Exploiting SNOVA's Structure in the Wedge Product Attack

Maxime Bros[1,2], Thai Hung Le[3,4], Jacob Lichtinger[5], Brice Minaud[4], Ray Perlner[5], Daniel Smith-Tone[5,6], and Cristian Valenzuela[6]

[1] Associate, National Institute of Standards and Technology (NIST), Maryland, USA
maxime.bros@nist.gov
[2] Contractor, Izum Inc., Potomac, Maryland, USA
[3] LTCI, Telecom Paris, Institut Polytechnique de Paris, France
[4] École normale supérieure, PSL University, CNRS, Inria, France
{hung.le,brice.minaud}@ens.fr
[5] National Institute of Standards and Technology (NIST), Maryland, USA
{jacob.lichtinger,ray.perlner,daniel.smith}@nist.gov
[6] University of Louisville, Louisville KY, USA
{dcsmit11,cristian.valenzuela}@louisville.edu

Abstract. Post-quantum cryptography (PQC) aims to develop cryptographic schemes secure against quantum adversaries. One promising class of digital signature schemes is based on multivariate quadratic equations, where Unbalanced Oil and Vinegar (UOV) is a leading example. UOV has been extensively studied since its introduction by Kipnis, Patarin, and Goubin at Eurocrypt 1999, and it has remained secure. The signature sizes the scheme allows are quite small; however the key sizes are very large, making the scheme less attractive for some common protocols. To remediate this deficiency, some schemes—such as MAYO, QR-UOV, and SNOVA—add a structure to reduce the size of the public key. These multivariate schemes are candidates that made it to the third round of the National Institute of Standards and Technology's Call for Additional Digital Signature Schemes for the Post-Quantum Cryptography Standardization Process.

In this work, we revisit a recently proposed algebraic attack by Ran at Eurocrypt 2026 on UOV and extend this approach to a new attack on SNOVA by exploiting its block-ring structure. In addition to improving the attack complexity, the exploitation of the block-ring structure rules out spurious solutions, which prevents the generic version of Ran's attack from applying to SNOVA. This attack breaks 6 of the 11 second-round SNOVA parameter sets and improves on the previous best result for an additional 2 sets; it is significantly more effective against larger ℓ in comparison to several earlier attacks. For example, for SNOVA-V with parameters $(v, o, \ell) = (29, 6, 5)$, the estimated security drops to 181 bits, compared to 310 bits for the previous best known attack.

Keywords: Post-Quantum Cryptography · UOV · SNOVA · Algebraic Attack.

1 Introduction

In 2016, the National Institute of Standards and Technology (NIST) began the process for standardizing quantum-resistant public-key cryptography. After selecting some algorithms to standardize, NIST opened an additional standardization process [34] to increase the variety of digital signature standards available. A significant number of submissions belong to the family of multivariate public-key schemes, with many based on the Unbalanced Oil and Vinegar (UOV) construction [29]. Among those which have advanced to the third round, only UOV-based schemes [8,9,21,41] have short signatures; however, another multivariate scheme MQOM [5] has also progressed, albeit following the MPC-in-the-Head paradigm and without short signatures.

Identifying schemes with short signatures was identified as a specific secondary goal in the NIST call for proposals [34]. However, most post-quantum candidates have significantly larger signatures than those of widely-used quantum-vulnerable schemes such as ECDSA. Among the candidates that advanced to the third round, only two broad families offer signature sizes of roughly 200 bytes or less: isogeny-based schemes and multivariate schemes. Isogeny-based schemes [1] are appealing for their combination of short signatures, compact public keys, and small expanded secret keys; however, they are significantly slower than UOV-based multivariate schemes, despite notable performance improvements. Moreover, their long-term security remains less understood, as highlighted by the recent attacks that broke SIDH and SIKE [13,32,37]. This uncertainty has motivated continued interest in fast and well-studied alternatives such as UOV-based designs within the multivariate family.

Unbalanced Oil and Vinegar. By comparison, UOV is a time-tested construction. Despite extensive cryptanalysis since its introduction in 1999—and renewed analysis using modern techniques following its submission to the NIST competition—its security has remained robust. In addition to its long-standing security record, UOV offers excellent performance, with some of the fastest signing and verification speeds among post-quantum signature schemes.

UOV is a hash-and-sign multivariate signature scheme defined over a field $\mathbb{F}$ with $n = v + o$ variables and $m = o$ quadratic equations. Its trapdoor structure partitions the variables into *vinegar* (v) and *oil* (o) parts, with each public polynomial quadratic in all variables except that oil–oil terms are absent. Knowledge of the secret affine transformation T and the central map F enables the signer to efficiently invert the public map $P = F \circ T$. Without this extra information, the inversion reduces to solving a multivariate quadratic (MQ) system that is conjectured to be indistinguishable from random and computationally hard.

A significant drawback to the UOV scheme is its large public key size, which typically ranges in the dozens of kilobytes for compressed-key variants for NIST category 1 parameters. To address this issue, structured variants of UOV have been introduced, such as MAYO [8], QR-UOV [21], and SNOVA [41]. Similarly to structured lattices and codes, this additional structure allows for more efficient and compact schemes, but careful analysis is required to ensure that security is

not weakened in the process. Previous notable attempts to reduce the public key size of UOV include Rainbow [15], LUOV [10], and VOX [35]; however, these schemes all suffered severe attacks [6,16,19] during the NIST PQC Standardization process.

The SNOVA scheme. SNOVA is a third-round candidate in the ongoing NIST standardization process for additional post-quantum signatures, noted for its significant public-key size reduction relative to UOV. For example, the smallest public-key size achieved by the submitted parameters of SNOVA is 1016 bytes (with a signature size of 248 bytes). SNOVA can be viewed as an augmentation of UOV with additional structure in two different ways: it embeds a MAYO-like "whipping" structure; and it uses a specific block-ring structure, similar to the one used in QR-UOV. Beullens [7] showed that, in the initial design, the whipping component admits rank-based attacks when the emulsifier matrices allow low-rank linear combinations. Following independent approaches, Cabarcas *et al.* [12] and Furue *et al.* [22] exploited the block-ring/group-action structure to accelerate polynomial-system solving for relevant SNOVA instances via extension-field changes of variables yielding ideals stable under a cyclic diagonal action and XL-style techniques leveraging multihomogeneity, leading to cheaper forgeries than earlier estimates in some settings. Additional reductions relate SNOVA to UOV over modified parameters: Ikematsu *et al.* [24] and Li and Ding [31] convert SNOVA to UOV with parameters $(q, \ell v, \ell o)$, improving intersection attacks; Nakamura *et al.* [33] reduces SNOVA to a smaller UOV system over $\mathbb{F}_{q^\ell}$ while Furue and Ikematsu [20] look at the truncated polynomial ring and showed that for $\ell = 2$ SNOVA parameters fail to meet NIST security targets. Nevertheless, these attacks appear to become less effective as ℓ grows, while larger ℓ underlies SNOVA's advantage over other UOV-like constructions.

Our Contributions. In this work, we build on the wedge attack against generic UOV systems by Ran [36], which attacks UOV by linearly solving for an exterior space corresponding to the annihilator of the oil space. We identify and exploit the block-ring structure intrinsic to SNOVA's *mini-UOV* map (using the terminology of [2]), and this insight allows us to reduce the number of variables in the linear system attacking SNOVA from $\binom{\ell(v+o)}{\ell v}$ to $\binom{v+o}{v}^\ell$. Furthermore, the algebraic structure of SNOVA also enables us to derive an accurate formula for the rank of the linear system. In consequence, our attack breaks 6 of the 11 second-round parameter sets and improves on the previous best results for an additional 2 sets. For example, for SNOVA-V with parameters $(v, o, \ell) = (29, 6, 5)$, our attack reduces the security to 181 bits, compared to 310 bits under the previous best known attack. A full breakdown of attack complexities can be found in

Table 1 (Section 4.7).The full implementation in Sage [38] and Magma[7] [11] of our attack can be found at:

https://gitlab.di.ens.fr/hle/snova_wedge_attack .

Related Work. The SNOVA scheme was presented in the introduction, and we focus here on other works building on the wedge attack. The wedge attack was introduced by Ran at Eurocrypt 2026 [36], after an initial version at the rump session of Eurocrypt 2025. Ran's attack applies to UOV and MAYO, but not QR-UOV and SNOVA. Other works have already built on the wedge attack, including [27,39] and [30] which is a preliminary version of this work.

2 Preliminaries

2.1 Notation

Let q be a fixed power of 2 and $\mathbb{F}_q$ the finite field with q elements. In this article, $q = 16$. For $\ell \in \mathbb{N}$, $\mathbb{F}_{q^\ell}$ is its degree-ℓ extension. To denote an arbitrary field of characteristic 2, $\mathbb{F}$ is used.

The set of matrices with a rows, b columns, and entries in $\mathbb{F}$ is denoted $\mathbb{F}^{a\times b}$. The general linear group of non singular matrices in $\mathbb{F}^{a\times a}$ is denoted $\mathrm{GL}_a(\mathbb{F})$. Matrices and vectors are written in boldface font: $\mathbf{M}$ and $\mathbf{v}$, respectively. The zero matrix $\mathbf{0}$ is also written in boldface.) Additionally, vectors are column vectors by default.

The canonical basis for $\mathbb{F}^a$ is given by $\{\mathbf{e}_i\}_{i=1}^a$ where the ith coordinate is 1 and the rest are 0. We also denote $\mathbb{F}^{*a}$, with the canonical basis $\{\mathbf{e}_i^*\}_{i=1}^a$, to be the dual vector space of $\mathbb{F}^a$. For some $k \in \mathbb{N}$, let $[k] = \{1, \ldots, k\}$, and let $\mathfrak{S}_k$ be the group of permutations on the set $[k]$. For $\mathbf{M} \in \mathbb{F}^{a\times b}$ and $J \subset [a]$ of size $|J| = k$, let $\mathbf{M}_{J*} \in \mathbb{F}^{k\times b}$ be the submatrix of $\mathbf{M}$ whose rows are indexed by J. Additionally, if $b \leq a$ and $|J| = b$, $\det(\mathbf{M}_{J*})$ is a maximal minor of $\mathbf{M}$. For $1 \leq j \leq v$, the jth column of $\mathbf{M}$ is denoted $\mathbf{m}_j$. For a matrix (or vector) $\mathbf{B} \in \mathbb{F}^{a\times b}$, $\mathbf{B}^{\otimes c} \in \mathbb{F}^{ac\times bc}$ has c copies of $\mathbf{B}$ along the diagonal and 0 everywhere else.

A vector space $\mathcal{V}$ over $\mathbb{F}$ is denoted with the math calligraphic font. Often, the same letter is used for the matrix $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$ representing a basis for $\mathcal{V}$ as column vectors. If $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{F}^a$, the subspace of $\mathbb{F}^a$ generated by the $\mathbb{F}$-linear span of these vectors is denoted by $\langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle$.

2.2 Tensor Algebra, Exterior Algebra, and Symmetric Algebra

Tensor Algebra. Let $\mathcal{V}$ be a vector space over $\mathbb{F}$, $T^0(\mathcal{V}) = \mathbb{F}$, and for $k > 0$, $T^k(\mathcal{V})$ be the vector space over $\mathbb{F}$ generated by $\{\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_k : \mathbf{x}_i \in \mathcal{V}\}$ that

[7] Certain commercial software is identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the software identified is necessarily the best available for the purpose.

respects the multilinear structure. An element of $T^k(\mathcal{V})$ is called a k-tensor. The graded algebra $T(\mathcal{V}) = \bigoplus_{k\in\mathbb{N}} T^k(\mathcal{V})$ is the tensor algebra. For $\mathcal{V}$ and $\mathcal{W}$ being two vector spaces over $\mathbb{F}$, the tensor product $\mathcal{V} \otimes \mathcal{W}$ is the vector space over $\mathbb{F}$ generated by $\{\mathbf{x} \otimes \mathbf{y} : \mathbf{x} \in \mathcal{V},\ \mathbf{y} \in \mathcal{W}\}$ that respects the bilinear structure.

Exterior Algebra. Let I be the ideal of $T(\mathcal{V})$ generated by elements $\mathbf{x} \otimes \mathbf{x}$ for $\mathbf{x} \in \mathcal{V}$. The exterior algebra over $\mathcal{V}$ is a quotient algebra of the tensor algebra $T(\mathcal{V})$. Formally, it is defined as:

$$\bigwedge(\mathcal{V}) = T(\mathcal{V})/I \quad \text{and} \quad \bigwedge^k(\mathcal{V}) = T^k(\mathcal{V})/(I \cap T^k(\mathcal{V})).$$

The product inherited from the tensor algebra via this quotient is called the *wedge product*, written as $\wedge$. This product satisfies

$$\mathbf{x}_{\sigma(1)} \wedge \cdots \wedge \mathbf{x}_{\sigma(k)} = \operatorname{sgn}(\sigma)\, \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k,$$

for every $\sigma \in \mathfrak{S}_k$, where $\operatorname{sgn}(\sigma)$ is the sign of the permutation. Since $\mathbb{F}$ has characteristic 2, the sign can be omitted. An element of $\bigwedge^k(\mathcal{V})$ is called a *k-form*. For the canonical basis $\{\mathbf{e}_i\}_{i=1}^a$ of $\mathbb{F}^a$ and $J \subset [a]$, let $\mathbf{e}_J = \bigwedge_{j\in J} \mathbf{e}_j$. For $k \leq a$, the elements $\{\mathbf{e}_J : |J| = k\}$ form a basis of $\bigwedge^k(\mathcal{V})$.

