

BASS: Boolean Automorphisms Signature Scheme

DIMA GRIGORIEV, ILIA ILMER, ALEXEY OVCHINNIKOV, AND VLADIMIR SHPILRAIN

ABSTRACT. We offer a digital signature scheme using Boolean automorphisms of a multivariate polynomial algebra over integers. Verification part of this scheme is based on the approximation of the number of zeros of a multivariate Boolean function.

Keywords: digital signature, multivariate polynomial, Boolean function

1. INTRODUCTION

Due to the concern that if large-scale quantum computers are ever built, they will compromise the security of many commonly used cryptographic algorithms, NIST had begun in 2016 a process to develop new cryptography standards and, in particular, solicited proposals for new digital signature schemes [5] resistant to attacks by known quantum algorithms, such as e.g. [6].

One possible way to avoid quantum attacks based on solving the hidden subgroup problem (including the attacks in [6]) is *not* to use one-way functions that utilize one or another (semi)group structure. The candidate one-way function that we use in our scheme here takes a private polynomial automorphism φ as the input and outputs $\varphi(P)$ for a public multivariate polynomial P .

To avoid any parallels with the encryption scheme of [4], we say up front that since ours is just a signature scheme (i.e., is not a spin-off of any encryption scheme), we do not need our candidate one-way function to have a trapdoor because the private key holder does not need to invert the function. Also, in [4], the candidate one-way function was φ itself, and the private (decryption) key was φ^{-1} . In contrast, in our signature scheme φ^{-1} does not play any role and does not have to be computed.

The main novelty of our signature scheme is manifested in the verification part. First, note that any polynomial P has as many zeros as $\varphi(P)$ does, where φ is any automorphism of the polynomial algebra. To balance between security and efficiency, we do not want the number of zeros to be either too small or too large. To that end, we use polynomials over integers, but we count zero values on Boolean tuples only. Since the number of Boolean tuples is exponential in the number of variables, it can still be too large to process deterministically. Instead, we use a non-deterministic (Monte Carlo) method to count the number of zero (or nonzero) values of a polynomial in question. We note that the accuracy of the Monte Carlo method for estimating the number of zeros of a multivariate polynomial was studied and quantified in [1].

2. SCHEME DESCRIPTION

Let $K = \mathbb{Z}[x_1, \dots, x_n]$ denote the algebra of polynomials in n variables over the ring \mathbb{Z} of integers, and let $B(K)$ denote the factor algebra of K by the ideal generated by all polynomials of the form $(x_i^2 - x_i)$, $i = 1, \dots, n$. Informally, one can call $B(K)$ the

“Booleanization” of K . We note that the ring $B(K)$ is isomorphic (as a ring) to the direct sum of 2^n copies of the ring \mathbb{Z} .

The signature scheme is as follows.

Private: an automorphism φ of the algebra $B(K)$. We note that φ is defined by the polynomials $y_i = \varphi(x_i)$, $i = 1, \dots, n$.

Public:

- 3 sparse polynomials $P_i = P_i(x_1, \dots, x_n)$ with coefficients ± 1 .
- 3 polynomials $\varphi(P_i)$, where φ is a private automorphism of the algebra $B(K)$. We note that $\varphi(P_i) = P_i(y_1, \dots, y_n)$, where $y_i = \varphi(x_i)$.
- a hash function H with values in the algebra $B(K)$ and a (deterministic) procedure for converting values of H to sparse polynomials from the algebra $B(K)$.

Remark 1. *We emphasize that the automorphism φ , the 3 sparse polynomials P_i , and the 3 polynomials $\varphi(P_i)$ are all generated/computed in the offline phase. The hash function H is one of the standard hash functions (we suggest SHA-256), with values converted from a bit string to a polynomial in $B(K)$ (see Section 3.3).*

Signing a message m :

1. Apply the hash function H to the message m . Convert $H(m)$ to a polynomial $Q = Q(x_1, \dots, x_{n+1})$ with integer coefficients using a deterministic public procedure (see Section 3.3). That is, the polynomial Q has an extra variable compared to the polynomials P_i .
2. The automorphism φ is extended to the “Booleanization” of the algebra $\mathbb{Z}[x_1, \dots, x_{n+1}]$ by taking x_{n+1} to $x_{n+1} + r(x_1, \dots, x_n) - 2x_{n+1} \cdot r(x_1, \dots, x_n)$, where $r(x_1, \dots, x_n)$ is a polynomial from the set \mathbb{G} of polynomials (see Section 4). This extended automorphism we denote by the same letter φ . (The fact that this is, indeed, an automorphism of the “Booleanization” is part of Proposition 1 in Section 4.)
3. The signature is $\varphi(Q)$.

