	106857	<b>0.7×</b>
	F	$\checkmark$	(5, 14)	3.5	207	234	-13%	9732	11224	<b>0.9×</b>
$\text{tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ $x \in [-50, +50], F(x) \in [-1.0, 1.0]$ Non-linear building-blocks: 3	A	$\times$	(9, 8)	4.5	1876	216	<b>90%</b>	335	21	<b>15.7×</b>
	B	$\times$	(5, 9)	3.2	13	1	<b>92%</b>	800	80	<b>10.0×</b>
	C	$\times$	(5, 9)	3.2	19	3	<b>83%</b>	5901	597	<b>9.9×</b>
	D	$\times$	(5, 9)	3.2	64	14	<b>78%</b>	8882	1115	<b>8.0×</b>
	E	$\times$	(5, 9)	3.2	996	197	<b>80%</b>	337530	68550	<b>4.9×</b>
	F	$\times$	(5, 9)	3.2	966	150	<b>84%</b>	45486	7309	<b>6.2×</b>
$\text{soft\_sign}(x) = \frac{x}{1+ x }$ $x \in [-50, 50], F(x) \in [-1.0, 1.0]$ Non-linear building-blocks: 2	A	$\times$	(8, 8)	1.9	518	231	<b>60%</b>	131	21	<b>6.1×</b>
	B	$\checkmark$	NA	1.3	1	1	0%	79	78	<b>1.0×</b>
	C	$\checkmark$	NA	1.3	2	2	0%	451	437	<b>1.0×</b>
	D	$\checkmark$	NA	1.3	8	8	0%	741	753	<b>1.0×</b>
	E	$\checkmark$	NA	1.3	52	52	0%	15507	15520	<b>1.0×</b>
	F	$\checkmark$	NA	1.3	49	49	0%	2315	2373	<b>1.0×</b>
$\text{Normal\_dis}(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ $x \in [-10, +10], F(x) \in [0.0, 0.4]$ Non-linear building-blocks: 1	A	$\times$	(8, 12)	5.2	420	295	<b>30%</b>	67	24	<b>2.8×</b>
	B	$\times$	(8, 12)	3.6	3	2	<b>45%</b>	4906	156	<b>31.5×</b>
	C	$\times$	(8, 12)	3.6	7	5	<b>27%</b>	5029	970	<b>5.2×</b>
	D	$\times$	(8, 12)	3.6	24	23	<b>5%</b>	6588	1846	<b>3.6×</b>
	E	$\times$	(5, 22)	3.6	257	481	-87%	89740	166328	<b>0.5×</b>
	F	$\times$	(8, 12)	3.6	249	301	-21%	14908	14861	<b>1.0×</b>
$Bs\_dis(x) [5] = \left( \frac{\sqrt{x} + \sqrt{\frac{1}{x}}}{2\gamma x} \right) \phi \left( \frac{\sqrt{x} - \sqrt{\frac{1}{x}}}{\gamma} \right)$ $\gamma = 0.5, x \in [10^{-6}, 30], F(x) \in [0.0, 0.2]$ Non-linear building-blocks: 3	A	$\times$	(10, 8)	4.0	2815	263	<b>90%</b>	630	22	<b>29.1×</b>
	B	$\times$	(7, 11)	3.2	13	1	<b>89%</b>	11463	133	<b>86.1×</b>
	C	$\times$	(5, 16)	3.2	23	5	<b>79%</b>	14631	915	<b>16.0×</b>
	D	$\times$	(5, 16)	3.2	65	22	<b>66%</b>	19167	1763	<b>10.9×</b>
	E	$\times$	(5, 16)	3.2	741	352	<b>53%</b>	239549	122325	<b>2.0×</b>
	F	$\times$	(5, 16)	3.2	718	268	<b>63%</b>	42157	13136	<b>3.2×</b>

\*  $T_{\text{Fit}}$  is the time for  $\hat{p}_k^m$  fitting in seconds.  $\checkmark$  indicates whether baseline achieves the accuracy requirements.

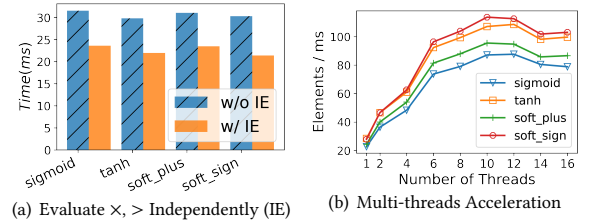
between  $\times$  and  $>$  is about 1 : 1 in these settings. Thus computing  $O(\log k)$  rounds of  $\times$ s is more expensive than a single round of vectorized  $>$ .