Symmetric Algebra. Let J be the ideal of $T(\mathcal{V})$ generated by elements $\mathbf{x} \otimes \mathbf{y} - \mathbf{y} \otimes \mathbf{x}$ for $\mathbf{x}, \mathbf{y} \in \mathcal{V}$. Similar to the exterior algebra, the symmetric algebra over $\mathcal{V}$ is a quotient algebra of the tensor algebra $T(\mathcal{V})$. Formally, it is defined as:

$$\operatorname{Sym}(\mathcal{V}) = T(\mathcal{V})/J \quad \text{and} \quad \operatorname{Sym}^k(\mathcal{V}) = T^k(\mathcal{V})/(J \cap T^k(\mathcal{V})).$$

For a vector space $\mathcal{V} \cong \mathbb{F}^m$ having a canonical basis $\{\mathbf{e}_1, ..., \mathbf{e}_m\}$, we have $\operatorname{Sym}(\mathcal{V}) \cong \mathbb{F}[\mathbf{e}_1, \ldots, \mathbf{e}_m]$.

2.3 Description of UOV

Let $n, m \in \mathbb{N}$ and $\{p_i\}_{i=1}^m$ be quadratic maps on $\mathbb{F}^n$. Then for all $\mathbf{x} \in \mathbb{F}^n$, there are $\mathbf{P}_i \in \mathbb{F}^{n\times n}$ such that $p_i(\mathbf{x}) = \mathbf{x}^\intercal \mathbf{P}_i \mathbf{x}$. The quadratic maps are called UOV maps if there exists an oil space $\mathcal{O} \subset \mathbb{F}^n$ of dimension $o = m$ such that for each $\mathbf{x} \in \mathcal{O}$ and $1 \leq i \leq m$, $p_i(\mathbf{x}) = 0$. Let the *annihilator space* be defined as

$$\mathcal{A} := \{\, \mathbf{a} \in \mathbb{F}^n : \mathbf{a}^\intercal \mathbf{x} = 0 \text{ for all } \mathbf{x} \in \mathcal{O} \,\},$$

which has dimension $v = n - o$. In [36], $\mathcal{A}$ was called the "vinegar space." We decided to not use this term to avoid conflating $\mathcal{A}$ with the span of the vinegar variables.

Without loss of generality, there exists an $\mathbf{X} \in \mathbb{F}^{v \times o}$ such that $\mathcal{O}$ and $\mathcal{A}$ are given by the column spaces of

$$\mathbf{O} := \begin{bmatrix} \mathbf{X} \\ \mathbf{I}_o \end{bmatrix} \text{ and } \mathbf{A} := \begin{bmatrix} \mathbf{I}_v \\ -\mathbf{X}^\intercal \end{bmatrix},$$

respectively. Since $\mathbb{F}$ has characteristic 2, omitting the negative sign in the definition of $\mathbf{A}$ preserves the property that $\mathbf{A}^\intercal \mathbf{O} = \mathbf{0}$.

2.4 Description of SNOVA

This work recovers the private key from the public key, so only the public key and its trapdoor structure are described below. The comprehensive definition of the scheme can be found in the second-round specification document [41]. SNOVA's public key consists of $m = o$ bilinear forms in $n\ell = (v + o)\ell$ variables, given by matrices $\mathbf{P}_i \in \mathbb{F}_q^{n\ell \times n\ell}$. For $\mathbf{x}, \mathbf{y} \in \mathcal{O}$, a secret (oil) subspace of dimension $o\ell$, the equation $\mathbf{x}^\intercal \mathbf{P}_i \mathbf{y} = 0$ holds. This oil subspace is structured so that it can be expressed as the column space of a matrix $\mathbf{O} \in \mathbb{F}_q^{n\ell \times o\ell}$, where each $\ell \times \ell$ block is restricted to be an element of the field $\mathbb{F}_q[\mathbf{S}]$, where $\mathbf{S}$ is a public symmetric $\ell \times \ell$ matrix with an irreducible characteristic polynomial over $\mathbb{F}_q$. Knowledge of the oil space (with the public key) enables generation of valid signatures.

While each element of $\mathbb{F}_q[\mathbf{S}] \subset \mathbb{F}_q^{\ell \times \ell}$ has ℓ^2 components over $\mathbb{F}_q$, the field $\mathbb{F}_q[\mathbf{S}]$ is a degree ℓ extension of $\mathbb{F}_q$ and is therefore an ℓ-dimensional $\mathbb{F}_q$-vector space. Using this field structure, an $m\ell^2$-dimensional space of bilinear forms can be generated from the original m bilinear forms. This space is spanned by the bilinear forms $\mathbf{P}_{i,a,b} = (\mathbf{S}^a)^{\otimes n} \mathbf{P}_i (\mathbf{S}^b)^{\otimes n}$, where $0 \leq a, b < \ell$ and $1 \leq i \leq m$. These bilinear forms share the same oil space as is noted by [7] and several earlier attacks [24,31]. We additionally define a related space of symmetric bilinear forms spanned by $\mathbf{Q}_{i,a,b} = \mathbf{P}_{i,a,b} + \mathbf{P}_{i,a,b}^\intercal$

3 Generic Wedge Product Attack

This section starts by recalling Ran's wedge attack on UOV [36]. This attack is empowered by special structure exhibited by the public quadratic forms of a UOV instance. We then introduce an alternative method for determining the rank of the main system to be solved in the wedge attack, which will serve as a basis for analyzing our attack against SNOVA.

3.1 Exterior Product Structure of UOV

One may associate with any quadratic form $p_i : \mathbb{F}^n \to \mathbb{F}$ its *polar form* $Q_i : (\mathbb{F}^n)^2 \to \mathbb{F}$ given by $Q_i(\mathbf{x}, \mathbf{y}) = p_i(\mathbf{x} + \mathbf{y}) - p_i(\mathbf{x}) - p_i(\mathbf{y}) + p_i(\mathbf{0})$. Note that the polar form is sometimes called the discrete differential in the literature. In the relevant case of characteristic 2, the polar form is *an alternating bilinear form*

because $Q_i(\mathbf{x},\mathbf{x}) = p_i(2\mathbf{x}) - 2p_i(\mathbf{x}) + p_i(\mathbf{0}) = 2p_i(\mathbf{0}) - 2p_i(\mathbf{x}) = 0$. Each bilinear form can be expressed via the matrix expression

$$Q_i(\mathbf{x},\mathbf{y}) := \mathbf{x}^\intercal \mathbf{Q}_i \mathbf{y},$$

where $\mathbf{Q}_i := \mathbf{P}_i + \mathbf{P}_i^\intercal$ and $\mathbf{P}_i$ is the matrix representation of the quadratic form p_i. More precisely, an alternating bilinear form on $\mathbb{F}^n$ is an element of $\bigwedge^2(\mathbb{F}^n)^*$. We use the canonical coordinate pairing to identify the dual basis $\mathbf{e}_j^*$ with $\mathbf{e}_j$. Thus the matrix $\mathbf{Q}_i$ and the associated 2-form $\widehat{Q}_i$ are two representations of the same alternating form. This property implies that $0 = Q_i(\mathbf{e}_j,\mathbf{e}_j) = \mathbf{e}_j^\intercal \mathbf{Q}_i \mathbf{e}_j = (\mathbf{Q}_i)_{j,j}$, so each Q_i can be represented as a 2-form

$$\widehat{Q}_i := \sum_{1 \le j < k \le n} (\mathbf{Q}_i)_{j,k} \mathbf{e}_j \wedge \mathbf{e}_k. \tag{1}$$

Ran leveraged this generic fact in the context of UOV and revealed that the public quadratic maps can be manipulated with the exterior algebra to derive a new attack [36]. In order to describe the wedge product attack, here is some background information that appeared in [36].

Lemma 1 ([36, Lemma 2]). *For each $i \in [m]$, let Q_i be the polar form of a UOV public coordinate map so that $Q_i(\mathbf{x},\mathbf{y}) = 0$ for all $\mathbf{x},\mathbf{y} \in \mathcal{O}$. Let $\widehat{Q}_i \in \bigwedge^2 \mathbb{F}^n$ be the associated alternating form and let $\mathbf{a}_1,\ldots,\mathbf{a}_v$ be a basis of the annihilator space $\mathcal{A}$ of $\mathcal{O}$. Then there exist $\mathbf{x}_{i,j} \in \mathbb{F}^n$ such that*

$$\widehat{Q}_i = \sum_{j=1}^{v} \mathbf{a}_j \wedge \mathbf{x}_{i,j}.$$

Interestingly, it is also beneficial to consider k-dimensional vector spaces and their corresponding basis matrices in the language of k-forms.

Lemma 2 ([42, Proposition 3.3]). *Let $\mathbf{y}_1,\ldots,\mathbf{y}_k \in \mathbb{F}^a$ and define the matrix*

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \cdots \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} y_{1,1} & \cdots & y_{k,1} \\ \vdots & \ddots & \vdots \\ y_{1,a} & \cdots & y_{k,a} \end{bmatrix}$$

then

$$\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_k = \sum_{\substack{J \subset [a] \\ |J| = k}} \det(\mathbf{Y}_{J*})\, \mathbf{e}_J.$$

Corollary 1 ([36, Corollaries 1 & 2]). *Let $\mathbf{y}_1,\ldots,\mathbf{y}_k \in \mathbb{F}^a$. The quantity $\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_k = 0$ if and only if the vectors are linearly dependent. Otherwise, for any basis $\mathbf{z}_1,\ldots,\mathbf{z}_k$ of $\langle \mathbf{y}_1,\ldots,\mathbf{y}_k \rangle$ there exists a nonzero $a \in \mathbb{F}$ such that the wedge product $\mathbf{z}_1 \wedge \cdots \wedge \mathbf{z}_k = a(\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_k)$.*

The *Plücker embedding* is a standard tool for connecting k-dimensional vector spaces with k-forms. If the vector space $\mathcal{Y}$ has a basis $\mathbf{y}_1, \ldots, \mathbf{y}_k$, the Plücker embedding maps

$$\mathcal{Y} \mapsto \mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_k,$$

which is well-defined up to a scalar as a result of Corollary 1. The symbol $\widehat{\cdot}$ is overloaded to associate the equivalence class of k-forms $\widehat{\mathcal{Y}}$ modulo these scalars with the k-dimensional vector space $\mathcal{Y}$ in a similar manner as the 2-form $\widehat{Q}_i$ is associated with the alternating bilinear form Q_i.

In particular, for the annihilator space $\mathcal{A}$ of $\mathcal{O}$, the following is obtained:

$$\mathcal{A} = \langle \mathbf{a}_1, \ldots, \mathbf{a}_v \rangle \mapsto \widehat{\mathcal{A}} := \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_v,$$

where the $\mathbf{a}_i$ are the columns of $\mathbf{A}$. Note that there is a slight abuse of notation here since $\widehat{\mathcal{A}}$ should be the equivalence class of $\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_v$ modulo scalars. While the attack works for any basis of $\mathcal{A}$, the basis corresponding to $\mathbf{A}$ is used for concreteness and clarity. Since the $\mathbf{a}_i$ are linearly independent, $\widehat{\mathcal{A}}$ is nonzero.

Each $\widehat{Q}_i$ has the form given above, so the following relation is obtained:

$$\widehat{\mathcal{A}} \wedge \widehat{Q}_i = (\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_v) \wedge \sum_{j=1}^{v} \mathbf{a}_j \wedge \mathbf{x}_{i,j} = 0. \tag{2}$$

Because all coefficients $(\mathbf{Q}_i)_{j,k}$ from Equation 1 are known, there are linear equations in the maximal minors of $\widehat{\mathcal{A}}$ using Equation 2 and Lemma 2:

$$\begin{aligned}
0 = \widehat{\mathcal{A}} \wedge \widehat{Q}_i &= \sum_{j<k\in[n]} (\mathbf{Q}_i)_{j,k}\, \widehat{\mathcal{A}} \wedge \mathbf{e}_j \wedge \mathbf{e}_k \\
&= \sum_{j<k\in[n]} (\mathbf{Q}_i)_{j,k} \left(\sum_{\substack{J\subset[n] \\ |J|=v}} \det(\mathbf{A}_{J*})\, \mathbf{e}_J \right) \wedge \mathbf{e}_j \wedge \mathbf{e}_k \\
&= \sum_{\substack{K\subset[n] \\ |K|=v+2}} \left(\sum_{\substack{j<k\in K \\ J=K\setminus\{j,k\}}} (\mathbf{Q}_i)_{j,k} \det(\mathbf{A}_{J*}) \right) \mathbf{e}_K.
\end{aligned}$$

Since $\{\mathbf{e}_K\}_{\substack{K\subset[n] \\ |K|=v+2}}$ forms a basis of $\bigwedge^{v+2} \mathbb{F}^n$, each set K gives an equation:

$$\sum_{\substack{j<k\in K \\ J=K\setminus\{j,k\}}} (\mathbf{Q}_i)_{j,k} \det(\mathbf{A}_{J*}) = 0. \tag{3}$$

Ranging over all $i \in [m]$, a linear system is obtained with $m\binom{n}{v+2}$ equations over $\binom{n}{v}$ variables, i.e., the maximal minors $\det(\mathbf{A}_{J*})$ of $\mathbf{A}$. Since $\mathbf{A}$ contains an identity block, every coordinate of $\mathbf{A}$ outside the identity block corresponds

to one of the maximal minors. Thus, recovering the maximal minors reveals the coordinates of $\mathbf{A}$, which are also the coordinates for the oil space.