Remark 2. *The reason why we extend the automorphism φ by adding an extra variable x_{n+1} at Step 2 is to prevent the forger from accumulating many pairs $(Q, \varphi(Q))$ with the same φ . Now, with each new signature, we have a different φ because of a random choice of the polynomial $r(x_1, \dots, x_n)$ at Step 2.*

Verification.

1. The verifier computes $H(m)$ and converts $H(m)$ to $Q = Q(x_1, \dots, x_{n+1})$ using a deterministic public procedure.
2. The verifier selects a random 4-variable polynomial $u(x, y, z, t)$ with coefficients from \mathbb{Z} and computes $u(\varphi(P_1), \varphi(P_2), \varphi(P_3), \varphi(Q))$. Note that this is equal to $\varphi(u(P_1, P_2, P_3, Q))$. Denote the polynomial $\varphi(u(P_1, P_2, P_3, Q))$ by S .
3. The verifier also computes $u(P_1, P_2, P_3, Q)$. Denote this polynomial by R . (Note that S should be equal to $\varphi(R)$ if the signature is valid.)

4. The verifier then compares the proportion of positive values on Boolean tuples for the polynomials R and S . That is, the proportion of positive values on $(n + 1)$ -tuples (x_1, \dots, x_{n+1}) , where each x_i is 0 or 1. It is not immediately obvious how the verifier can do that, given that with suggested parameters, the number of Boolean $(n + 1)$ -tuples is quite large (2^{n+1} , to be exact) and given that this kind of counting is #P-hard, see [7], so it is unlikely that there are any “short cuts” to this counting. We explain a non-deterministic method in the following subsection; here we just say that the verifier accepts the signature if and only if the proportions of positive values for R and S differ by no more than 3% of the total number of trials. See Section 4 for an explanation of why these proportions should be the same (when computed deterministically) if S is an automorphic image of R .

Correctness. While it is obvious that polynomials P and $\varphi(P)$ have the same number of zeros, it is not at all obvious why they have the same number of positive values on Boolean tuples. Indeed, this may not be true for an arbitrary automorphism φ , so we have a special algorithm for sampling φ . This is explained in Section 4, and correctness is formally proved in the Appendix.

2.1. Monte Carlo method for counting positive values of a polynomial on Boolean tuples. Our non-deterministic method for estimating the proportion of positive values on Boolean tuples for a given polynomial P is pretty standard. Just plug in a large number (we suggest 1,000) of random Boolean tuples into P and count how many of them yield a positive value of P . Then divide the obtained number by the total number of Boolean tuples used; this is your proportion.

We note that the accuracy of the Monte Carlo method for counting *zeros* of Boolean polynomials was studied and quantified in [1].

3. KEY GENERATION

First we note that, since the algebra $B(K)$ is the factor algebra of K by the ideal generated by all polynomials of the form $(x_i^2 - x_i)$ and since we only count values of a polynomial on Boolean tuples, when we generate the public polynomials P_i it makes sense to only generate monomials where no x_j occurs with an exponent higher than 1. Then generating P_i will look as follows.

3.1. Generating a random t -sparse polynomial.

- (1) Select, uniformly at random, an integer d between 1 and b (where b is one of the parameters of the scheme). This integer will be the degree of our monomial. (Note that the degree of a monomial cannot be higher than n since our monomials are square-free because of factoring by the ideal generated by all polynomials of the form $(x_i^2 - x_i)$.)
- (2) To select a monomial of degree d , do a selection of x_i , uniformly at random from $\{x_1, \dots, x_n\}$, d times, avoiding repetition of x_i . Then build the monomial as a product of the selected x_i .
- (3) Finally, build a t -sparse polynomial as a linear combination of t selected monomials with coefficients ± 1 , selected at random.

