Alternating moduli PRFs and their polynomial representations

Håvard Raddum

26.01.2025

Simula UiB

Outline

- weak pseudo-random functions (wPRF)
- Constructions mixing linear functions over $\mathbb{F}_2$ and $\mathbb{F}_3$
- Polynomial representation of mappings
  - Impossibility result: is it sufficient?
- Ideas for further study

weak pseudorandom function (wPRF)

A mapping $F \colon \mathscr{K} \times \mathscr{X} \longrightarrow \mathscr{Y}$ where

$\mathscr{K}$ is the key space, $\mathscr{X}$ is the input space and $\mathscr{Y}$ is the output space

Property:

For fixed $k \in \mathscr{K}$, can not distinguish

$$\begin{bmatrix} (x_1, F_k(x_1)) \\ (x_2, F_k(x_2)) \\ \vdots \\ (x_s, F_k(x_s)) \end{bmatrix} \quad \text{from} \quad \begin{bmatrix} (x_1, g(x_1)) \\ (x_2, g(x_2)) \\ \vdots \\ (x_s, g(x_s)) \end{bmatrix}$$

where $g : \mathscr{X} \longrightarrow \mathscr{Y}$ is a random function,

the x_i are drawn uniformly at random from $\mathscr{X}$, and $s < 2^{\lambda}$

wPRFs from mixing $\mathbb{F}_2$ and $\mathbb{F}_3$

Main idea

Build wPRF by combining *linear* mappings over $\mathbb{F}_2$ and $\mathbb{F}_3$

- Simple design
- Very efficient for use in MPC (few communication rounds)
- Gives high algebraic degree when expressed over a single field

Generalization: mappings over $\mathbb{F}_p$ and $\mathbb{F}_q$

Notation: elements and computations are red in $\mathbb{F}_2$, and blue in $\mathbb{F}_3$

DarkMatter (2018)

TCC 2018, eprint 2018/1281

BIP+18 presents idea and first construction (single output)

$\mathcal{X} = \mathbb{F}_2^n, \mathcal{K} = \mathbb{F}_2^{m \times n} \quad \mathcal{Y} = \mathbb{F}_3$

$$y = [1 1 \cdots 1] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ w_m \end{bmatrix} \overset{\mathrm{Id}}{\leftarrow} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} \mathbf{k} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Special matrices

k is circulant matrix, given by top row

k is Toeplitz matrix,

given by top row and leftmost column

suggested (optimistic) parameters for λ-bit security: $n = m = 2\lambda$

DarkMatter alternative constructions

basic LPN variant

$\mathcal{X} = \mathbb{F}_2^n, \mathcal{K} = \mathbb{F}_2^n, \mathcal{Y} = \mathbb{F}_2$