The above discussion corresponds to the key recovery attack [36] which begins by solving for the kernel of the following linear map.

Definition 1 (The Map Φ_{UOV} [36, Equation 3]). *Let $\widehat{Q}_1, \ldots, \widehat{Q}_m \in \bigwedge^2 \mathbb{F}^n$ be the 2-forms associated with the a UOV instance. For any v-form α, define the linear map $\Phi_{\mathrm{UOV}} : \bigwedge^v \mathbb{F}^n \to \left(\bigwedge^{v+2} \mathbb{F}^n\right)^m$ by*

$$\Phi_{\mathrm{UOV}}(\alpha) := (\, \alpha \wedge \widehat{Q}_1, \ldots, \alpha \wedge \widehat{Q}_m \,). \tag{4}$$

Due to Equation (2), $\widehat{\mathcal{A}} \in \ker(\Phi_{\mathrm{UOV}})$. Thus, if $\ker(\Phi_{\mathrm{UOV}})$ is 1-dimensional, then $\ker(\Phi_{\mathrm{UOV}}) = \{a\widehat{\mathcal{A}} \mid a \in \mathbb{F}\}$ and the attack succeeds.

3.2 A Conjecture for rank Φ_{UOV}

Observe that not all of these equations in 3 are linearly independent and previous works [36] rely on the following conjecture.

Conjecture 1 ([36], Section 4.2) *With high probability, the number of linear independent equations in such a linear system when $v < \min\{\frac{o-1}{2}\ ,\ 2\}m$ is*

$$\operatorname{rank} \Phi_{\mathrm{UOV}} = \min\left\{ \binom{v+o}{v} - 1\ ,\ \sum_{i=1}^{\lfloor \frac{o}{2} \rfloor} (-1)^{i+1} \binom{m+i-1}{i} \binom{v+o}{v+2i} \right\}. \tag{5}$$

Hence, the linear system is expected to be solvable with a unique solution for those (v, o, m)-UOV instances that satisfy

$$\sum_{i=0}^{\lfloor \frac{o}{2} \rfloor} (-1)^{i} \binom{m+i-1}{i} \binom{v+o}{v+2i} \ \leq\ 1. \tag{6}$$

In [36], Conjecture 1 is reduced to the well-known conjecture on the existence of semi-regular sequences in Boolean algebras. In this section, we present a different perspective on Conjecture 1 by reducing it to the exactness of a cochain complex. In general, the two reductions yield the same formula for the rank of the the linear map Φ_{UOV}. However, the cochain-complex viewpoint will be important later for the analysis of the wedge attack against SNOVA.

Let $\mathcal{F} = \mathcal{A} \oplus \mathcal{B} \cong \mathbb{F}^n$, where $\mathcal{A}$ is the annihilator space with $\dim \mathcal{A} = v$ and $\mathcal{B}$ is a complement of $\mathcal{A}$, so $\dim \mathcal{B} = n - v = o$. For $i \geq 0$, set

$$\mathcal{W}_i := \bigwedge^i \mathcal{A} \wedge \bigwedge^{v+i} \mathcal{F} = \left\langle \alpha \wedge \beta : \alpha \in \bigwedge^i \mathcal{A}, \beta \in \bigwedge^{v+i} \mathcal{F} \right\rangle \subseteq \bigwedge^{v+2i} \mathcal{F}.$$

Since the parameters are coming from an instance of UOV, $o \leq v$; more generally the same argument applies in every degree i for which $o \leq v + i$.

Remark 1. Here, we overload the notation $\wedge$ by defining the wedge product between two subspaces of the same space as the linear span of the wedge products of elements from these two subspaces.

Lemma 3. *For every degree i satisfying $o \leq v + i$, $\mathcal{W}_i = \bigwedge^{v+2i} \mathcal{F}$.*

Proof. The inclusion $\mathcal{W}_i \subseteq \bigwedge^{v+2i} \mathcal{F}$ is immediate. Conversely, using $\mathcal{F} = \mathcal{A} \oplus \mathcal{B}$,

$$\bigwedge^{v+2i} \mathcal{F} = \bigoplus_{\substack{s+t=v+2i \\ 0 \leq t \leq o}} \bigwedge^{s} \mathcal{A} \otimes \bigwedge^{t} \mathcal{B}.$$

For each nonzero summand, $t \leq o \leq v + i$, hence $s = v + 2i - t \geq i$. Thus every decomposable tensor in that summand has at least i factors from $\mathcal{A}$, so it can be written as an element of $\bigwedge^{i} \mathcal{A} \wedge \bigwedge^{v+i} \mathcal{F}$. This proves the reverse inclusion. □

Fix m two-forms $\widehat{Q}_1, \ldots, \widehat{Q}_m \in \mathcal{A} \wedge \mathcal{F} \subseteq \bigwedge^2 \mathcal{F}$. Let $\mathcal{E} = \mathbb{F}^m$ with basis $\mathbf{q}_1, \ldots, \mathbf{q}_m$, and define

$$\mathcal{C}^i := \mathrm{Sym}^i \mathcal{E} \otimes \mathcal{W}_i.$$

which is isomorphic to $\mathcal{W}_i^{\binom{m+i-1}{i}}$. In particular, $\mathcal{C}^0 \cong \mathcal{W}_0$ and $\mathcal{C}^1 \cong \mathcal{W}_1^m$.

Since $\widehat{Q}_j \in \mathcal{A} \wedge \mathcal{F}$, wedging by $\widehat{Q}_j$ maps $\mathcal{W}_i$ into $\mathcal{W}_{i+1}$, and hence the linear map $\delta_i : \mathcal{C}^i \to \mathcal{C}^{i+1}$ defined by

$$\delta_i(\eta \otimes \mu) = \sum_{j=1}^{m} (\mathbf{q}_j \eta) \otimes (\widehat{Q}_j \wedge \mu),$$

are well defined. We can see that δ_0 is the linear map Φ_{UOV} up to the vector space isomorphisms given above.

In the degrees relevant for Conjecture 1, Lemma 3 gives

$$\mathcal{C}^i = \mathrm{Sym}^i \mathcal{E} \otimes \bigwedge^{v+2i} \mathcal{F} \text{ and } \dim \mathcal{C}^i = \binom{m+i-1}{i} \binom{v+o}{v+2i}.$$

Lemma 4. *For every 2-form $\widehat{Q}$ over $\mathbb{F}$, $\widehat{Q} \wedge \widehat{Q} = 0$.*

Proof. For a 1-form x, the exterior algebra directly gives $x \wedge x = 0$. However, it does not imply the same identity for an arbitrary 2-form except over characteristic 2. Let $\widehat{Q} = \sum_{1 \leq i < j \leq n} \mathbf{Q}_{i,j}\, \mathbf{e}_i \wedge \mathbf{e}_j \in \bigwedge^2 \mathbb{F}^n$ where $\mathbf{e}_i$ is the canonical basis for $\mathbb{F}^n$. Then

$$\widehat{Q} \wedge \widehat{Q} = \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < \ell \leq n}} (\mathbf{Q}_{i,j}\, \mathbf{e}_i \wedge \mathbf{e}_j) \wedge (\mathbf{Q}_{k,\ell}\, \mathbf{e}_k \wedge \mathbf{e}_\ell)$$

When $\{i, j\} \neq \{k, \ell\}$, these terms come in pairs. Since $\mathbf{e}_i \wedge \mathbf{e}_j \wedge \mathbf{e}_k \wedge \mathbf{e}_\ell = \mathbf{e}_k \wedge \mathbf{e}_\ell \wedge \mathbf{e}_i \wedge \mathbf{e}_j$, the coefficient of $\mathbf{e}_i \wedge \mathbf{e}_j \wedge \mathbf{e}_k \wedge \mathbf{e}_\ell$ is 0 over characteristic 2. When $\{i, j\} = \{k, \ell\}$, these terms are 0 since $\mathbf{e}_i \wedge \mathbf{e}_i = 0$. □

Using Lemma 4, we can see that for $i \geq 0$ and a basis element $\eta \otimes \mu \in \mathcal{C}^{i-1}$,

$$\begin{aligned}\delta_i \circ \delta_{i-1}(\eta \otimes \mu) &= \sum_{j \neq k} (\mathbf{q}_j \mathbf{q}_k \eta) \otimes (\widehat{Q}_j \wedge \widehat{Q}_k \wedge \mu) \\ &= \sum_{j<k} (2\mathbf{q}_j \mathbf{q}_k \eta) \otimes (\widehat{Q}_j \wedge \widehat{Q}_k \wedge \mu) \\ &= 0.\end{aligned}$$

Hence we obtain a cochain complex

$$0 \longrightarrow \mathcal{C}^0 \xrightarrow{\delta_0} \mathcal{C}^1 \xrightarrow{\delta_1} \cdots \xrightarrow{\delta_{r-1}} \mathcal{C}^r \longrightarrow 0, \tag{7}$$

where $r = \left\lfloor \frac{o}{2} \right\rfloor$, since for $i > r$, the space $\mathcal{C}^i$ is trivial.

We introduce the following conjecture whose validity is supported by our experimental evidence, see Figure 1 and Tables 4 and 5.

Conjecture 2 *With high probability (over the random coins used to sample the UOV public key), the cochain complex of Equation (7) is exact in every positive degree.*

Conjecture 2 provides the desired result:

$$\begin{aligned}\operatorname{rank} \Phi_{\mathrm{UOV}} = \operatorname{rank} \delta_0 &= \sum_{i=1}^{r} (-1)^{i+1} \dim \mathcal{C}^i \\ &= \sum_{i=1}^{\lfloor o/2 \rfloor} (-1)^{i+1} \binom{m+i-1}{i} \binom{v+o}{v+2i}.\end{aligned}$$

This discussion shows that Conjecture 2 implies Conjecture 1. As mentioned earlier, the point of the stronger Conjecture 2 is that it introduces a new viewpoint, based on (co)chain complexes, that will be fruitful for the analysis of SNOVA. It should be noted that Koszul-type cochain complexes are a standard tool in Algebraic Geometry and other areas to analyze relations in polynomial systems, dating back to Hilbert's introduction of (what can be retrospectively viewed as) the Koszul complex [23], exactly for that purpose.

Although it is not strictly needed for the proof, some intuition regarding the use of cochain complexes may be helpful. As also noted earlier, the map δ_0 is equal to Φ_{UOV} (up to composing with isomorphisms at the input and output spaces). At an intuitive level, the next map δ_1 describes linear relations among the equations defining Φ_{UOV}, the next map δ_2 describes linear relations among those relations themselves, and so forth. This type of nested linear relations is typical of syzygies in polynomial systems.

In particular, the elements of $\mathcal{C}^1$ correspond to the relations of the form $\widehat{\mathcal{A}} \wedge \widehat{Q}_i$ where those of $\mathcal{C}^2$ correspond to $\widehat{\mathcal{A}} \wedge \widehat{Q}_i \wedge \widehat{Q}_j$. The relation $\widehat{\mathcal{A}} \wedge \widehat{Q}_i \wedge \widehat{Q}_j$ is within the span of both $\widehat{\mathcal{A}} \wedge \widehat{Q}_i$ and $\widehat{\mathcal{A}} \wedge \widehat{Q}_j$, and extends analogously to the other $\mathcal{C}^i$ for $i > 1$. This causes soundness of the cochain complex as well as the

alternating sum in Conjecture 1. The analysis in this section proves that the cochain complex is valid: the image of each map is included in the kernel of the next map. In other words, we prove that the nested linear relations described by the cochain do exist in Φ_{UOV} (we will proceed similarly for the cochain complex describing Φ_{SNOVA}, in the next section).

Conjecture 2 states that the cochain is exact except at degree 0. This means there are no other linear relationships in the target equation system, beside the ones described by the cochain complex. This holds experimentally for both UOV and SNOVA.

Remark 2. [36,39] also rely on a conjecture of the same nature. This is expected. Indeed, even for well-studied "normal" polynomial systems (not in the exterior algebra), proving the regularity of a random system is a famous open problem (see e.g. [4]). Standard regularity assumptions also amount to assuming that a certain cochain complex is exact. For those reasons, we do not expect that an exactness proof is within reach from current knowledge. The cohomology of the cochain complex in this section seems to relate to that of the Eagon-Northcott complex [17,18]. In short, the Eagon-Northcott complex can be regarded as analyzing relations in the symmetric algebra, using the exterior algebra as an auxiliary. We are doing exactly the opposite. In general, polynomial systems in the exterior algebra are less well-studied.