3.2. Generating a random polynomial from the set \mathbb{G} . The set \mathbb{G} of polynomials in $\mathbb{Z}[x_1, \dots, x_n]$ will play a crucial role in generating automorphisms of the algebra $B(K)$, see Section 4. This set can be defined recursively as follows. Assign all variables x_1, \dots, x_n to \mathbb{G} . Then keep adding more polynomials to \mathbb{G} using the following rules: (1) if a polynomial P belongs to \mathbb{G} , then $1 - P$ belongs to \mathbb{G} , too; (2) if both polynomials P_1 and P_2 belong to \mathbb{G} , then their product P_1P_2 belongs to \mathbb{G} , too.

Remark 3. *The number of multiplications in the above procedure for generating a polynomial from the set \mathbb{G} (see Step 3 in the procedure below) is one of the parameters of our scheme; denote it by r .*

Note that the set \mathbb{G} consists of polynomials P such that $P(x_1, \dots, x_n) = 0$ or 1 for any Boolean n -tuple (x_1, \dots, x_n) . This easily follows by induction from the above recursive definition of the set \mathbb{G} . In other words, any polynomial from \mathbb{G} induces an n -variable Boolean function and, conversely, any n -variable Boolean function is induced by a polynomial from \mathbb{G} .

Based on this description, we suggest the following procedure for sampling a polynomial, depending on variables from a subset X of the set of variables, from the set \mathbb{G} . We emphasize again that in our scheme, this is done in the offline phase.

- (1) Select a random monomial as in the previous Section 3.1, except that the degree d should be really small, between 1 and 2. Denote this monomial by M .
- (2) With probability $\frac{1}{2}$, select between M and $1 - M$. Denote the result by M' .
- (3) Select, uniformly at random, a variable x_i not from the subset X of variables. Then, with probability $\frac{1}{2}$, multiply M' by either x_i or $1 - x_i$.
- (4) Repeat steps (2) through (3) r times for some small r (one of the parameters of the scheme).

3.3. Converting $H(m)$ to a polynomial. We suggest using a hash function H from the SHA-2 family, specifically SHA-256. We assume the security properties of SHA-256, including collision resistance and preimage resistance. Below is an ad hoc procedure for converting a hash $H(m)$ to a polynomial. We assume there is a standard way to convert $H(m)$ to a bit string of length 256.

Let B be a bit string of length 256. We will convert B to a polynomial from the factor algebra of $K = \mathbb{Z}[x_1, \dots, x_{n+1}]$ by the ideal generated by all polynomials of the form $(x_i^2 - x_i)$, $i = 1, \dots, n + 1$. We note that this process is deterministic.

(1) Split 256 bits in 32 8-bit blocks. The 5 leftmost bits will be responsible for a coefficient of the corresponding monomial, while the 3 rightmost bits will be responsible for a collection of variables x_i in the monomial.

(2) After Step (1), we have 32 3-bit blocks corresponding to monomials of degree 3 that we now have to populate with 3 variables each. Enumerate 96 bits in these 32 3-bit blocks by $x_1, \dots, x_{32}, x_1, \dots, x_{32}, x_1, \dots, x_{32}$ (in this order, going left to right). Now each 3-bit block is converted to a monomial that is a product of x_i corresponding to the places in the bit string where the bit is “1”. In particular, each monomial will be of degree at most 3.

(3) Now we have to use 5 remaining bits in each 8-bit block to obtain an integer coefficient for each monomial of degree ≤ 3 obtained at Step 2. This is done as follows. First, we compute the sum of these 5 bits. Then, we reduce it modulo 3. If the result is 0, then the coefficient is 0. If the result is 1, then the coefficient is 1. If the result is 2, then the coefficient is -1.

(4) Combine all monomials and coefficient obtained at Step (2) into a polynomial.

4. GENERATING AN AUTOMORPHISM φ

An automorphism φ is generated in the offline phase, as follows.