$x_i \overset{\text{Id}}{\to} x_i \qquad k_i \overset{\text{Id}}{\to} k_i$

$w = x_1k_1 + x_2k_2 + \ldots + x_nk_n$

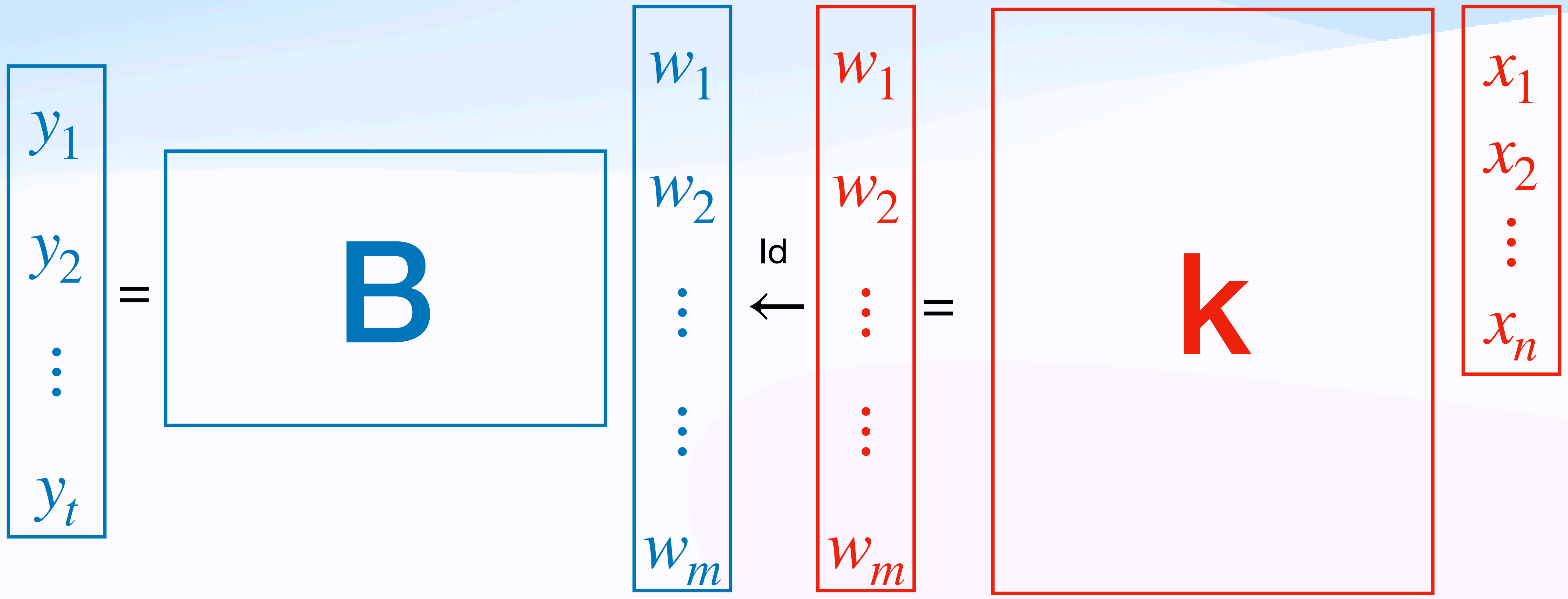
$w \overset{\text{mod 2}}{\to} w$

$y = x_1k_1 + x_2k_2 + \ldots + x_nk_n + w$

«LPN with error rate 1/3»

multi output variant

$\mathcal{X} = \mathbb{F}_2^n, \mathcal{K} = \mathbb{F}_2^{m\times n} \quad \mathcal{Y} = \mathbb{F}_3^t$