3.3 Recovering the Annihilator Space from Φ

The v-form $\widehat{\mathcal{A}} \in \bigwedge^v \mathbb{F}_q$ is always in the kernel, due to Equation (2). Therefore $\operatorname{rank} \Phi_{\mathrm{UOV}} \leq \binom{v+o}{v} - 1$. Combining the alternating-sum prediction with this forced upper bound gives precisely the minimum in Equation (5). Condition (6) is the statement that the alternating-sum term is at least $\binom{v+o}{v} - 1$; in that range the expected kernel dimension is 1, so no kernel vector other than the one forced by the UOV structure is expected for a generic instance.

Let E denote the number of linear equations and U the number of variables in the maximal minors modeling of $\ker \Phi_{\mathrm{UOV}}$ via Equation (3). This straightforward construction produces a sparse matrix of size $E \times U$ whose row density is $\binom{v+2}{2}$, and whose rank is given by Conjecture 1. Using the Wiedemann Algorithm [14,28,40] for sparse linear algebra, the cost of finding a nonzero kernel vector is

$$\mathcal{O}\left(\binom{v+2}{2} E\, U\right)$$

field operations in $\mathbb{F}$. Since the computation of the maximal minors of $\mathbf{A}$ allows the recovery of $\mathcal{A}$, this calculation completes the key recovery. Therefore, the wedge-product viewpoint turns the algebraic cryptanalysis of UOV into a single structured linear problem that can be solved by modern sparse-matrix solvers.

4 Wedge Product Attack on SNOVA

In this section, we apply the wedge product attack to SNOVA and demonstrate that the scheme's specific algebraic structure can be exploited with great efficiency. For the concrete SNOVA system in this section, the ground field is denoted by $\mathbb{F}_q$ instead of $\mathbb{F}$ as in the generic discussion. A description of SNOVA was provided in Section 2.4. Recall that the public key of SNOVA can be viewed as an UOV instance where the matrices $\mathbf{P}_i \in \mathbb{F}_q^{n\ell \times n\ell}$ defining the public bilinear forms have dimension $n\ell$ instead of n, and the oil space has dimension $o\ell$ instead of o. Also recall from Section 2.4 that the oil space has a blockwise structure coming from $\mathbb{F}_q[\mathbf{S}]$, and the public key can be scaled by a factor ℓ^2 to yield $m\ell^2$ alternating bilinear forms $\mathbf{Q}_{i,a,b}$, also with a special structure. The crux of this section will be to take advantage as well as analyze the effect of those structures when trying to translate the wedge attack to SNOVA.

We will reuse the same notation for the oil space and its annihilator as in the previous section (as we already did in Section 2.4), although we are now dealing with the structured SNOVA public key instead of the more generic UOV public key. When applied to SNOVA, an attacker attempts to find a v-form $\widehat{\mathcal{A}}$ in the kernel of the following linear map.

Definition 2 (The map Φ_{SNOVA}). *Let $\widehat{Q}_{0,0,0}, \ldots, \widehat{Q}_{m,\ell-1,\ell-1} \in \bigwedge^2 \mathbb{F}_q^n$ be the 2-forms associated with the a SNOVA instance. Define the linear map*

$$\begin{aligned} \Phi_{\text{SNOVA}} : \bigwedge^{\ell v} \mathbb{F}_q^{\ell n} &\to \left(\bigwedge^{\ell v+2} \mathbb{F}_q^{\ell n} \right)^{m\ell^2} \\ \alpha &\mapsto (\, \alpha \wedge \widehat{Q}_{1,0,0}, \ldots, \alpha \wedge \widehat{Q}_{m,\ell-1,\ell-1} \,). \end{aligned}$$

In the above expression, all $m\ell^2$ forms $\widehat{Q}_{i,a,b}$'s for $1 \le i \le m$ and $0 \le a, b < \ell$ are wedged on the right-hand side. We also define a variant with m forms, using only the $\widehat{Q}_{i,0,0}$'s:

$$\begin{aligned} \Phi^{\langle m \rangle}_{\text{SNOVA}} : \bigwedge^{\ell v} \mathbb{F}_q^{\ell n} &\to \left(\bigwedge^{\ell v+2} \mathbb{F}_q^{\ell n} \right)^{m} \\ \alpha &\mapsto (\, \alpha \wedge \widehat{Q}_{1,0,0}, \ldots, \alpha \wedge \widehat{Q}_{m,0,0} \,), \end{aligned}$$

and a variant with $m\ell$ forms, using only the $\widehat{Q}_{i,a,a}$ (i.e. setting $a = b$):

$$\begin{aligned} \Phi^{\langle m\ell \rangle}_{\text{SNOVA}} : \bigwedge^{\ell v} \mathbb{F}_q^{\ell n} &\to \left(\bigwedge^{\ell v+2} \mathbb{F}_q^{\ell n} \right)^{m\ell} \\ \alpha &\mapsto (\, \alpha \wedge \widehat{Q}_{1,0,0}, \ldots, \alpha \wedge \widehat{Q}_{m,\ell-1,\ell-1} \,). \end{aligned}$$

The input space of Φ_{SNOVA} has dimension $\binom{\ell n}{\ell v}$. We will be looking for a special element arising from the annihilator space of the oil space. Due to the

block structure of oil space, we will be able to restrict the input Φ_{SNOVA} to smaller subspace that contains the desired element, thus effectively reducing the dimension of the domain of Φ_{SNOVA}, as well as its rank. This reduction will ultimately yield a decrease in the cost of the kernel computation.

4.1 Dimension of the Domain Space of Φ_{SNOVA}

In analogy to the approach of Section 3, the variables in the wedge attack when applied to SNOVA correspond to the maximal minors of an $\ell(v+o) \times \ell v$ matrix, denoted $\mathbf{A}$, the column space of which is the annihilator space $\mathcal{A}$ of $\mathcal{O}$. In the UOV setting, this matrix is generic; the dimension of the domain space of Φ_{UOV} is $\binom{v+o}{v}$, which is the number of linearly independent maximal minors and thus the number of variables in the attack. There is an extra structure built into, however.

Due to the fact that $\mathcal{O}$ lies over $\mathbb{F}_q[\mathbf{S}]$, its annihilator space $\mathcal{A}$ does as well. Recall that each $\ell \times \ell$ block of $\mathbf{A}$ is given by an element of $\mathbb{F}_q[\mathbf{S}]$, which is defined as:

$$\mathbb{F}_q[\mathbf{S}] = \{b_0 + b_1\mathbf{S} + \cdots + b_{\ell-1}\mathbf{S}^{\ell-1} \mid b_i \in \mathbb{F}_q\} \cong \mathbb{F}_{q^\ell}.$$

This extra structure from the SNOVA construction changes the analysis of the resulting system. Observe that this structure induces linear dependencies among the maximal minors, effectively reducing the number of variables in the system. In particular, for any $\mathbf{\Gamma} \in \mathbb{F}_q[\mathbf{S}]$, there is the relation:

$$\mathbf{\Gamma}^{\otimes(v+o)}\mathbf{A} = \mathbf{A}\mathbf{\Gamma}^{\otimes v} \tag{8}$$

since elements of $\mathbb{F}_q[\mathbf{S}]$ commute and $\mathbf{A} \in \left(\mathbb{F}_q[\mathbf{S}]\right)^{(v+o)\times v}$. This identity can be used to express a linear relation among the minors of $\mathbf{A}$, specifically:

$$(\mathbf{e}_{i_1}^\intercal \cdot \mathbf{A}) \wedge \cdots \wedge (\mathbf{e}_{i_{v\ell}}^\intercal \cdot \mathbf{A}) = \det(\mathbf{\Gamma})^{-v} \cdot (\mathbf{e}_{i_1}^\intercal \cdot \mathbf{\Gamma}^{\otimes(v+o)}\mathbf{A}) \wedge \cdots \wedge (\mathbf{e}_{i_{v\ell}}^\intercal \cdot \mathbf{\Gamma}^{\otimes(v+o)}\mathbf{A}).$$

One therefore expects that the number of linearly independent minors in SNOVA is less than $\binom{\ell(v+o)}{\ell v}$, the expected number of linearly independent minors for a generic UOV scheme with the same dimensions. The next proposition quantifies this effect precisely.

Proposition 1. *Let $\iota : \mathbb{F}_q[\mathbf{S}] \hookrightarrow \mathbb{F}_q^{\ell\times\ell}$ be the canonical injection. A matrix $\mathbf{M} \in (\mathbb{F}_q[\mathbf{S}])^{a\times b}$ induces a matrix $\iota(\mathbf{M}) \in \mathbb{F}_q^{a\ell\times b\ell}$ with an $\ell \times \ell$ block structure. By construction, the matrix $\mathbf{A}$ corresponding to the annihilator space is an element of $(\mathbb{F}_q[\mathbf{S}])^{n\times v}$, so $\iota(\mathbf{A}) \in \mathbb{F}_q^{n\ell\times v\ell}$. The number of linearly independent maximal minors of $\iota(\mathbf{A})$ is*

$$\binom{n}{v}^\ell.$$

Proof. The field $\mathbb{F}_{q^\ell}$ is a splitting field of the characteristic polynomial of $\mathbf{S}$. Let $\lambda_1, \ldots, \lambda_\ell$ be the eigenvalues of $\mathbf{S}$ over $\mathbb{F}_{q^\ell}$, with multiplicity. There exists

$\mathbf{B} \in \mathrm{GL}_\ell(\mathbb{F}_{q^\ell})$ such that

$$\mathbf{B}^{-1}\mathbf{S}\mathbf{B} = \mathrm{diag}(\lambda_1, \ldots, \lambda_\ell).$$

Note that all matrices in $\mathbb{F}_q[\mathbf{S}]$ are simultaneously diagonalized by $\mathbf{B}$.

We first make two reductions that do not change the desired dimension. Extending scalars from $\mathbb{F}_q$ to $\mathbb{F}_{q^\ell}$ preserves linear dependence among the minor coordinate functions. Also, multiplication by invertible matrices on the left and on the right preserves the dimension of their span: right multiplication by an invertible $v\ell \times v\ell$ matrix multiplies every maximal minor by its determinant, and left multiplication acts on the vector of maximal minors through the invertible exterior-power operator $\bigwedge^{v\ell}$. Equivalently, this is the Cauchy-Binet formula.

Put $\mathbf{U}_{a\ell} = \mathbf{B}^{\otimes a}$. After replacing $\iota(\mathbf{A})$ by

$$\mathbf{A}_1 = \mathbf{U}_{n\ell}^{-1}\iota(\mathbf{A})\mathbf{U}_{v\ell},$$

all $\ell \times \ell$ blocks become diagonal. Reordering rows and columns then gives a block diagonal matrix

$$\mathbf{A}_2 = \mathrm{diag}(\mathbf{B}_1, \ldots, \mathbf{B}_\ell),$$

where each $\mathbf{B}_s$ is an $n \times v$ matrix. Specifically, if the (i,j) entry of $\mathbf{A}$ is $a_{ij}(\mathbf{S})$, then the (i,j) entry of $\mathbf{B}_s$ is $a_{ij}(\lambda_s)$. For fixed (i,j), the map from the coefficients of a_{ij} to the values $(a_{ij}(\lambda_1), \ldots, a_{ij}(\lambda_\ell))$ is a Vandermonde map, so it is invertible over $\mathbb{F}_{q^\ell}$. Thus the entries of $\mathbf{B}_1, \ldots, \mathbf{B}_\ell$ are independent generic coordinates after this change of variables.