**Independence and parallelism.** We adopt two types of optimizations to accelerate the online phase performance: 1) Independent evaluation of  $\times$  and  $>$ : As we have discussed in Section 4.3, it is independent to perform the comparisons and to compute  $x$  to the  $k$ -th power. We evaluate the speedup on PrivPy by running them independently<sup>3</sup>. Figure 2(a) shows the performance results with/without independent evaluation. We can see the optimization provides a 1.3 – 1.4 $\times$  speedup even on small 100-dimensional input.

2) Concurrency on large inputs: Figure 2(b) shows the performance with varied numbers of threads on  $10^6$  input. We observe good speedup up to 10 threads (about 4 $\times$  speedup comparing with the single-thread) for all four functions we evaluate. Using more than 10 threads decreases performance as given the vector size, cost of threading overweights the speedup.

### 7.3 Accuracy

<sup>3</sup>MP-SPDZ does not support customized multi-threading in its user-level language so we do not adopt this optimization.



**Figure 2: Performance Optimizations**

**Accuracy against the baseline.** We compute the soft relative distance (SRD in Section 3) on each of the 10,000 input samples. We see a large variation of SRDs on different functions, and Figure 3 shows the *cumulative distribution functions (CDFs)* of the SRDs on each function for both the baseline and NFGen.

We observe the following: 1) 100% of the SRD of all 10k samples are under  $10^{-3}$  (the target  $\epsilon$ ), ranging between  $10^{-12}$  and  $10^{-3}$ , as expected. 2) In comparison, the baseline shows diverse errors across functions. The *Cauchy\_dis* is very accurate as it is just a  $\frac{1}{x}$  (NFGen falls back to direct evaluation in this case). However, in some functions, the baseline shows errors way exceeding the

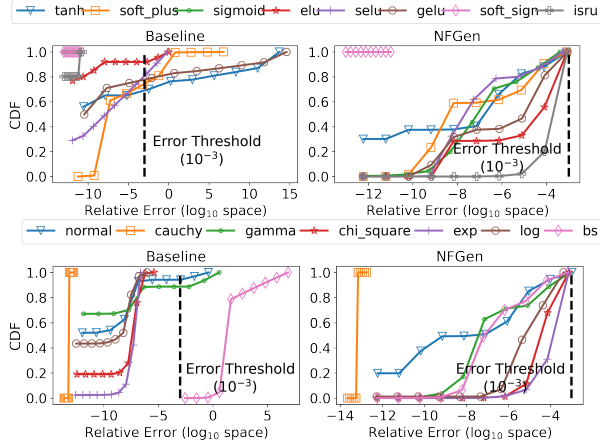


Figure 3: CDF of the Relative Errors

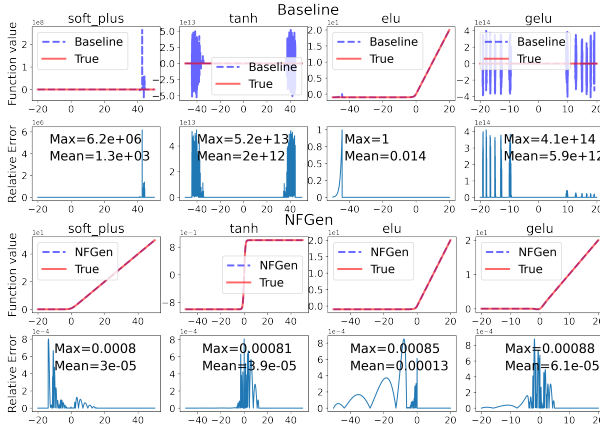


Figure 4: Overflow Error Examples

$10^{-3}$  limit (dashed vertical line in Figure 3) on over 10% samples (Function  $bs\_dis$  has 100% and  $tanh$  has more than 35%). There are two reasons: precision loss due to concatenating multiple non-linear functions and errors due to overflow/underflow. 3) Baseline has a larger percentage of samples with smaller errors (e.g.,  $< 10^{-9}$ ) than NFGen. This is as expected too: NFGen is based on a regression model to approximate, while baseline is a direct computation. However, we believe the predictable accuracy is more important than sometimes getting smaller errors.

**Errors due to overflows/underflows.** When the evaluation of  $F(x)$  uses an intermediate result in a large range, like  $e^x$ , it overflows given an  $x$  with  $|x| > \ln 2^{(96-48-1)} \approx 33.2$  in MP-SPDZ with  $\langle 96, 48 \rangle$ -FXP. The top two rows in Figure 4 plot SRD vs.  $x$  value for the baselines, and show the overflow cases, where the SRD can exceed  $10^5$ , even if the range of  $F(x)$  is perfectly representable. In comparison, the bottom two rows in Figure 4 show that SRDs are less than  $10^{-3}$  in NFGen. It also uncovers that NFGen has larger SRD when  $F(x)$  is close to zero, as the accuracy of polynomial approximation are constrained by the limited resolution of FXP.