Recall that the set \mathbb{G} consists of polynomials P such that $P(x_1, \dots, x_n) = 0$ or 1 for any Boolean n -tuple (x_1, \dots, x_n) , see Section 3.2.

Then we have:

Proposition 1. *Let $h = h(x_1, \dots, x_n)$ be a polynomial from the set \mathbb{G} . Suppose h does not depend on x_k . Let α be the map that takes x_k to $x_k + h - 2x_k \cdot h$ and fixes all other variables. Then:*

(a) *α defines an automorphism of $B(K)$, the factor algebra of the algebra $\mathbb{Z}[x_1, \dots, x_n]$ by the ideal generated by all polynomials of the form $(x_i^2 - x_i)$, $i = 1, \dots, n$. Denote this automorphism also by α .*

(b) *The group of automorphisms of $B(K)$ is generated by all automorphisms as in part (a) and is isomorphic to the group of permutations of the vertices of the n -dimensional Boolean cube.*

(c) *For any polynomial P from $\mathbb{Z}[x_1, \dots, x_n]$, the number of positive values of P on Boolean tuples (x_1, \dots, x_n) equals that of $\alpha(P)$.*

For the proof of Proposition 1, see the Appendix.

4.1. Generating triangular automorphisms. Our (private) automorphism φ will be a composition of “triangular” automorphisms and permutations on the set of variables. Below is how we generate an “upper triangular” automorphism α .

(1) Let $k = 1$.

(2) With probability $\frac{1}{2}$, either take x_k to itself or take x_k to $x_k + h(x_1, \dots, x_n) - 2x_k \cdot h(x_1, \dots, x_n)$, where $h(x_1, \dots, x_n)$ is a random t -sparse polynomial from the set \mathbb{G} not depending on any x_j with $j \leq k$ (see Section 3.2). Fix all other variables.

(3) If $k < n$, increase k by 1 and go to Step (2). Otherwise, stop.

Generating a “lower triangular” automorphism β is similar:

(1) Let $k = n$.

(2) With probability $\frac{1}{2}$, either take x_k to itself or take x_k to $x_k + h(x_1, \dots, x_n) - 2x_k \cdot h(x_1, \dots, x_n)$, where $h(x_1, \dots, x_n)$ is a random t -sparse polynomial from the set \mathbb{G} not depending on x_j with $j \geq k$. Fix all other variables.

(3) If $k > 1$, decrease k by 1 and go to Step (2). Otherwise, stop.

4.2. Generating φ as a composition of triangular automorphisms and permutations. Having generated an upper triangular automorphism α and a lower triangular automorphism β , we generate our private automorphism φ as a composition $\alpha\beta\pi$, where π is a random permutation on the set of variables. Here α is applied first, followed by β , followed by π .

At the end of the whole procedure, we will have n polynomials $y_i = \varphi(x_i)$ that define the automorphism φ .

5. SUGGESTED PARAMETERS

For the hash function H , we suggest SHA-256.

For the number n of variables, we suggest $n = 31$.

For the number t of monomials in t -sparse polynomials, we suggest $t = 3$.

For the bound b on the degree of monomials in t -sparse polynomials, we suggest $b = 3$.

For the degree d of the monomial M in the procedure for generating a polynomial from the set \mathbb{G} (Section 3.2), we suggest $d = 2$.

For the number r of the number of multiplications in the procedure for generating a polynomial from the set \mathbb{G} (Section 3.2), we suggest $r = 1$.

For the number of trials in Monte Carlo method for counting positive values of a polynomial on Boolean tuples, we suggest 1,000.

6. PERFORMANCE AND SIGNATURE SIZE

For our computer simulations, we used Apple MacBook Pro, M1 CPU (8 Cores), 16 GB RAM computer. Julia code is available, see [2].

With the suggested parameters, signature verification takes under 0.2 sec on average, which is not bad, but the polynomial $\varphi(Q)$ (the signature) is rather large, about 4 Kb on average.

The size of the private key (the automorphism φ) is about 1.5 Kb, and the size of the public key is about 15 Kb.

We note that we have measured the size of a signature, as well as the size of private/public keys, as follows. We have counted the total number of variables that occurred in relevant polynomial(s) and multiplied that number by 5, the number of bits sufficient to describe the index of any variable (except x_{32}). To that, we added the number of monomials times 3 (the average number of bits needed to describe a coefficient at a monomial in our construction(s)).