Now consider a maximal minor of $\mathbf{A}_2$. The matrix $\mathbf{A}_2$ has ℓ column blocks, each of width v. Suppose the chosen rows contain k_s rows from the s-th n-row block. The selected square matrix has rank at most $\sum_s \min(k_s, v)$. Since $\sum_s k_s = v\ell$, this upper bound is $v\ell$ only when $k_s = v$ for every s. Thus every nonzero maximal minor is indexed by a tuple $(J_1, \ldots, J_\ell)$ with $J_s \subset [n]$ and $|J_s| = v$. For such a row choice, the submatrix is block diagonal, and its determinant is, up to the sign of the row permutation,

$$\prod_{s=1}^{\ell} \det((\mathbf{B}_s)_{J_s,*}).$$

This is the coordinate form of the standard exterior decomposition. If $\mathcal{F}_s = \mathbb{F}_{q^\ell}^n$, then

$$\bigwedge^{v\ell}\left(\bigoplus_{s=1}^{\ell}\mathcal{F}_s\right) \cong \bigoplus_{k_1+\cdots+k_\ell=v\ell}\ \bigotimes_{s=1}^{\ell}\bigwedge^{k_s}\mathcal{F}_s.$$

It remains to prove that these products are linearly independent. For one generic $n \times v$ matrix $\mathbf{B}_s$, the maximal minors are linearly independent: the minor indexed by $J = \{j_1 < \cdots < j_v\}$ contains the monomial $\prod_{c=1}^{v} b_{j_c,c}$, and this monomial occurs in no maximal minor with a different row set. Let R_s be the polynomial ring in the entries of $\mathbf{B}_s$, and let $V_s \subset R_s$ be the span of its maximal

minors. The variables of the different blocks are disjoint, so the full polynomial ring is the tensor product $R_1 \otimes_{\mathbb{F}_{q^\ell}} \cdots \otimes_{\mathbb{F}_{q^\ell}} R_\ell$. Here we use the notation $\otimes_{\mathbb{F}_{q^\ell}}$ to indicate that we view $R_1, \ldots, R_\ell$ as $\mathbb{F}_{q^\ell}$-vector spaces. Under this identification, products of one minor from each block are the pure tensors in $V_1 \otimes_{\mathbb{F}_{q^\ell}} \cdots \otimes_{\mathbb{F}_{q^\ell}} V_\ell$. Hence they are linearly independent, and their number is

$$\prod_{s=1}^{\ell} \binom{n}{v} = \binom{n}{v}^{\ell}.$$

Returning from $\mathbb{F}_{q^\ell}$ to the base field does not change this dimension, which proves the proposition. □

4.2 Rank of Φ_{SNOVA}

In the linear system induced by Φ_{SNOVA}, we can use any set of 2-forms which vanish on the same oil space as the bilinear forms corresponding to the public key of SNOVA. Recall that an attacker can generate equations using $m\ell^2$ different bilinear forms that share the same oil space. These forms correspond to matrices of the type $\mathbf{P}_{i,a,b} = (\mathbf{S}^a)^{\otimes n}\mathbf{P}_i(\mathbf{S}^b)^{\otimes n}$, where $0 \le a, b < \ell$ and $1 \le i \le m$. The 2-forms $\widehat{Q}_{i,a,b}$ associated with $\mathbf{Q}_{i,a,b} = \mathbf{P}_{i,a,b} + \mathbf{P}_{i,a,b}^{\intercal}$ can be used to wedge with $\widehat{\mathcal{A}}$ to generate linear equations. (Here, like in the UOV setting, $\widehat{\mathcal{A}}$ denotes the form associated with the annihilator space of the oil space via the Plücker embedding, which is unique up to a scalar.)

However, it should be noted that, using $m\ell^2$ 2-forms introduces a scenario where many of the linear equations are dependent. Indeed, we note that at most $m\ell$ 2-forms of the form $\widehat{Q}_{i,0,b}$ corresponding to $\mathbf{P}_i(\mathbf{S}^b)^{\otimes n} + (\mathbf{S}^b)^{\otimes n}\mathbf{P}_i^{\intercal}$ can create additional linearly independent equations by observing that the equations arising from $\widehat{Q}_{i,a,b}$ can be obtained as linear combinations of the equations arising from $\widehat{Q}_{i,0,b-a}$, as follows:

$$\begin{aligned}
0 &= \widehat{Q}_{i,0,b-a} \wedge \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{v\ell} \\
&= \left(\mathbf{P}_i(\mathbf{S}^{b-a})^{\otimes n} + (\mathbf{S}^{b-a})^{\otimes n}\mathbf{P}_i^{\intercal}\right) \wedge \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{v\ell} \\
&= \left((\mathbf{S}^a)^{\otimes n}\mathbf{P}_i(\mathbf{S}^b)^{\otimes n} + (\mathbf{S}^b)^{\otimes n}\mathbf{P}_i^{\intercal}(\mathbf{S}^a)^{\otimes n}\right) \wedge (\mathbf{S}^a)^{\otimes n}\mathbf{a}_1 \wedge \cdots \wedge (\mathbf{S}^a)^{\otimes n}\mathbf{a}_{v\ell} \\
&= \det(\mathbf{S})^{a(v\ell+2)}\widehat{Q}_{i,a,b} \wedge \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{v\ell}
\end{aligned}$$

This immediately implies that all of the $m\ell^2$ 2-forms are obtainable from at most $m\ell$ 2-forms $\mathbf{P}_i(\mathbf{S}^b)^{\otimes n} + (\mathbf{S}^b)^{\otimes n}\mathbf{P}_i^{\intercal}$ where $0 \le b < \ell$. Therefore, we analyze two scenarios where the attacker selects two different sets of 2-forms to attack SNOVA via solving the kernel vector in Φ_{SNOVA}.

Scenario 1: Using m many 2-forms $\widehat{Q}_{i,0,0}$. In this scenario, we use m base 2-forms $\mathbf{P}_i + \mathbf{P}_i^{\intercal}$ of SNOVA and build the linear map $\Phi_{\text{SNOVA}}^{\langle m \rangle}$ (Definition 2).

These m equations serve as a starting point for the analysis of the rank of the full system Φ_{SNOVA}, see Scenario 2.

The previous section has already considered how the dimension of the domain of $\Phi^{\langle m\rangle}_{\mathrm{SNOVA}}$ could be reduced to take into account the block structure of the oil space. It remains to analyze the rank of the map. This analysis is provided by the next proposition, under an exactness assumption of the same nature Conjecture 2. (The cochain complex that this exactness assumption bears on is precisely defined in the proof.)

Proposition 2. *Under an exactness assumption analogous to Conjecture 2, the following holds. Let* $0 \le a_i \le n-v$, $\mathbf{a} = (a_1, ..., a_\ell)$, $\sum_{i=1}^{\ell} a_i = 2r$, *and*

$$N^{\langle m\rangle}_{r,k}(t, \mathbf{z}, y; \mathbf{a}) = \left[z_1^{a_1} \cdots z_\ell^{a_\ell}\, t^{\,r} y^{\,k}\right] \left(\prod_{i=1}^{\ell} \frac{1}{1 - t\, z_i^2} \prod_{1\le i<j\le \ell} \frac{1}{1 - t\, y\, z_i z_j}\right)^m$$

$$R^{\langle m\rangle}_{\mathrm{SNOVA}} = \sum_{r=1}^{\lfloor \frac{\ell o}{2} \rfloor} (-1)^{\,r+1} \sum_{\substack{0\le a_i\le o \\ \sum_i a_i = 2r}} \left(\sum_{k=0}^{r} N^{\langle 1\rangle}_{r,k}(t, \mathbf{z}, y; \mathbf{a})\right) \prod_{i=1}^{\ell} \binom{n}{v + a_i}$$

then

$$\operatorname{rank} \Phi^{\langle m\rangle}_{\mathrm{SNOVA}} = \min\left\{ \binom{n}{v}^{\ell} - 1\ ,\ R^{\langle m\rangle}_{\mathrm{SNOVA}} \right\}. \tag{9}$$

Proof. After the change of basis used in Proposition 1, we have $\mathcal{F} = \mathbb{F}^{\ell n} = \bigoplus_{i=1}^{\ell} \mathcal{F}_i$, $\mathcal{A} = \bigoplus_{i=1}^{\ell} \mathcal{A}_i$, and $\mathcal{F}_i = \mathcal{A}_i \oplus \mathcal{B}_i$, with $\dim \mathcal{F}_i = n$, $\dim \mathcal{A}_i = v$, and $\dim \mathcal{B}_i = o$. The SNOVA parameters considered here satisfy $o \le v$; this restriction is used below to identify the relevant wedge subspaces with full block-degree components. By Proposition 1, the reduced Plücker coordinates are the coordinates on

$$\mathcal{G} := \bigotimes_{i=1}^{\ell} \bigwedge^{v} \mathcal{F}_i \subseteq \bigwedge^{\ell v} \mathcal{F} \quad \text{where} \quad \dim \mathcal{G} = \binom{n}{v}^{\ell}.$$

Lemma 5. *For every* $r \ge 0$,

$$\bigwedge^{r} \mathcal{A} \wedge \bigwedge^{r} \mathcal{F} \wedge \mathcal{G} = \bigoplus_{\substack{0\le a_i\le o \\ \sum_i a_i = 2r}} \bigotimes_{i=1}^{\ell} \bigwedge^{v+a_i} \mathcal{F}_i.$$

In particular,

$$\dim\left(\bigwedge^{r} \mathcal{A} \wedge \bigwedge^{r} \mathcal{F} \wedge \mathcal{G}\right) = \sum_{\substack{0\le a_i\le o \\ \sum_i a_i = 2r}} \prod_{i=1}^{\ell} \binom{n}{v + a_i}.$$

Proof. Distribute the r factors from $\mathcal{A}$ and the r additional factors from $\mathcal{F}$ among the ℓ blocks. This gives a sum of subspaces

$$\bigotimes_{i=1}^{\ell} \left(\bigwedge_{\sum_i p_i = r}^{p_i} \mathcal{A}_i \wedge \bigwedge_{\sum_i q_i = r}^{v+q_i} \mathcal{F}_i \right)$$

For one block, the same argument as in Lemma 3 gives

$$\bigwedge^{p_i} \mathcal{A}_i \wedge \bigwedge^{v+q_i} \mathcal{F}_i = \bigwedge^{v+p_i+q_i} \mathcal{F}_i,$$

because $o \leq v + q_i$. If $a_i = p_i + q_i$, then $0 \leq a_i \leq o$ for every nonzero summand and $\sum_i a_i = 2r$. Conversely, every such $\mathbf{a}$ arises by summing the corresponding block-degree pieces. The equality follows from the decomposition of the exterior algebra by the block degrees. □

For each base form $\widehat{Q}_j$, wedging one component of $\widehat{Q}_j \in \mathcal{A} \wedge \mathcal{F}$ either raises one block degree by 2 or raises two distinct block degrees by 1. We fix the m 2-forms $\widehat{Q}_1, ..., \widehat{Q}_m \in \mathcal{A} \wedge \mathcal{F} \subset \bigwedge^2 \mathcal{F}$ spanned by the vectors $v \wedge f$, for $v \in \mathcal{A}$ and $f \in \mathcal{F}$.

Because $\mathcal{G}$ is now split into ℓ blocks where wedging by any fixed $\widehat{Q}_i$ either "increases" a_i by 2 in a single block (a "double") or "adds" 1 to a_i and a_j in two distinct blocks (a "split"), indexing $\widehat{Q}_i$ with $\mathbb{F}_q^m$ is not sufficient: we have to index them with $\mathcal{E} = \mathrm{Sym}^2(\mathbb{F}_q^\ell)^m$. Furthermore, among ℓ single blocks, no block can exceed $a_i \leq n - v$. That means to encode the behaviour of the wedges, we need the variables z_i to record block increments, the variable m to record which $\widehat{Q}_i$ is being used, and the marker y to record the splits. Therefore, for a single $\widehat{Q}$, the ordinary generating function is

$$G(t, \mathbf{z} = (z_1, \ldots, z_l), y) = \prod_{i=1}^{\ell} \frac{1}{1 - tz_i^2} \prod_{1 \leq i < j \leq \ell} \frac{1}{1 - tyz_i z_j}.$$

The series expansion for G can be obtained by substituting the geometric series expansion for $\dfrac{1}{1 - tz_i^2}$ and $\dfrac{1}{1 - tyz_i z_j}$.

Over all m forms these choices multiply, hence the number of degree-r selections that (i) use exactly k splits and (ii) land in occupancy $\mathbf{a} = (a_1, ..., a_\ell)$ where $\sum_{i=1}^{\ell} a_i = 2r$ with $0 \leq a_i \leq n - v$ equals

$$N_{r,k}(t, \mathbf{z}, y; \mathbf{a}) = \left[\mathbf{z}^{\mathbf{a}}\, t^r y^k\right]\, G(t, \mathbf{z}, y)^m$$

Now, we define a bigraded vector space

$$\mathcal{C}^a = \mathrm{Sym}^{2m} \mathbb{F}_q^{\ell} \otimes \left(\bigwedge^a \mathcal{A} \wedge \bigwedge^a \mathcal{F} \wedge \mathcal{G} \right)$$

and the operator

$$\delta(\eta \otimes \mu) \coloneqq \sum_{i=1}^{m} (x_i \cdot \eta) \otimes (\widehat{Q}_i \wedge \mu).$$

Note that we still have $\delta^2 = 0$. Therefore, we fix ℓ, v and then build a cochain complex

$$0 \longrightarrow \mathcal{C}^0 \xrightarrow{\delta_0} \mathcal{C}^1 \xrightarrow{\delta_1} \cdots \mathcal{C}^m \xrightarrow{\delta_m} \ldots \longrightarrow 0 \,.$$

Here, every step raises the $\left(\bigwedge^{\bullet} \mathcal{A} \wedge \bigwedge^{\bullet} \mathcal{F} \wedge \mathcal{G} \right)$-degree by $+2$.