**Error accumulation.** The error accumulation problem widely exists in scientific computing [3], even with double precision numbers.

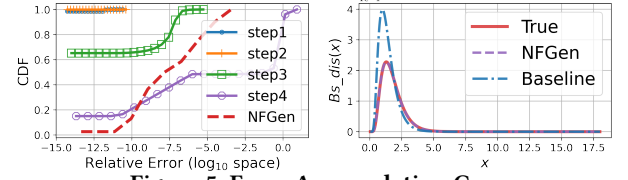


Figure 5: Error Accumulation Case

Table 3: Secret Sharing Reconstruction Loss

	Simulation		Secret Sharing	
	Max( $\times 10^{-4}$ )	Mean( $\times 10^{-5}$ )	Max( $\times 10^{-4}$ )	Mean( $\times 10^{-5}$ )
tanh	8.06	3.88	8.06	3.88
soft_plus*	7.67	2.94	8.00	2.98
sigmoid	3.98	1.87	3.98	1.87
elu	8.50	12.87	8.50	12.87
selu	5.24	1.94	5.24	1.94
gelu	8.79	6.11	8.79	6.11
isru	8.61	28.18	8.61	28.18
normal_dis*	3.40	1.01	10.49	2.53
gammas_dis	8.80	3.11	8.80	3.11
chi_square	7.27	10.70	7.50	10.70
exp_dis	7.39	19.76	7.39	19.77
log_dis	4.24	3.28	4.24	3.28
bs_dis*	5.96	1.98	6.91	1.97

Using FXP makes the problem worse, especially when the calculation of the target functions has many steps. As NFGen approximate the entire  $F(x)$  in one shot, it does not accumulate errors.

We take the  $bs\_dis$  function (Row 5 in Table 2), whose calculation has four steps, as an example. 1)  $x_{11} \leftarrow \sqrt{x}$ ;  $x_{12} \leftarrow \frac{1}{x}$ ; 2)  $x_2 \leftarrow \sqrt{x_{12}}$ ; 3)  $x_3 \leftarrow \phi(\frac{x_{11}-x_2}{y})$  and 4)  $x_4 \leftarrow (\frac{x_{11}+x_2}{2y*x}) * x_3$ . Using 10k  $\hat{x}$  samples, we compute the SRD after each step, and plot the CDF in Figure 5 (left). We can see that although the first two steps results in negligible SRD that is smaller than  $10^{-10}$ , more samples starts to show larger SRD after steps 3 and 4. After step 4, only less than half of the samples meet the accuracy requirement of  $10^{-3}$ . Figure 5 (right) shows the evaluation results, and we find obvious inaccuracies for  $\hat{x} \in [0.054, 18]$ , without overflow or underflow. In comparison, NFGen successfully limits the error below  $3.8 \times 10^{-4}$  using  $\hat{p}_5^{11}$ .

**Accuracy loss in secret sharing.** The secret sharing and reconstruction processes in MPC platforms may introduce extra inaccuracies as they approximate each real value in fixed-point shares [11, 33] (we call it *secret sharing reconstruction error (SSRE)*). For example, we observe that the value of exactly  $1 \times 10^{-14}$ , after secret sharing and reconstruction, may become  $1.06581 \times 10^{-14}$ . This SSRE also adds inaccuracy to the final evaluation result. To quantify the contribution of SSRE, we compare the SRD using real MPC to our simulated FXP (without SSRE). Table 3 shows the comparison (\* highlights cases with different results). We can see that SSRE does increase the inaccuracy, but the contribution is small comparing to other sources of inaccuracy.

**Accuracy with different FXP widths.** Many MPC platforms offer configurable  $\langle n, f \rangle$  for FXP numbers. While reducing  $n$  saves computation cost, a small  $n$  limits the input domain for both baseline methods and NFGen, because we need to represent all inputs, intermediate results and outputs with the number of bits. We conduct