As usual, there is a trade-off between the size of the private key φ and its security. The size of φ can be reduced to just a few hundred monomials, but then security becomes a concern since some of $\varphi(x_i)$ may be possible to recover more or less by inspection of the public pairs $(P_i, \varphi(P_i))$.

In the table below, we have summarized performance data for most reasonable (in our opinion) parameter sets. Most columns are self-explanatory; the last column shows memory usage during verification.

Performance metrics for various parameter values								
# monomials in P_i	max degree of P_i	max degree of monomials M	parameter m	verification time (sec)	signature size (Kbytes)	public key size (Kbytes)	private key size (Kbytes)	memory usage (Mbytes)
3	3	1	1	0.12	4.3	17.5	1.2	5.7
3	3	2	1	0.12	3.7	12.6	1.6	5.7
3	4	1	1	0.25	4.3	25	1.2	5.8
4	3	1	1	1.6	4.2	38	1.25	7.1
5	3	1	1	3.2	6	46	1.3	8.2
3	3	1	2	0.7	20	56	5.5	6.3

6.1. Accuracy of the Monte Carlo method. We have run numerous computer simulations to estimate the probability of a “false positive” result, in particular accepting a forged signature from somebody who knows only some of $\varphi(x_i)$. In our experiments, the difference between the number of positive values of u and u' for a u' obtained by using a wrong private key φ was always above 9%. Recall that the threshold difference for accepting a signature in our scheme is 3%.

“False negative” results (i.e., rejecting a valid signature because the difference was more than 3%) occurred in 1 out of 1000 experiments. Increasing the number of trials in the Monte Carlo method obviously reduces the probability of false negative as well as false positive results. With 10,000 trials (instead of suggested 1000), we did not detect any false negative or false positive results. However, the verification time then increases to 1.2 sec on average.

7. WHAT IS THE HARD PROBLEM HERE?

Recall that the candidate one-way function that we use in our scheme takes a private polynomial automorphism φ as the input and outputs $\varphi(P)$ for a public multivariate polynomial P . Thus, the (allegedly) hard problem here is: given a public pair (or several pairs) $(P, \varphi(P))$, recover φ . We note that such a φ does not have to be unique, although most of the time it is.

The problem of recovering φ from a pair $(P, \varphi(P))$, as well as the relevant decision problem to find out whether or not, for a given pair of polynomials (P, Q) , there is an automorphism that takes P to Q , was successfully addressed only for two-variable polynomials [3]. For polynomials in more than two variables the problem is unapproachable at this time, and there are no even partial results in this direction.

Of course, in a cryptographic context one is typically looking not for general theoretical results, but rather for practical ad hoc, often non-deterministic, attacks. The most straightforward non-deterministic attack that comes to mind here is as follows. Recall that monomials in the polynomial P have low degree (bounded by 3). Thus, given a monomial, say, $x_1x_2x_3$ in the polynomial P , one can try to replace each x_i by a hypothetical $\varphi(x_i)$ of the form $\sum(c_ix_i + c_{ij}x_ix_j + c_{ijk}x_ix_jx_k)$, with indeterminate coefficients c_i, c_{ij}, c_{ijk} . Given that φ is “sparse”, this may yield a number of equations in the indeterminate coefficients that is not huge. However, these equations will include not just linear

equations, but also equations of degree 2 and 3 (since $\varphi(x_1x_2x_3) = \varphi(x_1)\varphi(x_2)\varphi(x_3)$), and given a large number (hundreds) of unknowns, there is no computationally feasible way known to solve such a system.

In the next Section 8, we offer a “linearization” of this attack where all equations become linear, at the expense of making the number of unknowns and the number of equations very large.

8. LINEAR ALGEBRA ATTACK

One can attempt to recover the private automorphism φ from the public pairs $(P_i, \varphi(P_i))$ by using linear algebra, more specifically by trying to replace φ by a linear transformation of the linear space of monomials involved in P_i and in the polynomials $\varphi(x_i)$. The latter polynomials are not known to the adversary, but at least the degrees of monomials in those polynomials can be bounded based on the public polynomials $\varphi(P_i)$.