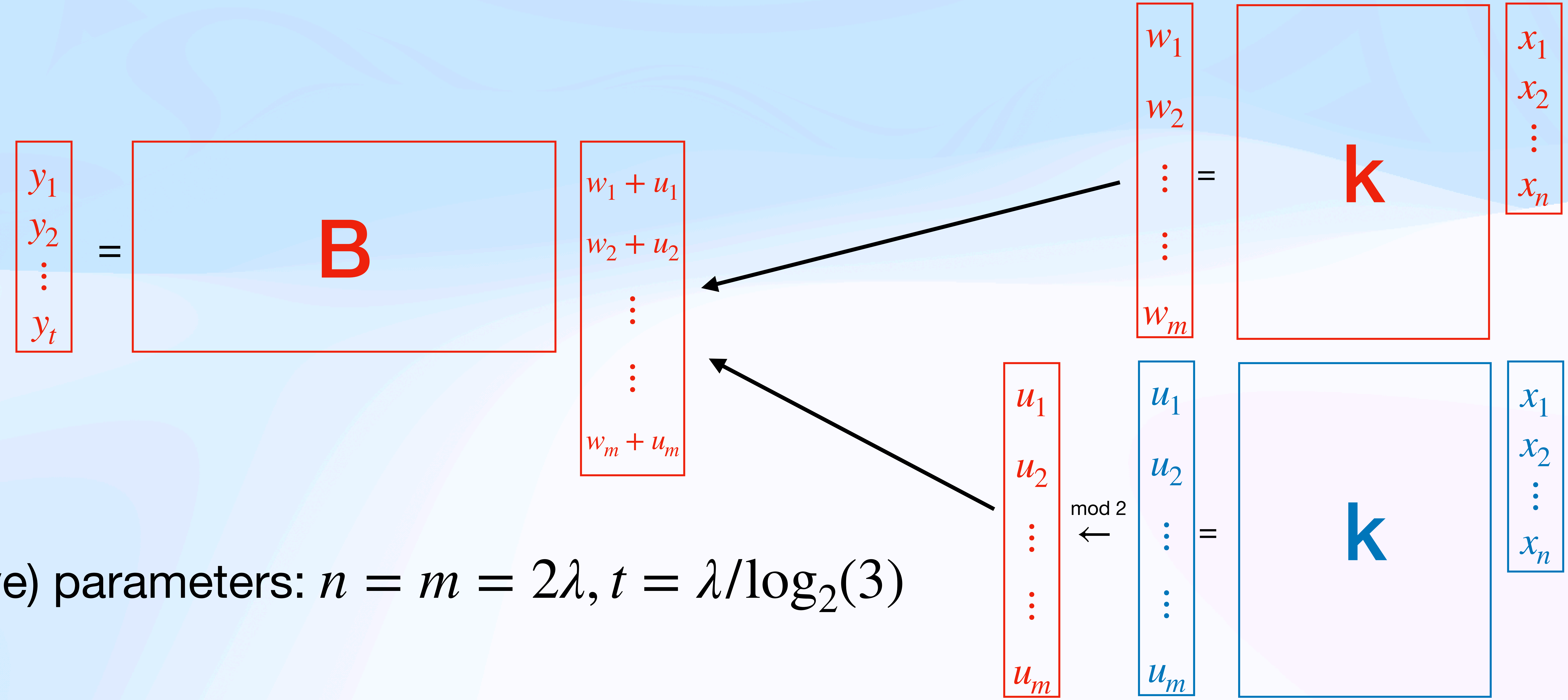


where $t \leq m - \lambda$

DGH+21 construction

CRYPTO 2021, eprint 2021/885

multi output LPN variant $\mathcal{X} = \mathbb{F}_2^n, \mathcal{K} = \mathbb{F}_2^{m\times n}, \mathcal{Y} = \mathbb{F}_2^t$



(agressive) parameters: $n = m = 2\lambda, t = \lambda/\log_2(3)$

APRR24 construction

CRYPTO 2024, eprint 2024/582

$\mathcal{X} = \mathbb{F}_2^n, \mathcal{K} = \mathbb{F}_2^n \qquad \mathcal{Y} = \mathbb{F}_3^t$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{bmatrix} = \mathbf{B} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ w_m \end{bmatrix} \xleftarrow{\text{Id}} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ w_m \end{bmatrix} = \mathbf{A} \begin{bmatrix} x_1 \cdot k_1 \\ x_2 \cdot k_2 \\ \vdots \\ x_n \cdot k_n \end{bmatrix}$$

1-to-1 parameters: $n = 2\lambda, m = 7.06\lambda, t = 2\lambda/\log_2(3)$

many-to-1 parameters: $n = 4\lambda, m = 2\lambda, t = \lambda/\log_2(3)$

Cryptanalysis so far

- CCKK20 PKC 2020, eprint 2020/783
  - Attacks basic wPRF of BIP+18 with circulant matrix and basic LPN version
  - Exploits biases in the modular reductions
  - Parameters in original constructions must be increased
- MR24 eprint 2024/2055
  - Attacks 1-to-1 parameter set of APRR24 construction
  - Exploits collisions in output (mapping is not 1-to-1)
  - wPRF gives only $\lambda/2$-bit security

Polynomial representations

On polynomial representation

- BIP+18 argues the mixed moduli wPRFs do not admit representation by low-degree polynomials over a fixed field
- CCKK20 does not consider polynomial representations
- DGH+21 refers to BIP+18, and does not consider polynomial representation further
- APRR24 shows polynomial representation over $\mathbb{F}_3$ that is surprisingly compact, but does not investigate further

BIP+18 argument

Smo87

$MOD_{(s,p)}$ - outputs one iff the number of ones in the input is congruent to s mod p. $MOD_p = NOT(MOD_{(0,p)})$.

Theorem 2: Let p be a prime number and r is not a power of p then computing MOD_r by depth k circuit with basic operations AND, OR, NOT and MOD_p requires $exp(O(n^{\frac{1}{2k}}))$ AND and OR gates.

4.2 Inapproximability by Low-Degree Polynomials

Another necessary condition for a PRF family is that the family should be hard to approximate by low-degree polynomials. Specifically, assume there exists a degree-d multivariate polynomial

19

f over GF(2) such that $\mathsf{F}_k(x) = f(x)$ for all $x \in \{0,1\}^n$. Then, given (sufficiently many) PRF evaluations $(x_i, \mathsf{F}_k(x_i))$ on uniformly random values x_i, an adversary can set up a linear system where the unknowns corresponds to the coefficients of f. Since f has degree d, the resulting system has $N = \sum_{k=0}^{d} \binom{n}{k}$ variables. Thus, given $O(2^d \cdot N)$ random samples, the adversary can solve the linear system and recover the coefficients of f (and therefore, a complete description of F_k). We note that this attack still applies even if F_k is $1/O(2^d \cdot N)$-close to a degree d polynomial. In this case, the solution to the system will be $1/O(2^d \cdot N)$-close to F_k with constant probability (which still suffices to break pseudorandomness). Thus, for a candidate PRF family to be secure, the family should not admit a low-degree polynomial approximation.