**Table 4: Domain Restriction with Data Representation**

Config	$F(x)$	$(k, m)$	Origin $D$	Ours $D$
$\langle 128, 64 \rangle$	$\tanh$	(6, 11)	$[-44.4, 44.4]$	$(-10^{19}, 10^{19})$
$\hat{0} = 10^{-6}$	$\text{soft\_plus}$	(9, 9)	$[-44.4, 44.4]$	$(-10^{19}, 10^{19})$
$\epsilon = 10^{-4}$	$\text{Normal\_dis}$	(6, 48)	$[-9.4, 9.4]$	$(-10^{19}, 10^{19})$
	$\text{Bs\_dis}$	(10, 9)	$[0.0, 24.1]$	$(0, 10^{19})$
$\langle 96, 48 \rangle$	$\tanh$	(6, 11)	$[-33.3, 33.3]$	$(-10^{14}, 10^{14})$
$\hat{0} = 10^{-6}$	$\text{soft\_plus}$	(6, 11)	$[-33.3, 33.3]$	$(-10^{14}, 10^{14})$
$\epsilon = 10^{-3}$	$\text{Normal\_dis}$	(10, 9)	$[-8.2, 8.2]$	$(-10^{14}, 10^{14})$
	$\text{Bs\_dis}$	(5, 14)	$[0.1, 18.6]$	$(0, 10^{14})$
$\langle 64, 32 \rangle$	$\tanh$	(5, 9)	$[-22.2, 22.2]$	$(-10^9, 10^9)$
$\hat{0} = 10^{-5}$	$\text{soft\_plus}$	(7, 9)	$[-22.2, 22.2]$	$(-10^9, 10^9)$
$\epsilon = 10^{-3}$	$\text{Normal\_dis}$	(10, 9)	$[-6.7, 6.7]$	$(-10^9, 10^9)$
	$\text{Bs\_dis}$	(5, 14)	$[0.1, 13.0]$	$(0, 10^9)$
$\langle 32, 16 \rangle$	$\tanh$	(4, 6)	$[-11.1, 11.1]$	$(-10^4, 10^4)$
$\hat{0} = 10^{-2}$	$\text{soft\_plus}$	(4, 6)	$[-11.1, 11.1]$	$(-10^4, 10^4)$
$\epsilon = 5 \cdot 10^{-2}$	$\text{Normal\_dis}$	(4, 7)	$[-4.7, 4.7]$	$(-10^4, 10^4)$
	$\text{Bs\_dis}$	(5, 5)	$[0.1, 7.4]$	$(0, 10^4)$

\* The function range are actually defined by  $\hat{0}$ , we set outer range to default constant (e.g.,  $\tanh(\hat{x}) = -1$  for  $\forall \hat{x} \leq -18.79$ ).

experiments with  $n$  ranging from 32 to 128 and compare the supported  $\hat{x}$  domain between the baseline and NFGGen and summarize the results in Table 4.

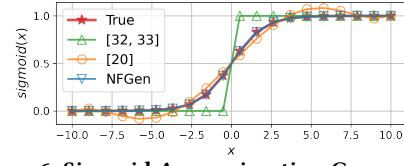
We observe: 1) As  $n$  gets smaller, the ranges shrink for both cases. NFGGen supports a much larger domain even for  $n = 32$ . This is because the baseline domain is severely limited by the range of intermediate results, e.g.  $\frac{1}{x}$  overflows when  $x$  gets close to zero and  $e^x$  overflows when  $x > \ln 2^{n-f-1}$ . 2) In contrast,  $n$  affects the domain in NFGGen in two different ways: a) it directly limits representable  $x$  to  $[2^{n-f-1}]$ ; b) it limits the representation of  $x^k$ , forcing us to use only small  $k$  values (Column 3). We see that NFGGen can automatically adapt to the  $n$  settings by reducing the  $k$  values to prevent overflows. 3) As expected, a small  $k$  limits the accuracy we can achieve. We empirically determine the minimal possible accuracy (both  $\hat{0}$  and  $\epsilon$ ) when we require  $m < 1000$  (Column 1), and find that even in the  $\langle 32, 16 \rangle$  setting (i.e. the max representable number is only  $2^{15}$ ), we can still maintain an  $\epsilon \approx 5\%$ , covering almost the entire representable domain between  $-10^4$  and  $10^4$ .

**Limitations on accuracy.** As NFGGen uses approximations, there is no guarantee that it will find a workable polynomial with small  $\epsilon$ s. We have shown that when  $\epsilon = 10^{-3}$ , we can successfully find approximations for all 15 functions. When we set  $\epsilon = 10^{-4}$ , we fail to find a good  $\hat{p}_k^m$  for  $\text{Gamma\_dis}$ . If we further limit  $\epsilon$  to  $10^{-5}$ , six out of the 15 functions fails to fit. However, considering that MPC is mostly employed on data mining applications that do not require high precision, we believe NFGGen strikes the right balance between efficiency and predictable accuracy.