Projecting to the $\mathbf{a}$-summand of $\bigwedge^r \mathcal{A} \wedge \bigwedge^{v\ell+r} \mathcal{F}$, the span contributed from $\mathcal{E}^r$ has dimension at most

$$\sum_{k=0}^{r} N_{r,k}(t, \mathbf{z}, y; \mathbf{a})$$

while the exterior factor has dimension given by Equation (5). Under the assumption that the earlier cochain complex is exact, we have

$$\operatorname{rank} \Phi_{\mathrm{SNOVA}} = \sum_{r=1}^{\lfloor \frac{\ell o}{2} \rfloor} (-1)^{r+1} \sum_{\substack{\mathbf{a} \in \{0,\ldots,o\}^{\ell} \\ \sum_i a_i = 2r}} \Big(\sum_{k=0}^{r} N_{r,k}(t, \mathbf{z}, y; \mathbf{a}) \Big) \prod_{i=1}^{\ell} \binom{v+o}{v+a_i}.$$

□

Scenario 2: Using ℓm many 2-forms $\widehat{Q}_{i,\boldsymbol{a},\mathbf{0}}$ In this scenario, we use ℓm 2-forms $\mathbf{P}_i(\mathbf{S}^b)^{\otimes n} + (\mathbf{S}^b)^{\otimes n}\mathbf{P}_i^{\intercal}$ where $0 \le b < \ell$ of SNOVA and build the linear mapping $\Phi_{\mathrm{SNOVA}}^{\langle \ell m \rangle}$ (Definition 2).

Proposition 3 analyzes the rank of this map. As in the previous scenario, the analysis relies on an exactness assumption of the same nature as Conjecture 2 (which is precisely defined in the proof). While the proof involves further analysis, ultimately, it should be noted that the only difference between Proposition 2 and Proposition 3 lies in the definition of $N_{r,k}^{\langle \ell m \rangle}(t, \mathbf{z}, y; \mathbf{a})$, where Proposition 3 replaces the factor $(1 - t\, y\, z_i z_j)$ with $(1 - t\, y\, z_i z_j)^2$.

Proposition 3. *Under an exactness assumption analogous to Conjecture 2, the following holds. Let* $0 \le a_i \le n - v$, $\mathbf{a} = (a_1, \ldots, a_\ell)$, $\sum_{i=1}^{\ell} a_i = 2r$, *and*

$$N_{r,k}^{\langle \ell m \rangle}(t, \mathbf{z}, y; \mathbf{a}) = \left[z_1^{a_1} \cdots z_\ell^{a_\ell}\, t^r y^k \right] \left(\prod_{i=1}^{\ell} \frac{1}{1 - t\, z_i^2} \prod_{1 \le i < j \le \ell} \frac{1}{(1 - t\, y\, z_i z_j)^2} \right)^m$$

$$R_{\text{SNOVA}}^{\langle \ell m \rangle} = \sum_{r=1}^{\lfloor \frac{\ell o}{2} \rfloor} (-1)^{r+1} \sum_{\substack{0 \le a_i \le o \\ \sum_i a_i = 2r}} \left(\sum_{k=0}^{r} N_{r,k}^{\langle 2 \rangle}(t, \mathbf{z}, y; \mathbf{a}) \right) \prod_{i=1}^{\ell} \binom{n}{v + a_i}$$

then

$$\operatorname{rank} \Phi_{\text{SNOVA}}^{\langle \ell m \rangle} = \min \left\{ \binom{n}{v}^{\ell} - 1 \ , \ R_{\text{SNOVA}}^{\langle \ell m \rangle} \right\}. \tag{10}$$

Proof. Because $\mathbf{S}$ is diagonalizable, we have

$$\mathbf{S} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1} \quad \text{where } \mathbf{D} = \operatorname{Diag}(\lambda_1, ..., \lambda_l)$$

Let

$$\mathbf{\Pi}_r = \mathbf{U}\mathbf{E}_r\mathbf{U}^{-1} \quad \text{where } \mathbf{E} = \begin{pmatrix} 0 \cdots 0 \cdots 0 \\ \vdots \ddots \vdots \quad \vdots \\ 0 \cdots 1 \cdots 0 \\ \vdots \quad \vdots \ddots \vdots \\ 0 \cdots 0 \cdots 0 \end{pmatrix} \quad \text{with 1 in position } r$$

then we have $\mathbf{\Pi}_r^2 = \mathbf{\Pi}_r$, $\mathbf{\Pi}_r\mathbf{\Pi}_s = 0 \ \forall r \neq s$ and $\sum_{i=1}^{\ell} \lambda_i \mathbf{\Pi}_i = \mathbf{S}$. By consequence, we have

$$\mathbf{S}^b = \sum_{i=1}^{\ell} \lambda_i^b \mathbf{\Pi}_i \ .$$

Let $Q^{\langle b \rangle} = \widehat{Q}_{\bullet,b,0}$ be a 2-form scaled by the matrix $(\mathbf{S}^b)^{\otimes n}$, then we have

$$\begin{aligned} \mathbf{Q}^{\langle b \rangle} &= (\mathbf{S}^b)^{\otimes n}\mathbf{P} + \mathbf{P}^{\intercal}(\mathbf{S}^b)^{\otimes n} \\ &= \left(\sum_{i=1}^{\ell} \lambda_i^b \mathbf{\Pi}_i \right)^{\otimes n} \mathbf{P} + \mathbf{P}^{\intercal} \left(\sum_{i=1}^{\ell} \lambda_i^b \mathbf{\Pi}_i \right)^{\otimes n} \\ &= \sum_{i=1}^{\ell} \bigwedge_{i}^{b} \left(\mathbf{\Pi}_i^{\otimes n}\mathbf{P} + \mathbf{P}^{\intercal}\mathbf{\Pi}_i^{\otimes n} \right) \\ &= \sum_{i=1}^{\ell} \bigwedge_{i}^{b} \mathbf{Z}^{\langle i \rangle} \quad \text{where } \mathbf{Z}^{\langle i \rangle} = \left(\mathbf{\Pi}_i^{\otimes n}\mathbf{P} + \mathbf{P}^{\intercal}\mathbf{\Pi}_i^{\otimes n} \right) \ . \end{aligned}$$

Let $\mathbf{V}$ be the Vandermonde matrix where $\mathbf{V}_{b,i} = \lambda_i^b$, then

$$\begin{pmatrix} \mathbf{Q}^{\langle 1 \rangle} \\ \mathbf{Q}^{\langle 2 \rangle} \\ \vdots \\ \mathbf{Q}^{\langle \ell \rangle} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \cdots & \lambda_\ell \\ \lambda_1^2 & \cdots & \lambda_\ell^2 \\ \vdots & \ddots & \vdots \\ \lambda_1^\ell & \cdots & \lambda_\ell^\ell \end{pmatrix} \begin{pmatrix} \mathbf{Z}^{\langle 1 \rangle} \\ \mathbf{Z}^{\langle 2 \rangle} \\ \vdots \\ \mathbf{Z}^{\langle \ell \rangle} \end{pmatrix} = \mathbf{V} \begin{pmatrix} \mathbf{Z}^{\langle 1 \rangle} \\ \mathbf{Z}^{\langle 2 \rangle} \\ \vdots \\ \mathbf{Z}^{\langle \ell \rangle} \end{pmatrix}.$$

By consequence, we have

$$\begin{pmatrix} \mathbf{Q}_i^{\langle 1 \rangle} \\ \mathbf{Q}_i^{\langle 2 \rangle} \\ \vdots \\ \mathbf{Q}_i^{\langle \ell \rangle} \end{pmatrix} = \mathbf{V} \begin{pmatrix} \mathbf{Z}_i^{\langle 1 \rangle} \\ \mathbf{Z}_i^{\langle 2 \rangle} \\ \vdots \\ \mathbf{Z}_i^{\langle \ell \rangle} \end{pmatrix} \qquad \forall i \in \{1, ..., m\}\ .$$

Since the λ_i are distinct in the splitting field, $\mathbf{V}$ is invertible, so the ℓ-tuple $(\mathbf{Q}^{\langle b \rangle})_i$ spans the same ℓ-dimensional space as $(\mathbf{Z}^{\langle b \rangle})_i$.

For a fixed ℓv-form α in $\bigwedge^{\ell v} F$, let φ be the linear mapping

$$\begin{aligned} \varphi : \left(\bigwedge^{2} F \right)^m &\longrightarrow \left(\bigwedge^{\ell v+2} F \right)^m \\ (\widehat{Q}_1, \ldots, \widehat{Q}_m) &\longmapsto (\alpha \wedge \widehat{Q}_1, \ldots, \alpha \wedge \widehat{Q}_m) \end{aligned}$$

then we have

$$\varphi \begin{pmatrix} Q_i^{\langle 1 \rangle} \\ Q_i^{\langle 2 \rangle} \\ \vdots \\ Q_i^{\langle \ell \rangle} \end{pmatrix} = \varphi \begin{pmatrix} Z_i^{\langle 1 \rangle} \\ Z_i^{\langle 2 \rangle} \\ \vdots \\ Z_i^{\langle \ell \rangle} \end{pmatrix}.$$

In Proposition 1, we proved that, up to some change-of-basis, the vinegar subspace achieves the block diagonal form

$$\mathrm{diag}(\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_\ell).$$

Again, we consider the linear map corresponding to SNOVA where each $\mathcal{F}_i$ corresponds to $\begin{pmatrix} 0 & \cdots & \mathbf{B_i} & \cdots & 0 \end{pmatrix}^\mathsf{T}$, then we build a cochain complex related to the one in Proposition 2 with the following operator

$$\begin{aligned} \delta : \bigotimes_{j=1}^{\ell} \bigwedge^{v} \mathcal{F}_j &\longrightarrow \bigoplus_{\sum a_j = 2} \bigotimes_{j=1}^{\ell} \bigwedge^{v+a_j} \mathcal{F}_j \\ \mu &\longmapsto \mu \wedge \widehat{Q} \end{aligned}$$

Because $\mathbf{\Pi}_i$ is the projector onto $\mathcal{F}_i$, hence $\mathbf{\Pi}_i|_{\mathcal{F}_i} = \mathrm{Id}$ and $\mathbf{\Pi}_i|_{\mathcal{F}_j} = 0 \ \ (j \neq i)$. Let B_P be a bilinear form

$$\begin{aligned} B_P : \mathcal{F} \times \mathcal{F} &\longrightarrow k \\ B_P(x, y) &\longmapsto x^\intercal P y. \end{aligned}$$

We need two following lemmas:

Lemma 6.

$$Z^{\langle i \rangle}|_{\bigwedge^2 \mathcal{F}_j} = 0 \text{ for } j \neq i \ .$$

Proof. The bilinear form corresponding to $Z^{\langle i \rangle}$ is

$$Z^{\langle i \rangle}(x, y) = B_P(\mathbf{\Pi}_i x, y) + B_P(x, \mathbf{\Pi}_i y) \ .$$

Then, $\forall x, y \in \mathcal{F}_j$ where $j \neq i$, we have $\mathbf{\Pi}_i x = \mathbf{\Pi}_i y = 0 \implies Z^{\langle i \rangle}(x, y) = 0$. So the restriction of $Z^{\langle i \rangle}$ to $\mathcal{F}_j \times \mathcal{F}_j$ is identically zero when $j \neq i$. Equivalently, $Z^{\langle i \rangle}|_{\bigwedge^2 \mathcal{F}_j} = 0 \quad \text{for } j \neq i.$ □

Lemma 7.

$$Z^{\langle i \rangle}|_{\mathcal{F}_i \wedge \mathcal{F}_j} = B_P|_{\mathcal{F}_i \times \mathcal{F}_j} \text{ for all } j \neq i \ .$$

Proof. Indeed, restrict to $x \in \mathcal{F}_i, y \in \mathcal{F}_j$. Using $\mathbf{\Pi}_i x = x$, $\mathbf{\Pi}_i y = 0$ and $\mathbf{\Pi}_j x = 0$, $\mathbf{\Pi}_j y = y$, we get

$$\begin{aligned} Z^{\langle i \rangle}(x, y) &= B_P(\mathbf{\Pi}_i x, y) + B_P(x, \mathbf{\Pi}_i y) = B_P(x, y), \\ Z^{\langle j \rangle}(x, y) &= B_P(\mathbf{\Pi}_j x, y) + B_P(x, \mathbf{\Pi}_j y) = B_P(x, y). \end{aligned}$$

So both projectors i and j hit the same cross-block wedge space $\mathcal{F}_i \wedge \mathcal{F}_j$. □

Recall that with the block splitting $\mathcal{F} = \bigoplus_{j=1}^{\ell} \mathcal{F}_j$, every 2-form is still decomposed into: Type 1 ("double") $\bigwedge^2 \mathcal{F}_i$, and Type 2 ("split") $\mathcal{F}_i \wedge \mathcal{F}_j$ with $i < j$ (see Section 4.6). According to Lemma 6, for $j \neq i$, $(Z^{\langle i \rangle})|_{\mathcal{F}_j \wedge \mathcal{F}_j} = 0$, so the $\bigwedge^2 \mathcal{F}_j$ piece only exists when $j = i$. Thus, for each i, each $Z^{\langle i \rangle}$ provides exactly one "double channel" $\bigwedge^2 \mathcal{F}_i$. However, per Lemma 7, for each fixed pair $i < j$, *for a single base form* k, we have two independent ways to realize a split into (i, j): one via $Z_k^{\langle i \rangle}$ and one via $Z_k^{\langle j \rangle}$. Therefore, similar to the analysis of the previous scenario, the "double" part contributes a factor

$$\frac{1}{1 - t z_i^2}$$

while the "split" part contributes a factor

$$\left(\frac{1}{1 - t y z_i z_j}\right) \times \left(\frac{1}{1 - t y z_i z_j}\right) = \frac{1}{(1 - t y z_i z_j)^2}.$$

Now, we re-organize ℓm 2-form as

$$\begin{pmatrix} Z_1^{\langle 1\rangle} & Z_1^{\langle 2\rangle} & \cdots & Z_1^{\langle \ell\rangle} \\ Z_2^{\langle 1\rangle} & Z_2^{\langle 2\rangle} & \cdots & Z_2^{\langle \ell\rangle} \\ \vdots & \vdots & \ddots & \vdots \\ Z_m^{\langle 1\rangle} & Z_m^{\langle 2\rangle} & \cdots & Z_m^{\langle \ell\rangle} \end{pmatrix}$$

then every 2-forms in the column will contribute a factor of $\frac{1}{1-tz_i^2}$ and $\frac{1}{1-tyz_iz_j}$ while every 2-forms in the row will contribute a factor of $\frac{1}{1-tyz_iz_j}$ only.