In our setting, we are able to rule out low-degree polynomial approximations by appealing to the classic Razborov-Smolensky lower bounds for ACC^0 [Raz87, Smo87], which essentially says that for distinct primes p and q, MOD_p gates cannot be computed in $\mathsf{ACC}^0[q^\ell]$ for any $\ell \geq 1$. Translated to our setting, this essentially says that our "modulus-switching" mapping $\mathsf{map}_p \colon \{0,1\}^n \to \mathbb{Z}_p$, which implements the mapping $x \mapsto \sum_{i\in[n]} x_i \pmod p$, is hard to approximate over $\mathrm{GF}(q^\ell)$ as long as $p \neq q$. We formalize this in the following lemma.

Lemma 4.2 (Inapproximability by Low-Degree Polynomials). *For $n > 0$ and $d < n/2$, let $B(n,d) = \frac{1}{2^n} \cdot \sum_{i=0}^{n/2-d-1} \binom{n}{i}$. Then, for all primes $p \neq q$, the function $\mathsf{map}_p \colon \{0,1\}^n \to \mathbb{Z}_q$ on n-bit inputs that maps $x \mapsto \sum_{i\in[n]} x_i \pmod p$ is $B(n,d)$-far from all degree-d polynomials over $\mathrm{GF}(q^\ell)$ for all $\ell \geq 1$.*

BIP+18 conjecture

4.3 Inapproximability by Low-Degree Rational Functions

The low-degree polynomial approximation attack described in Section 4.2 generalizes to the setting where the PRF F_k can be approximated (sufficiently well) by a low-degree *rational* function. For instance, suppose there exist multivariate polynomials f, g over GF(2) of degree at most d such that $f(x) = \mathsf{F}_k(x) \cdot g(x)$ for all $x \in \{0,1\}^n$. Then, a similar attack can be mounted, as any random

Conjecture 4.3 (Inapproximability by Rational Functions)**.** For any distinct primes $p \neq q$, any integer $\ell \geq 1$, and any $d = o(n)$, there exists a constant $\alpha < 1$ such that the function $\mathsf{map}_p \colon \{0,1\}^n \to \mathbb{Z}_p$ that maps $x \mapsto \sum_{i \in [n]} x_i \pmod p$ is $1/(2^d \cdot N)^\alpha$-far from all degree-d rational functions over $\mathrm{GF}(q^\ell)$.

We believe that studying this conjecture is a natural and well-motivated complexity problem. Proving or disproving this conjecture would lead to a better understanding of ACC^0.