## 7.4 Applied in Real Algorithms

NFGGen benefits MPC algorithms mainly in two ways. First, it improves both performance and accuracy for existing MPC algorithms. Second, it allows people to evaluate advanced non-linear functions that we cannot construct with simple built-ins. We use logistic regression (LR) as an example to show the first benefit, and use a series of special functions and the  $\chi^2$  test to show the second.

**Logistic Regression (LR) Accuracy.** LR [6] is one of the most utilized data mining algorithms, both in plaintext and MPCs. The

**Figure 6: Sigmoid Approximation Comparison****Table 5: Performance Analysis of sigmoid**

Data setting*	Real	NFGGen	SecureML [32, 33]	[20]
(0.5, 3, 0.2)	59.2	59.2	49.2	58.6
(0.6, 3, 0.1)	62.5	62.5	50.8	62.2
(0.7, 3, 0.1)	65.8	65.8	51.4	65.3
(0.7, 4, 0.1)	61.9	61.9	50.0	61.5

\* The three numbers are arguments `class_sep`, `clusters_per_class` and `learning_rate` passed to the Python sklearn library's `make_classification()` function to generate the dataset.

**Table 6: Logistic Regression Speedups**

Dataset	Method	Train(sec)	Test(sec)
Adult [23] (48, 842 × 65)	PrivPy	413.1	1.8
	NFGGen	43.6 / 9.5×	0.8 / 2.3×
Bank [34] (41, 188 × 63)	PrivPy	72.8	1.6
	NFGGen	20.4 / 3.6×	0.8 / 2.0×
Branch [36] (400, 000 × 480)	PrivPy	703.8	12.2
	NFGGen	199.9 / 3.5×	6.9 / 1.8×

major challenge for MPC is the slow performance of evaluating *sigmoid*. Prior projects use a 3-piece linear function [32, 33], and [20] uses single *Chebyshev* polynomial to approximate the *sigmoid*. Figure 6 compares the *sigmoid* function with different approximations.

People argue that the accuracy of *sigmoid* does not affect the LR accuracy [20, 32, 33]. To evaluate this argument, we generate four datasets using Python sklearn's `make_classification()` method, and use them to train LR models using different approximations. Table 5 reports the LR prediction accuracy. We see that the 3-piece approximation can lead to significant LR accuracy loss, while the approximation in [20] slightly reduces accuracy. In comparison, NFGGen achieves almost the same accuracy as the plaintext result. Thus, *sigmoid* accuracy does affect LR performance in some cases. NFGGen provides an efficient way to evaluate *sigmoid* with high accuracy, eliminating the need for *ad hoc* approximations.

**LR performance.** We use 3 real datasets to evaluate LR training and inference time on PrivPy. We omit evaluation on MP-SPDZ as it needs to pre-compile all input data into the program, but the compiler fails on large datasets. Independent of the dataset, we set the  $\hat{x}$  domain to  $[-10, 10]$ , which is a typical setting in practice when the distribution of dataset is unknown (when  $x \notin [-10, 10]$ , output 0 or 1). Table 6 shows the results. We can see that NFGGen achieves 3.5× to 9.5× speedup in training and 1.8× to 2.3× speedup for inference using  $\hat{p}_8^0$  and with same Accuracy as plaintext LR.

**Hard-to-implement functions.** A big problem MPC practitioners face is how to implement some commonly-used but hard-to-implement functions, such as  $\gamma(x)$ ,  $\Gamma(x)$  and  $\Phi(x)$ . These functions are defined as integrals, and it takes much mathematical skills to

**Table 7: Special Functions Demonstration**

Target Function	Parameter	(k, m)	T <sub>Fit</sub> (sec)
$\gamma(x, z) = \int_0^x t^{z-1} e^t dt, x \in [0, 15]$	z = 1	(6, 4)	1.1
	z = 2	(5, 6)	1.6
	z = 3	(6, 6)	2.0
$\Gamma(x, z) = \int_x^\infty t^{z-1} e^t dt, x \in [0, 10]$	z = 1	(6, 6)	1.1
	z = 2	(8, 4)	1.1
	z = 3	(7, 4)	1.2
$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt, x \in [0, 5]$	NA	(4, 6)	0.8
$\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt, x \in [-5, 5]$	NA	(8, 6)	1.2

**Table 8:  $\chi^2$  Test using Real Datasets**

Dataset	Feature	Dof	(k, m)	Error	Time(sec)
Cervical [29] (72 × 20) 5 features for demo	Sexual behavior	5	(6, 11)	0	13.8
	Eating behavior	7	(5, 13)	0	
	Personal hygiene	11	(5, 14)	0	
	Social support	11	(5, 14)	0	
	Attitude	6	(5, 11)	0	
Sepsis [12] (110, 204 × 3)	age	10	(5, 12)	0	63.3
	sexual	1	(5, 98)	0	
	episode number	4	(5, 9)	0	

approximate them using the limited operators in MPC (and impossible sometimes). NFGen naturally solves the problem for all Lipschitz continuous functions as long as there is a plaintext implementation available. We demonstrate 8 hard-to-implement functions in Table 7. We see that like other functions, it only takes about 1-2 seconds to generate the approximation and achieve small  $k$  and  $m$  meeting the same accuracy requirement.