Let us compute the dimension of the linear space of monomials of degree at most 27 in 31 variables. This is because a polynomial P_i has monomials of degree at most 3, and in the polynomials $\varphi(x_i)$ there can be monomials of degree up to 9 (with the suggested parameters), so in $\varphi(P_i)$ there can be monomials of degree up to 27.

By a well-known formula of counting combinations with repetitions, the number of monomials of degree at most 27 in 31 variables is equal to $\binom{57}{30} \approx 1.4 \cdot 10^{16} > 2^{53}$. This is how many variables the attacker will have should (s)he use a linear algebra attack. The number of equations will be about triple of this number.

Solving a system of linear equations with that many variables and equations would require more than $2^{53 \cdot 2.3} \approx 2^{122}$ arithmetic operations, according to our understanding of the state-of-the-art in solving systems of linear equations.

We note that increasing the number of variables in the polynomial algebra will not seriously affect efficiency as long as the bound on the degrees of monomials remains the same. At the same time, the more variables the less feasible the linear algebra attack is.

9. SECURITY CLAIMS

The linear algebra “brute force” attack amounts to solving a system of linear equations (over \mathbb{Z}) with about 2^{53} variables and at least as many equations.

There could be ad hoc attacks on the public key aiming at recovering some of the $\varphi(x_i)$, but recovering only some of $\varphi(x_i)$ does not make the probability of passing verification non-negligible, according to our computer simulations.

We have not been able to come up with any ideas of forgery without getting a hold of the private key.

As for quantum security, it appears that no known quantum algorithms are applicable in our situation since there are no abelian (semi)groups in play here that could be exploited by one of the known quantum attacks.

10. CONCLUSION: ADVANTAGES AND LIMITATIONS OF THE SCHEME

10.1. Advantages.

1. A novel mathematical idea used for signature verification.
2. Efficiency of the signature verification (under 0.2 sec on average).

10.2. Limitations.

1. The main limitation is the size of the public key (about 15 Kbytes with suggested parameters).

The private key (the automorphism φ) is not too small either, about 1.5 Kbytes on average. There is a trade-off between the size of φ and its security. The size of φ can be, in principle, reduced to just a few hundred monomials, but then security becomes a concern since some parts of $\varphi(x_i)$ may be possible to recover more or less by inspection of the public pairs $(P_i, \varphi(P_i))$.

The signature size is about 4 Kb on average, which is decent but not record-breaking.

2. Another limitation is that using non-deterministic methods, such as a Monte Carlo type method, may result in errors, more specifically in false negative or even false positive results of the signature verification. So far, with suggested parameters, false positive results did not occur in our computer simulations, and false negative results occurred in about 1 out of 1000 experiments. (“False negative” means rejecting a valid signature.)

Increasing the number of trials in the Monte Carlo method obviously reduces the probability of false negative as well as false positive results. With 10,000 trials (instead of suggested 1000), we did not detect any false negative or false positive results. However, there is a trade-off with the verification time: it will then increase to 1.2 sec on average.

REFERENCES

- [1] D. Grigoriev, M. Karpinski, *An approximation algorithm for the number of zeroes of arbitrary polynomials over $GF[q]$* - Proc. 32 IEEE Symp. FOCS, 1991, pp. 662–669.
- [2] Julia code for the BASS, https://drive.google.com/file/d/1z3RWV9SRhSAxbB0tTXuw_HEIZhNFyB3a/view
- [3] L. Makar-Limanov, V. Shpilrain and J.-T. Yu, *Equivalence of polynomials under automorphisms of $K[x, y]$* , J. Pure Appl. Algebra **209** (2007), 71–78.
- [4] T. T. Moh, *A public key system with signature and master key functions*, Comm. Algebra **27** (1999), 2207–2222.
- [5] NIST: Post-Quantum Cryptography: Digital Signature Schemes, <https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/call-for-proposals-dig-sig-sept-2022.pdf>
- [6] P. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comput. **26** (1997), 1484–1509.
- [7] L. Valiant, *Complexity of computing the permanent*, Theor. Comput. Sci. **8** (1979), 189–201.