Finally, under the assumption that the earlier cochain complex is exact, we conclude that

$$\operatorname{rank} \Phi_{\mathrm{SNOVA}}^{\langle \ell m\rangle} = \sum_{r=1}^{\lfloor \frac{\ell o}{2} \rfloor} (-1)^{r+1} \sum_{\substack{0\le a_i\le o \\ \sum_i a_i = 2r}} \left(\sum_{k=0}^{r} N_{r,k}(t,\mathbf{z},y;\mathbf{a}) \right) \prod_{i=1}^{\ell} \binom{n}{v+a_i}$$

where

$$N_{r,k}(t,\mathbf{z},y;\mathbf{a}) = \left[z_1^{a_1}\cdots z_\ell^{a_\ell}\, t^{\,r}\, y^{\,k} \right] \left(\prod_{i=1}^{\ell} \frac{1}{1-t\,z_i^2} \prod_{1\le i<j\le \ell} \frac{1}{(1-t\,y\,z_iz_j)^2} \right)^m .$$

□

4.3 Experimental Validation

We conducted experiments with various parameters and found that the number of linearly independent equations always matches *exactly* Formula 9 for Scenario 1 and Formula 10 for Scenario 2: see Figure 1 and Tables 4 and 5 for all the considered parameters, and the corresponding ranks.

4.4 Unwanted Solutions

One important factor in the wedge attack is that we want the kernel of any linear system Φ built for UOV or SNOVA to have dimension 1. However, in Conjecture 1 when $v \ge 2m$, we can easily see that $\dim \ker \Phi_{\mathrm{UOV}} > 1$ because

$$\left(\widehat{Q}_1 \wedge \widehat{Q}_2 \wedge \cdots \wedge \widehat{Q}_m\right) \wedge \alpha \in \ker \Phi_{\mathrm{UOV}} \quad \text{for all } \alpha \in \bigwedge^{v-2m} \mathbb{F}^n.$$

Additionally, the oil space is not expected to be unique once $v \ge \frac{o-1}{2} m$. Such spurious solutions were identified in [36] and the author proposed restricting the attack to parameter choices satisfying

$$v < \min\left\{ \frac{o-1}{2},\, 2 \right\} m\,.$$

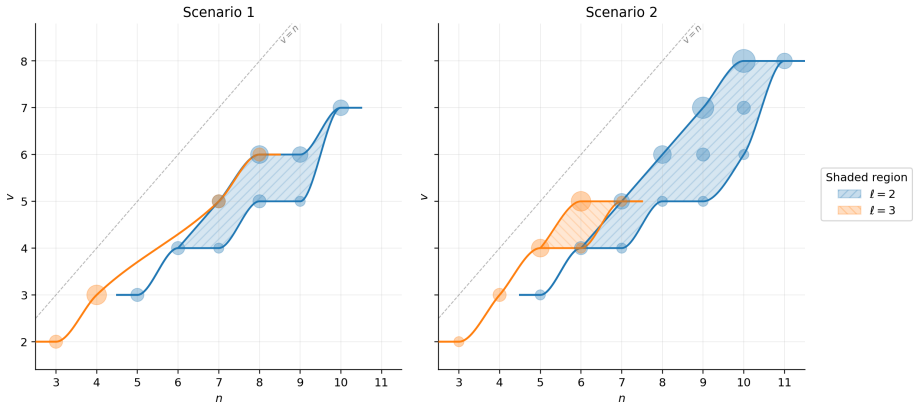

Fig. 1: The shaded area describes the (n, v, ℓ) triples, where $n = v + o$, that were tested for all relevant m under Scenarios 1 and 2. The bubbles mark the explicit tested parameters and larger bubbles correspond to larger tested values of m. The number of linearly independent equations always matches *exactly* Formula 9 and Formula 10.

Under this condition, the original wedge attack *does not apply* to SNOVA.

In our wedge attack against SNOVA, we do not expect $\left(\widehat{Q}_1 \wedge \cdots \wedge \widehat{Q}_m\right) \wedge \alpha$ to appear in the kernel if the structure of the annihilator is preserved along the wedge map. The reason is that the annihilator space $\mathcal{A}$ of $\mathcal{O}$ is highly structured with the appearance of elements from $\mathbb{F}_q[\mathbf{S}]$, while $\widehat{Q}_i \in \bigwedge^2 \mathbb{F}_q^{\ell n}$ do not share this property. Although, in the next section of the attack we propose two models of projections in which one is expected to have the spurious solutions in extreme cases, the absence of spurious solutions in the other model is supported by extensive experiments conducted specifically on SNOVA parameters, where v exceeds twice the number of 2-forms (see Table 4 and Table 5). The next section explores the uniqueness condition for the oil space of SNOVA, which is imposed both on v and the projection of oil coordinates.

4.5 Projection of Oil Coordinates

We may *project away* some oil coordinates: we keep the v vinegar coordinates together with only $o' \leq o$ oil coordinates. If we recover an o'-dimensional subspace of the *true* oil space, then the remaining $o - o'$ oil coordinates can be lifted (recovered) efficiently by linear algebra.

There is, however, a limit to how far we can project: when o' is too small, the attack can no longer reliably distinguish the genuine oil subspace from random ones, i.e., the oil space is no longer unique. In practice, the choice of o' is governed by three criteria:

1. The subspace output by the attack should be expected to lie inside the oil space, which is enforced by choosing o' large enough that the probability that such an o'-subspace exists "by chance" for a generic system is negligible.
   For $\mathbf{O} \supset \mathbf{O}' = \begin{bmatrix} \mathbf{X} \\ \mathbf{I}_{o'} \end{bmatrix} \in (\mathbb{F}_q[\mathbf{S}])^{n \times o'}$, there are $vo'\ell$ degrees of freedom. For each of the M polar forms $\mathbf{Q}_i$ (where $M = m$ or $m\ell$), $\iota(\mathbf{O}^\intercal \mathbf{Q}_i \mathbf{O}) = 0$ yields $\binom{o'\ell}{2} M$ quadratic equations over $\mathbb{F}_q$ since the $\mathbf{Q}_i$ are symmetric and have 0 on the diagonal. In order to expect $\mathbf{O}' \subset \mathbf{O}$, o' should be large enough so a solution is not expected for these equations. This is equivalent to $vo'\ell \leq \binom{o'\ell}{2} M$, so
   $$v \leq \frac{o'\ell - 1}{2} M \quad \text{or} \quad o' \geq \frac{\left(\frac{2v}{M} + 1\right)}{\ell} .$$
2. We can use the same argument as above, but using the equations $\mathbf{O}^\intercal \mathbf{P}_i \mathbf{O} = 0$, which corresponds to $m(o'\ell)^2$ equations. Thus o' must also satisfy
   $$vo'\ell \leq m(o'\ell)^2 \quad \text{or} \quad o' \geq \frac{v}{m\ell}.$$
   Notice that this condition is more restrictive than the first when $M = m\ell$ and $\ell \geq 3$.
3. After the projection, we must still obtain enough *independent* linear equations to recover the unknown Plücker coordinates. Equivalently, the linear system defining the candidate o'-dimensional subspace should be *overdetermined*: the number of linearly independent equations must be 1 less than the number of unknowns. Therefore, we require $R^{\langle m \rangle}_{\text{SNOVA}} \geq U_{\text{SNOVA}} - 1$ in Scenario 1 and $R^{\langle \ell m \rangle}_{\text{SNOVA}} \geq U_{\text{SNOVA}} - 1$ in Scenario 2.

Additionally, while we have parametrized our attack on the assumption that, when the oil space is represented in block form, an equal number o' of oil coordinates are projected out of each of the ℓ blocks, it is possible to consider the case that the number of coordinates projected out of each block could differ. In such case, rather than a single integer o', our attack would be parametrized by a vector $\mathbf{o}' = (o'_1, \ldots, o'_\ell)$. However, the conditions over the uniqueness of the oil subspace need to change when using $\mathbf{o}'$. We need to replace $o'\ell$ with $\sum_{i=1}^{\ell} o'_i$. This corresponds respectively to

$$\sum_{i=1}^{\ell} o'_i \geq \frac{2v}{M} + 1 \quad \text{and} \quad \sum_{i=1}^{\ell} o'_i \geq \frac{v}{m}.$$

It should be noted that if all but one of the entries of o' are zero, then the corresponding blocks no longer contain any projected oil coordinate. The equations restricted to the remaining block are then exactly the equations of an unstructured UOV system. We have implemented this projection technique and Table 2 in Appendix A reports five projected instances. It contains unbalanced projection cases with the desired one-dimensional kernel and extremely unbalanced

projection cases, where all but one of the entries of o' are zero, yielding kernels of dimension 2 and 17.

The added flexibility of unbalanced projection can enable a projection strategy that yields a better complexity than the strategies we have thus far considered. Thus, the vector $\mathbf{o}'$ will be selected so that the attack yields the lowest cost while satisfying the conditions mentioned above.

4.6 Density of the Linear System

Adapting the system from Equation (3) to SNOVA, each equation is now given by an index set $K \subset [n\ell]$ with $|K| = v\ell + 2$. In order to have a nonzero equation, at least v indices must be chosen from each block. Since there is the block-diagonal basis $\mathbf{A}_2 = \operatorname{diag}(\mathbf{B}_1, \ldots, \mathbf{B}_\ell)$ of the annihilator space $\mathcal{A}$, every nonzero $(v\ell)$-minor of $\mathbf{A}_2$ factors as a product of v-minors from the blocks; in particular, since $\mathbf{A} = \begin{bmatrix} \mathbf{I}_{v\ell} \\ -\mathbf{X}^\intercal \end{bmatrix}$, matrices $\mathbf{A}_1$ and each block of $\mathbf{A}_2$ inherit this structure, so $\mathbf{B}_i = \begin{bmatrix} \mathbf{I}_v \\ -\mathbf{X}_i^\intercal \end{bmatrix}$ for some $\mathbf{X}_i \in \mathbb{F}_{q^\ell}^{v \times o}$. As a result, each block has a v-minor with determinant 1. This yields two types for the equations:

1. $v + 2$ indices are chosen from one block, and v indices are chosen from the rest. Choosing the minors with determinant 1 from the remaining blocks yields
$$\sum_{j<j' \in J} (\mathbf{Q}_i)_{j,j'} \det((\mathbf{B}_k)_{J\setminus\{j,j'\},*}) = 0$$
for some $k \in [\ell]$ and index set $J \subset [n]$ with $|J| = v + 2$. There are $\ell\binom{n}{v+2}M$ many of these bilinear equation with (at most) $\binom{v+2}{2}$ nonzero terms.
2. $v + 1$ indices are chosen from two distinct blocks, and v indices are chosen from the rest. Choosing the minors with determinant 1 from the remaining blocks yields
$$\sum_{j \in J \text{ and } j' \in J'} (\mathbf{Q}_i)_{j,j'} \det((\mathbf{B}_k)_{J\setminus\{j\},*}) \det((\mathbf{B}_{k'})_{J'\setminus\{j'\},*}) = 0$$
for some $k < k' \in [\ell]$ and index sets $J, J' \subset [n]$ with $|J| = |J'| = v+1$. There are $\binom{\ell}{2}\binom{n}{v+1}^2 M$ many of these bilinear equation with (at most) $(v+1)^2$ nonzero terms.

Note that, $(v+1)^2 > \binom{v+2}{2}$ for all $v > 0$; moreover, because $2\binom{v+2}{2} \geq (v+1)^2$, treating all equations as having the greater density increases the complexity by less than one bit. Therefore, we use

$$\tau = (v+1)^2$$

in complexity estimation.

4.7 Complexity of the Attack

We now estimate the attack complexity for concrete SNOVA parameter sets in both scenarios. Let $U_{SNOVA} = \binom{v+o}{v}^{\ell}$ denote the number of variables, and let E_{SNOVA} denote the number of linear equations in the system induced by Φ_{SNOVA}. In Scenario 1, the attacker builds the linear system using only m base 2-forms, so the expected number of *linearly independent* equations is $R_{SNOVA}^{\langle m \rangle}$ (Proposition 2). In Scenario 2, the attacker instead uses all ℓm base 2-forms, in which case the expected number becomes $R_{SNOVA}^{\langle \ell m \rangle}$ (Proposition 3). The complexity of finding a unique kernel vector of the linear system using the Wiedemann [14,28,40] algorithm is therefore given by $\mathcal{O}\left(\tau E_{\mathrm{SNOVA}} U_{\mathrm{SNOVA}}\right)$ operations in $\mathbb{F}_{q^\ell}$, where $\tau = (v+1)^2$ is the density per row of the linear system.