Motivates further study on polynomial approximations of mod2/mod3-constructions

APRR24 observation

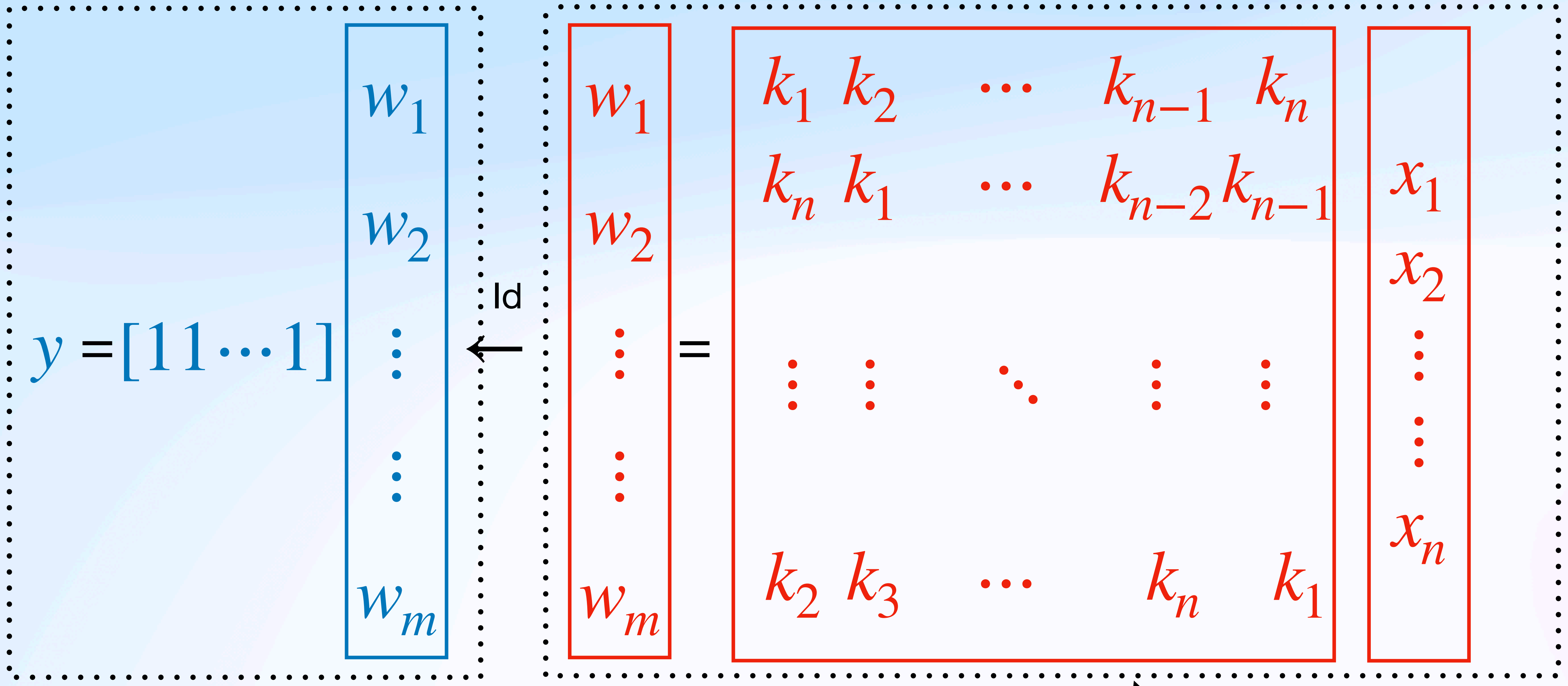
BIP+18 only considers approximating MOD$_p$ on inputs from $\{0,1\}^n$

no low-degree

polynomial

approximation

over $\mathbb{F}_q$ for $q \neq 3$



Polynomial approximation over $\mathbb{F}_3$?

APRR24 observation

$$(\mathbb{F}_2, +) \cong (\mathbb{F}_3^*, \times)$$
$$x \mapsto x + 1$$

$$x_1k_1 + x_2k_2 + \ldots + x_nk_n \cong \prod_{i=1}^{n} (k_i + 1)^{x_i}$$

Linear variable change $k_i + 1 = \bar{k}_i$

$$x_1k_1 + x_2k_2 + \ldots + x_nk_n \cong \prod_{i=1}^{n} \bar{k}_i^{x_i}$$

APRR24 observation

$$y = [11\cdots1]\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ w_n \end{bmatrix} \xleftarrow{\text{Id}} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} k_1 & k_2 & \cdots & k_{n-1} & k_n \\ k_n & k_1 & \cdots & k_{n-2} & k_{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ k_2 & k_3 & \cdots & k_n & k_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix}$$

$$\cong$$

$$y = [11\cdots1]\begin{bmatrix} w_1 - 1 \\ w_2 - 1 \\ \vdots \\ \vdots \\ w_n - 1 \end{bmatrix} \xleftarrow{\text{Id}} \begin{bmatrix} w_1 = \prod_{i=1}^{n} \bar{k}_i^{x_i} \\ w_2 = \prod_{i=1}^{n} \bar{k}_{i-1}^{x_i} \\ \vdots \\ w_n = \prod_{i=1}^{n} \bar{k}_{i+1}^{x_i} \end{bmatrix}$$