APPENDIX

Here we give a proof of Proposition 1.

Proposition 1. Let $h = h(x_1, \dots, x_n)$ be a polynomial from the set \mathbb{G} . Suppose h does not depend on x_k . Let α be the map that takes x_k to $x_k + h - 2x_k \cdot h$ and fixes all other variables. Then:

(a) α defines an automorphism of $B(K)$, the factor algebra of the algebra $\mathbb{Z}[x_1, \dots, x_n]$ by the ideal generated by all polynomials of the form $(x_i^2 - x_i)$, $i = 1, \dots, n$. Denote this automorphism also by α .

(b) The group of automorphisms of $B(K)$ is generated by all automorphisms as in part (a) and is isomorphic to the group of permutations of the vertices of the n -dimensional Boolean cube.

(c) For any polynomial P from $\mathbb{Z}[x_1, \dots, x_n]$, the number of positive values of P on Boolean tuples (x_1, \dots, x_n) equals that of $\alpha(P)$.

Proof. (a) Let B^n denote the Boolean n -cube, i.e., the n -dimensional cube whose vertices are Boolean n -tuples. The map α leaves the set of vertices of B^n invariant. Indeed, α fixes all x_i except x_k , and it is straightforward to see that if $x_k = 0$, then $\alpha(x_k) = h(x_1, \dots, x_n)$, and if $x_k = 1$, then $\alpha(x_k) = 1 - h(x_1, \dots, x_n)$. Since on any Boolean n -tuple (x_1, \dots, x_n) , one has $h(x_1, \dots, x_n) = 0$ or 1 (see Section 3.2), we see that α is a bijection of the set of vertices of B^n onto itself.

Next, observe that for any polynomial h from the set \mathbb{G} , one has $h^2 = h$ modulo the ideal generated by all polynomials of the form $(x_i^2 - x_i)$; this easily follows from the inductive procedure of constructing polynomials h , see Section 3.2. Therefore, α leaves the ideal generated by all $(x_i^2 - x_i)$ invariant since α takes x_i to $x_i + h - 2x_i \cdot h$, and then $\alpha(x_i^2 - x_i) = (x_i + h - 2x_i \cdot h)^2 - (x_i + h - 2x_i \cdot h) = (x_i^2 - x_i) + (h^2 - h) + 2x_i h - 4x_i^2 h - 4x_i h^2 + 4x_i^2 h^2 + 2x_i h = (x_i^2 - x_i) + (h^2 - h) + 4h(x_i - x_i^2) + 4h^2(x_i^2 - x_i)$.

(b) Consider the automorphism α again. Fix a particular Boolean n -tuple (x_1, \dots, x_n) . Suppose that $h(x_1, \dots, x_n) = 1$. Suppose $x_k = 0$ in this tuple. Then α takes this tuple to the tuple where all x_i , except x_k , are the same as before, and $x_k = 1$, i.e., just one of the coordinates in the tuple was flipped. Therefore, an appropriate composition of different α (with different x_k) can map any given Boolean n -tuple to any other Boolean n -tuple.

(c) This follows immediately from the argument in the proof of part (a). More specifically, since the set of vertices of B^n is invariant under α , there is a bijection between the sets of values of P and $\alpha(P)$ on Boolean n -tuples. □

CNRS, MATHÉMATIQUES, UNIVERSITÉ DE LILLE, 59655, VILLENEUVE D'ASCQ, FRANCE

Email address: Dmitry.Grigoryev@univ-lille.fr

DEPARTMENT OF COMPUTER SCIENCE, CUNY GRADUATE CENTER, 365 5TH AVE, NEW YORK, NY 10016

Email address: i.ilmer@icloud.com

DEPARTMENT OF MATHEMATICS, QUEENS COLLEGE, CITY UNIVERSITY OF NEW YORK, QUEENS, NY 11367

Email address: alexey.ovchinnikov@qc.cuny.edu

DEPARTMENT OF MATHEMATICS, THE CITY COLLEGE OF NEW YORK, NEW YORK, NY 10031

Email address: shpilrain@yahoo.com