**The  $\chi^2$  test on real datasets.**  $\chi^2$  test [31] is a classic statistical method. Unfortunately, the  $p$  value from  $\chi^2$  test depends on the  $\gamma$  and  $\Gamma$  functions in Table 7, as  $p = 1 - \frac{\gamma(\frac{k}{2}, \frac{x}{2})}{\Gamma(\frac{k}{2})}$ , where  $\frac{\gamma(\frac{k}{2}, \frac{x}{2})}{\Gamma(\frac{k}{2})}$  is the CDF of  $\chi^2$  distribution,  $x$  is the statistical value and  $k$  is the degree-of-freedom (Dof) parameter that is typically set to the number of classes minus 1. No current MPC framework supports  $\chi^2$  test yet, to our knowledge. We show that we can easily implement  $\chi^2$  test with NFGen on real datasets. The datasets contain features of a patient with certain diseases, and as a typical task in medical research, we use  $\chi^2$  test to determine whether the probability of a disease is correlated to a feature. Table 8 shows that we can evaluate cases with different Dofs, and achieve the same result as in plaintext (with 3 significant digits).

## 7.5 Effectiveness of Design Choices

We conduct an *ablation study* to evaluate the design choices in NFGen, including the *Scaling factor* (Algorithm 5), *Residual boosting* (Algorithm 6) and the *Merge stage* in Algorithm 1, using the same setting as in Section 7.2, on *Rep2k*.

We evaluate each technique on four functions, and Table 9 summarizes the result. We measure three metrics on each ablation case: 1)  $T_{Fit}$  is the runtime of Algorithm 1; 2) Best ( $k, m$ ) is the best-performance  $\hat{p}_k^m$  we find; 3) Failures count the number of cases where we do not find a feasible  $\hat{p}_k^m$  for  $k$  with  $m < m_{max}$ .

**Table 9: Ablation Studies**

F(x)	Metric	NFGen	no merge	no boosting	no scaling
sigmoid	$T_{Fit}$	3.1	2.2	14.0	2.6
	Best ( $k, m$ )	(7, 10)	(5, 20)	(7, 10)	(7, 12)
	Failures	0	0	1	0
soft_plus	$T_{Fit}$	2.7	1.8	14.1	2.0
	Best ( $k, m$ )	(7, 4)	(4, 30)	(7, 4)	(4, 9)
	Failures	0	0	1	0
selu	$T_{Fit}$	1.3	0.9	1.2	36.0
	Best ( $k, m$ )	(8, 9)	(6, 14)	(8, 9)	(6, 12)
	Failures	0	0	0	3
Gamma_dis	$T_{Fit}$	4.3	1.1	40.8	1.2
	Best ( $k, m$ )	(7, 21)	(5, 35)	NA	(5, 26)
	Failures	0	0	7	0

Key observations include: 1) Though each technique takes extra pre-computation time, the overhead is less than 1 second per technique. However, without them some functions are even slower to fit, e.g. *selu* w/o scaling and *Gamma\_dis* w/o residual boosting are both slower because some otherwise possible  $\hat{p}_k^m$  will not pass the check and results in more searching steps. 2) The merge step significantly reduces  $m$ , as the splitting strategy often unnecessarily increases the number of pieces. 3) Failures are more often if we remove residual boosting or scaling, showing that they effectively make some approximation possible by remedying inaccuracies introduced by the FLP-FXP conversion.

## 8 CONCLUSION AND FUTURE WORK

Creating general-purpose MPC platforms is analogous to creating a new computation system from scratch using MPC primitives instead of instructions. Non-linear function evaluation, akin to plaintext numeric libraries, is one of these systems' *foundations*. Prior approaches either naively attempted to reuse plaintext algorithms that resulted in erroneous results and/or slow performance, or developed *ad hoc* approximations that were tightly coupled with either specific functions or MPC platforms. Neither method possesses the generality or performance necessary to serve as a viable *foundation*. NFGen is, to our knowledge, the first attempt for a generic solution. We can accurately approximate general non-linear functions using piecewise polynomials by properly handling FXP and FLP operations. We achieve portability across multiple MPC systems and protocols by utilizing code generation and profiler-based performance prediction. Extensive evaluations verify our approach's effectiveness, accuracy, and generality.