Basic wPRF can be described by polynomial over $\mathbb{F}_3$

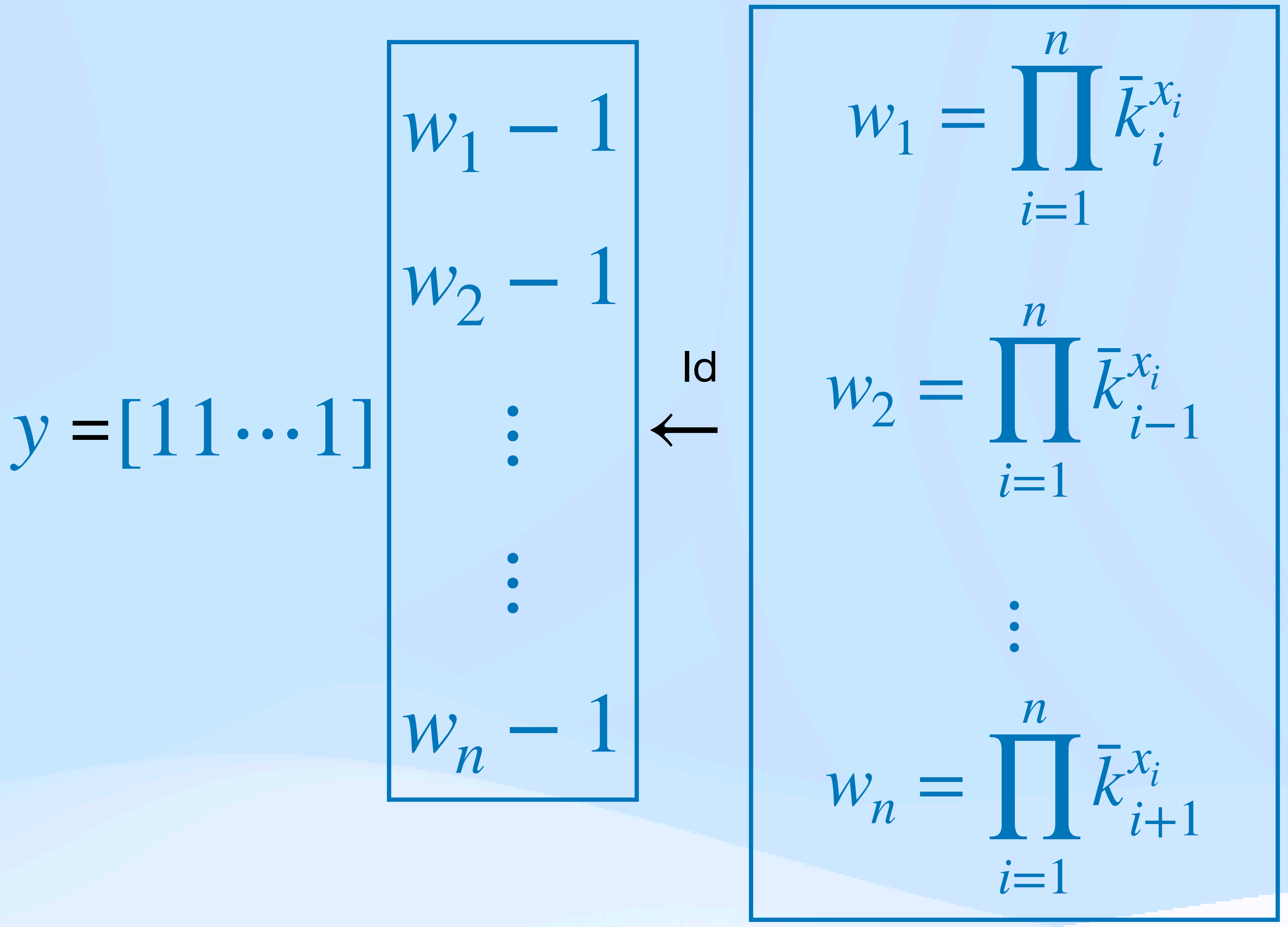
of degree $\approx n/2$, but only m terms

would correspond to *interpolating* sparse multilinear polynomials. While the connection between symmetric-key primitives (based on the alternating-moduli paradigm) and the hardness of interpolating sparse multilinear polynomials has already been observed by [BIP$^+$18], neither of [BIP$^+$18] or [DGH$^+$21] considered the dual problem of solving a system of sparse multilinear polynomial equations for their constructions.

Further observations

The set of terms $\{w_1, \ldots, w_n\}$ for x and $\{w'_1, \ldots, w'_n\}$ for x' are:

- equal if $x' = (x <<< i)$ for some i
- disjoint if $x' \neq (x <<< i)$ for any i

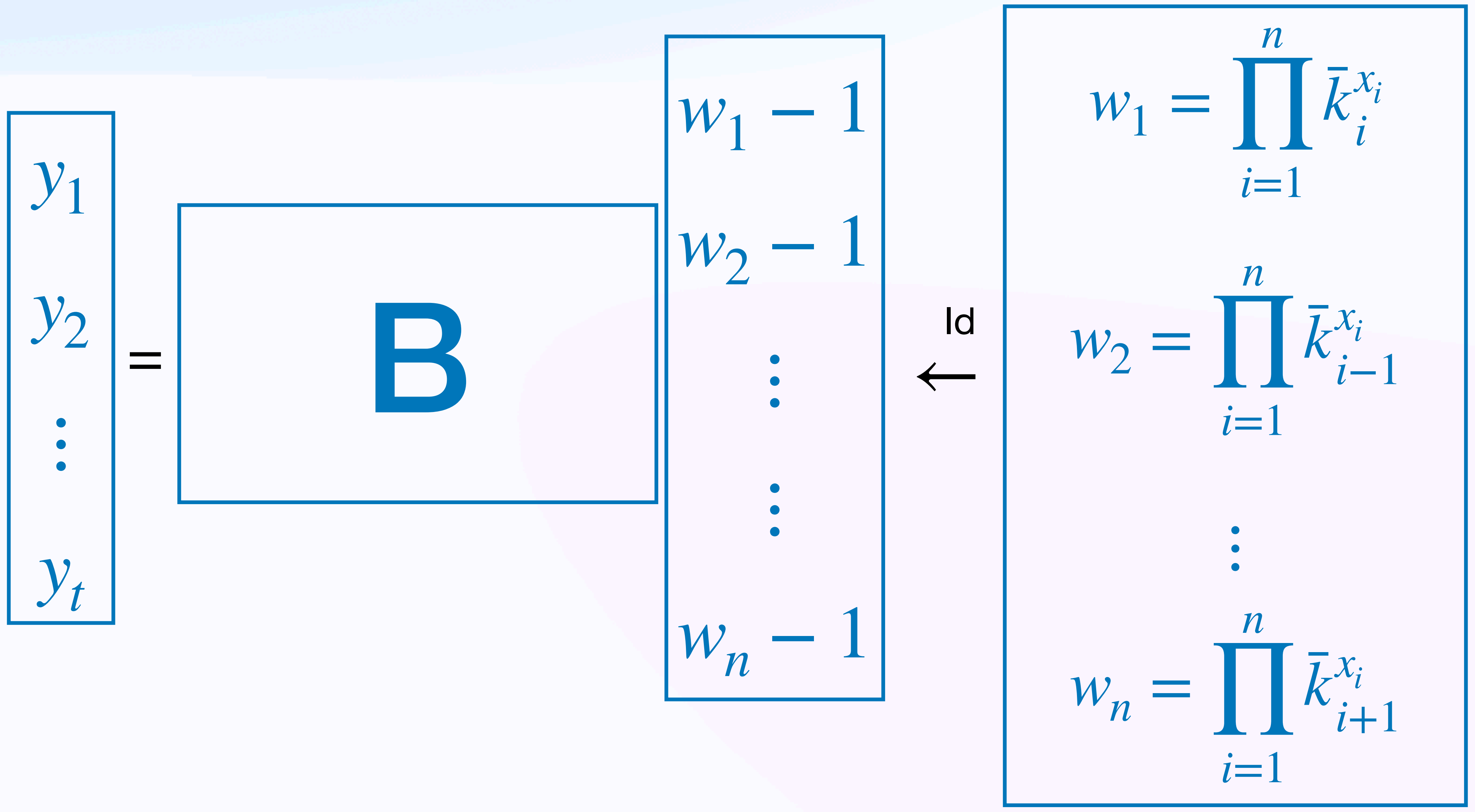


Multi-output version generates t polynomial equations in n terms for every query

$$n = 2\lambda, t = n - \lambda = \lambda$$

$$\Downarrow$$

Enough to find two queries where x and x' are rotations of each other to solve system



Ideas for further study

- Idea 1: Express each output element using multiple polynomials over $\mathbb{F}_2$
  - are we sure no such expression can consist of multiple low-degree polynomials?
- Idea 2: Investigate conjecture that $f(x) \cdot \text{wPRF}(x) = g(x)$ must have high-degree f, g
- Idea 3: Pursue $(\mathbb{F}_2, +) \cong (\mathbb{F}_3, \times)$ observation
  - how many queries must be made before we can expect to find x and x' that are rotations of each other?