As future work, we are going to support more MPC platforms, explore approximation algorithms with stronger theoretical guarantees, as well as support the evaluation of multi-dimensional non-linear functions.

## ACKNOWLEDGEMENTS

We thank Menghua Cao and Zhilong Chen for insightful discussions during the design of NFGen. We thank Yuanxi Dai and Xinze Li for their assistance with the security analysis. We thank Xiang Wang and Haoqing He for their help during the implementations.

This work is supported in part by the National Natural Science Foundation of China (NSFC) Grant 71872094 and gift funds from Nanjing Turing AI Institute.

## REFERENCES

- [1] 2019. IEEE Standard for Floating-Point Arithmetic. *IEEE Std 754-2019 (Revision of IEEE 754-2008)* (2019), 1–84. <https://doi.org/10.1109/IEEESTD.2019.8766229>
- [2] Mehrdad Aliasgari, Marina Blanton, Yihua Zhang, and Aaron Steele. 2013. Secure Computation on Floating Point Numbers. In *The Network and Distributed System Security Symposium (NDSS)*.
- [3] Kendall Atkinson and Weimin Han. 2005. *Theoretical Numerical Analysis*. Vol. 39. Springer.
- [4] Donald Beaver. 1991. Efficient Multiparty Protocols using Circuit Randomization. In *Annual International Cryptology Conference*. Springer.
- [5] Zygmunt W Birnbaum and Sam C Saunders. 1969. A New Family of Life Distributions. *Journal of applied probability* (1969).
- [6] Christopher M Bishop and Nasser M Nasrabadi. 2006. *Pattern Recognition and Machine Learning*, Chapter 4.3.4. Springer.
- [7] Dan Bogdanov, Sven Laur, and Jan Willemson. 2008. Sharemind: A Framework for Fast Privacy-preserving Computations. In *European Symposium on Research in Computer Security*. Springer.
- [8] Chillotti Ilaria Gama Nicolas Jetchev Dimitar Peceny Stanislav Petric Alexander Boura, Christina. 2018. High-Precision Privacy-Preserving Real-Valued Function Evaluation. In *International Conference on Financial Cryptography and Data Security (FC)*.
- [9] Ran Canetti. 2000. Security and Composition of Multiparty Cryptographic Protocols. *Journal of CRYPTOLOGY* (2000).
- [10] Ran Canetti. 2001. Universally Composable Security: A new Paradigm for Cryptographic Protocols. In *IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE.
- [11] Octavian Catrina and Amitabh Saxena. 2010. Secure Computation with Fixed-point Numbers. In *International Conference on Financial Cryptography and Data Security (FC)*. Springer.
- [12] Davide Chicco and Giuseppe Jurman. 2020. Survival Prediction of Patients with Sepsis from Age, Sex, and Septic Episode Number Alone. *Scientific reports* (2020).
- [13] Ronald Cramer, Ivan Damgård, Daniel Escudero, Peter Scholl, and Chaoping Xing. 2018.  $\text{SPDZ}_{k,k}$ : Efficient MPC mod  $2^k$  for Dishonest Majority. In *Advances in Cryptology—Crypto*.
- [14] Anders Dalskov, Daniel Escudero, and Marcel Keller. 2020. Secure Evaluation of Quantized Neural Networks. *Proceedings on Privacy Enhancing Technologies (PET)* (2020).
- [15] Ivan Damgård, Daniel Escudero, Tore Frederiksen, Marcel Keller, Peter Scholl, and Nikolaj Volgushev. 2019. New Primitives for Actively-secure MPC over Rings with Applications to Private Machine Learning. In *IEEE Symposium on Security and Privacy (S&P)*. IEEE.
- [16] Ivan Damgård, Matthias Fitzi, Eike Kiltz, Jesper Buus Nielsen, and Tomas Toft. 2006. Unconditionally secure constant-rounds multi-party computation for equality, comparison, bits and exponentiation. In *Theory of Cryptography Conference (TCC)*. Springer.
- [17] Daniel Demmler, Thomas Schneider, and Michael Zohner. 2015. ABY-A framework for Efficient Mixed-protocol Secure Two-party Computation.. In *The Network and Distributed System Security Symposium (NDSS)*.
- [18] Kyoohyung Han, Jinhyuck Jeong, Jung Hoon Sohn, and Yongha Son. 2020. Efficient Privacy Preserving Logistic Regression Inference and Training. *Cryptology ePrint Archive* (2020).
- [19] Marcella Hastings, Brett Hemenway, Daniel Noble, and Steve Zdancewicz. 2019. SoK: General Purpose Compilers for Secure Multi-party Computation. In *IEEE Symposium on Security and Privacy (S&P)*. IEEE.
- [20] Ehsan Hesamifard, Hassan Takabi, Mehdi Ghasemi, and Rebecca N Wright. 2018. Privacy-preserving Machine Learning as a Service. *Proceedings on Privacy Enhancing Technologies (PET)* (2018).
- [21] Marcel Keller. 2020. MP-SPDZ: A Versatile Framework for Multi-party Computation. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.
- [22] Brian Knott, Shobha Venkataraman, Awni Hannun, Shubho Sengupta, Mark Ibrahim, and Laurens van der Maaten. 2021. CrypTen: Secure Multi-party Computation Meets Machine Learning. *Advances in Neural Information Processing Systems (NIPS)* (2021).
- [23] Ron Kohavi et al. 1996. Scaling up the Accuracy of Naive-bayes Alassifiers: A Decision-tree Hybrid.. In *ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*.
- [24] The Scikit learn Documentation. 2011. Polynomial regression: extending linear models with basis functions. [https://scikit-learn.org/stable/modules/linear\\_model.html](https://scikit-learn.org/stable/modules/linear_model.html). (2011).
- [25] Yi Li and Wei Xu. 2019. PrivPy: General and Scalable Privacy-preserving Data Mining. In *ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*.
- [26] Yehuda Lindell and Ariel Nof. 2017. A Framework for Constructing Fast MPC over Arithmetic Circuits with Malicious Adversaries and an Honest-majority. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.
- [27] Yehuda Lindell and Benny Pinkas. 2000. Privacy Preserving Data Mining. In *Annual International Cryptology Conference*. Springer.
- [28] Donghang Lu, Thomas Yurek, Samarth Kulshreshtha, Rahul Govind, Aniket Kate, and Andrew Miller. 2019. Honeybadgermpc and asynchronix: Practical asynchronous mpc and its application to anonymous communication. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.
- [29] Rizanda Machmud, Adi Wijaya, et al. 2016. Behavior Determinant Based Cervical Cancer Early Detection with Machine Learning Algorithm. *Advanced Science Letters* (2016).
- [30] Wolfram MathWorld. 2015. Lagrange interpolating polynomial. (2015).
- [31] William Mendenhall, Robert J Beaver, and Barbara M Beaver. 2012. *Introduction to Probability and Statistics*, Chapter 14. Cengage Learning.
- [32] Payman Mohassel and Peter Rindal. 2018. ABY3: A Mixed Protocol Framework for Machine Learning. In *ACM SIGSAC conference on computer and communications security (CCS)*.
- [33] Payman Mohassel and Yupeng Zhang. 2017. SecureML: A System for Scalable Privacy-preserving Machine Learning. In *IEEE Symposium on Security and Privacy (S&P)*. IEEE.
- [34] Sérgio Moro, Paulo Cortez, and Paulo Rita. 2014. A Data-driven Approach to Predict the Success of Bank Telemarketing. *Decision Support Systems* (2014).
- [35] Deevashwer Rathee, Mayank Rathee, Rahul Kranti Kiran Goli, Divya Gupta, Rahul Sharma, Nishanth Chandran, and Aseem Rastogi. 2021. SIRONN: A Math Library for Secure RNN Inference. In *IEEE Symposium on Security and Privacy (S&P)*. IEEE.
- [36] DMITRY SHKADAREVICH. 2021. Branch prediction Binary Classification Dataset. <https://www.kaggle.com/dmitryshkadarevich/branch-prediction>. (2021).
- [37] Sijun Tan, Brian Knott, Yuan Tian, and David J Wu. 2021. CrypGPU: Fast Privacy-preserving Machine Learning on the GPU. In *IEEE Symposium on Security and Privacy (S&P)*. IEEE.
- [38] Lloyd N Trefethen. 2019. *Approximation Theory and Approximation Practice*, Extended Edition. Society for Industrial and Applied Mathematics (SIAM).
- [39] Fan Xiaoyu, Wang Guosai, Chen Kun, He Xu, and Xu Wei. 2021. PPCA: Privacy-preserving Principal Component Analysis using Secure Multiparty Computation (mpc). *arXiv preprint arXiv:2105.07612* (2021).
- [40] Fan Xiaoyu, Chen Kun, Wang Guosai, Zhuang Mingchun, Li Yi, and Xu Wei. 2022. NFGen: Automatic Non-linear Function Evaluation Code Generator for General-purpose MPC Platforms. *arXiv preprint arXiv:2210.09802* (2022).