After choosing the smallest possible o' per block, in accordance with Section 4.5, we observe that the difference between $\log_2(E_{\mathrm{SNOVA}})$ and $\log_2(U_{\mathrm{SNOVA}})$ is less than 4 bits. As in [36], we assume that the system is solvable by randomly choosing U_{SNOVA} many equations, and then the number of binary gates required for the attack against SNOVA can be accurately approximated by

$$\log_2\left(3 \cdot \left(2 \cdot (\log_2 q^\ell)^2 + \log_2 q^\ell\right) \cdot \tau \cdot U_{SNOVA}^2\right).$$

Here we assume that one multiplication in $\mathbb{F}_{q^\ell}$ costs approximately $(\log_2 q^\ell)^2$ bit operations, and we use the same order of magnitude for additions. Moreover, we assume that each field multiplication is accompanied by one field addition; this additional addition costs $\log_2 q^\ell$ bit additions. The factor 3 accounts for translating bit operations into an equivalent number of binary gates.

Table 1 summarizes our work by giving the complexity of our attack for "m forms" (Scenario 1) and "$m\ell$ forms" (Scenario 2) in comparison to the best attack as listed in the Round 2 specification of SNOVA [41]. Both our complexities and those of [41] assume free memory access and free on-the-fly matrix regeneration. The results are categorized by security levels of SNOVA under the names of SNOVA-I, SNOVA-III, and SNOVA-V. We also provide in Appendix D a more refined analysis of the cost incurred by memory accesses, using memory-locality optimization techniques based on those of [3]. For concreteness, we use a 2D model with a constant taken from [26], however, other models of the cost of memory access have been proposed and there is no consensus which memory model is the most accurate.

4.8 Modeling with Quadratic Equations

As mentioned in our discussion about density above, we get $\ell\binom{n}{v+2}M$ equations of the form

$$\sum_{j<j' \in J} (\mathbf{Q}_i)_{j,j'} \det((\mathbf{B}_k)_{J\setminus\{j,j'\},*}) = 0,$$

which can be used to reduce the number of variables. Each block behaves like a generic UOV instance, so the number of linearly independent equations of the first type is expected to satisfy Conjecture 1. Let E_0 denote the number of these

Table 1: Bit-complexity (gate count) estimates for our attack against SNOVA. The dash indicates that we found no o' that gives enough equations so that the kernel dimension is 1. Numbers in bold refer to the best attack while italicized ones refer to the cost below the required security level.

Variants	v	m	ℓ	Balanced projection				Unbalanced projection				Best attack in [41]
				o'	m forms	o'	$m\ell$ forms	o'	m forms	o'	$m\ell$ forms	
SNOVA-I	37	17	2	[10,10]	148	[8,8]	*130*	[1,15]	114	[1,14]	***110***	153
($q = 16$)	25	8	3	[7,7,7]	149	[5,5,5]	*122*	[4,8,8]	143	[0,7,8]	***110***	166
	24	5	4	-	-	[5,5,5,5]	**155**	-	-	[5,5,5,5]	**155**	180
SNOVA-III	56	25	2	[13,13]	201	[11,11]	*181*	[1,20]	152	[1,19]	***148***	221
($q = 16$)	49	11	3	-	-	[11,11,11]	251	-	-	[10,11,11]	246	**220**
	37	8	4	-	-	[7,7,7,7]	223	-	-	[3,8,8,8]	**214**	279
	24	5	5	-	-	[4,4,4,4,4]	*164*	-	-	[0,5,5,5,5]	***155***	257
SNOVA-V	75	33	2	[17,17]	262	[13,13]	*222*	[1,25]	***189***	[1,25]	***189***	287
($q = 16$)	66	15	3	-	-	[14,14,14]	324	-	-	[10,15,15]	313	**293**
	60	10	4	-	-	-	-	-	-	-	-	**343**
	29	6	5	-	-	[5,5,5,5,5]	*202*	-	-	[0,5,6,6,6]	***181***	310

linearly independent equations for a given block. Then the number of variables is given by

$$U = \left(\binom{n}{v} - E_0 \right)^{\ell}.$$

The rank is still expected to be $U - 1$, and the system is expected to be solvable if $U - 1$ equations are chosen at random. The drawback of this method is that eliminating variables does not preserve the sparsity of the equations. As a result, the complexity is given by $\mathcal{O}(U^{\omega})$. Some complexity estimates are given in Table 3, which do not outperform the sparse complexity estimates.

5 Conclusion

In this work, we revisit the wedge attack against UOV-like variants and extend this approach to a new attack on SNOVA by exploiting its block-ring structure. In addition to improving the attack complexity, the way we exploit the block-ring structure eliminates spurious solutions, which prevent generic versions of the wedge attack from working against SNOVA. Our attack breaks 6 of the 11 currently proposed SNOVA parameter sets and improves on the previous best result for an additional 2 sets; it is significantly more effective against larger ℓ values. For example, for SNOVA-V with $(v, o, \ell) = (29, 6, 5)$ the estimated security drops to 181 bits compared to 310 bits previously.

6 Future Work

There are several extensions of our attack, which remain outside the scope of the current state of our complexity analysis. In particular, when our attack is viewed as a quadratic system as described in Section 4.8, we have only computed

concrete complexities for parameterizations of our attack that correspond to solving the system at degree-ℓ. Nonetheless, it is possible that a better complexity may be obtained by solving the system at a higher degree, since doing so may allow us to use a smaller value of o' in the cases where an attack with solving degree-ℓ is possible. Allowing a higher solving degree may also allow us to solve the system in the cases where no parametrization of our attack works at degree-ℓ.

While we leave for future work detailed analysis of an extension of our quadratic modeling to higher degrees, a preliminary indication that such an extension is viable can be obtained by comparison with the p^L-truncated[8] polynomial ring attack [20]. In particular, the p^L-truncated polynomial ring variant of the reconciliation attack is closely related to our attack. This relationship can be seen from the fact that, if we combine our attack with the technique for removing spurious solutions [25], we obtain precisely the same attack complexities as in [25] for the $\ell = 2$ parameter sets of SNOVA; attacking SNOVA-I$(37, 17, 2)$ with $\mathbf{o}' = (0, 15)$, SNOVA-III$(56, 25, 2)$ with $\mathbf{o}' = (0, 20)$, and SNOVA-V$(75, 33, 2)$ with $\mathbf{o}' = (0, 26)$ yields matching attack complexities 2^{103}, 2^{140} and 2^{180} bit operations, respectively. For other parameters, our attack gives better complexities than the p^L-truncated attack, but we cannot solve our modeling at degree ℓ for one parameter set: SNOVA-V (60,10,4). In [20], the authors estimated that the complexity of the p^L-truncated reconciliation attack at 2^{695}, but did so using $L = 2$ rather than $L = 1$. We believe this indicates by analogy that the relevant system can be solved by going to higher degree, and since our attack often obtains better attack complexities against SNOVA when it is applicable, we believe extending our attack to higher degree can likely improve on the 2^{695} number. Nonetheless it remains unclear whether the improvement would be enough to improve on other attacks or bring the security of the parameter set below its claimed security level.

Acknowledgments. The authors would like to thank Lars Ran for helpful discussions. The second and fourth authors were supported by the France 2030 ANR Project ANR-22-PECY-003 SecureCompute and ANR grant PQMC (ANR-25-CE39-7662).

References

1. Aardal, M.A., Adj, G., Aranha, D.F., Basso, A., Martínez, I.A.C., Chávez-Saab, J., Santos, M.C.R., Dartois, P., Feo, L.D., Duparc, M., Eriksen, J.K., Fouotsa, T.B., , Filho, D.L.G., Hess, B., Kohel, D., Leroux, A., Longa, P., Maino, L., Meyer, M., Nakagawa, K., Onuki, H., Panny, L., Patranabis, S., Petit, C., Pope, G., Reijnders, K., Robert, D., Henríquez, F.R., Schaeffler, S., Wesolowski, B.: SQIsign: Specification Document – Round 2. Technical Report Round 2 (March 2025), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/round-2/spec-files/sqisign-spec-round2-web.pdf, available at NIST CSRC

[8] We replace the attack parameter ℓ from [20] by L to avoid confusion with the SNOVA parameter ℓ [41].

2. Alagic, G., Bros, M., Ciadoux, P., Cooper, D., Dang, Q., Dang, T., Kelsey, J., Lichtinger, J., Liu, Y.K., Miller, C., Moody, D., Peralta, R., Perlner, R., Robinson, A., Silberg, H., Smith-Tone, D., Waller, N.: Status report on the first round of the additional digital signature schemes for the nist post-quantum cryptography standardization process. NIST IR **8528** (2024)
3. Baena, J., Briaud, P., Cabarcas, D., Perlner, R., Smith-Tone, D., Verbel, J.: Improving support-minors rank attacks: applications to GeMSS and Rainbow. In: Annual International Cryptology Conference. pp. 376–405. Springer (2022)
4. Bardet, M., Faugère, J.C., Salvy, B.: Complexity of Gröbner basis computation for Semi-regular Overdetermined sequences over $\mathbb{F}_2$ with solutions in $\mathbb{F}_2$. Research Report RR-5049, INRIA (2003), https://inria.hal.science/inria-00071534
5. Benadjila, R., Bouillaguet, C., Feneuil, T., Rivain, M.: MQOM: Specification Document – Round 2. Technical Report Round 2 (March 2025), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/round-2/spec-files/mqom-spec-round2-web.pdf, available at NIST CSRC
6. Beullens, W.: Breaking Rainbow takes a weekend on a laptop. In: Dodis, Y., Shrimpton, T. (eds.) Advances in Cryptology - CRYPTO 2022 - 42nd Annual International Cryptology Conference, CRYPTO 2022, Santa Barbara, CA, USA, August 15-18, 2022, Proceedings, Part II. Lecture Notes in Computer Science, vol. 13508, pp. 464–479. Springer (2022). https://doi.org/10.1007/978-3-031-15979-4_16, https://doi.org/10.1007/978-3-031-15979-4_16
7. Beullens, W.: Improved cryptanalysis of SNOVA. In: Advances in Cryptology - EUROCRYPT 2025. Lecture Notes in Computer Science, vol. 15606, pp. 277–293. Springer (2025)
8. Beullens, W., Campos, F., Celi, S., Hess, B., Kannwischer, M.J.: MAYO: Specification Document – Round 2. Technical Report Round 2 (March 2025), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/round-2/spec-files/mayo-spec-round2-web.pdf, available at NIST CSRC
9. Beullens, W., Chen, M.S., Ding, J., Gong, B., Kannwischer, M.J., Patarin, J., Peng, B.Y., Schmidt, D., Shih, C.J., Tao, C., Yang, B.Y.: UOV: Specification document – round 2. Technical Report Round 2 (March 2025), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/round-2/spec-files/uov-spec-round2-web.pdf, available at NIST CSRC
10. Beullens, W., Preneel, B.: Field lifting for smaller UOV public keys. In: International Conference on Cryptology in India. pp. 227–246. Springer (2017)
11. Bosma, W., Cannon, J., Playoust, C.: The Magma algebra system. I. The user language. J. Symbolic Comput. **24**(3-4), 235–265 (1997). https://doi.org/10.1006/jsco.1996.0125, http://dx.doi.org/10.1006/jsco.1996.0125, computational algebra and number theory (London, 1993)
12. Cabarcas, D., Li, P., Verbel, J.A., Villanueva-Polanco, R.: Improved attacks for SNOVA by Exploiting Stability under a Group action. In: CRYPTO (2025)
13. Castryck, W., Decru, T.: An efficient key recovery attack on SIDH. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. pp. 423–447. Springer (2023)
14. Coppersmith, D.: Solving homogeneous linear equations over GF(2) via block Wiedemann algorithm. Mathematics of Computation **62**(205), 333–350 (1994)
15. Ding, J., Chen, M.S., Petzoldt, A., Schmidt, D., Yang, B.Y., Kannwischer, M., Patarin, J.: Rainbow: Proposal for NISTPQC: Digital signature algorithms version 3.0. Technical Report Round 3 (October 2020), https://csrc.nist.gov/Projects/post-quantum-cryptography/

post-quantum-cryptography-standardization/round-3-submissions, available at NIST CSRC
16. Ding, J., Deaton, J., Vishakha, Yang, B.: The nested subset differential attack - A practical direct attack against LUOV which forges a signature within 210 minutes. In: Canteaut, A., Standaert, F. (eds.) Advances in Cryptology - EUROCRYPT 2021 - 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17-21, 2021, Proceedings, Part I. Lecture Notes in Computer Science, vol. 12696, pp. 329–347. Springer (2021). https://doi.org/10.1007/978-3-030-77870-5_12, https://doi.org/10.1007/978-3-030-77870-5_12
17. Eagon, J.A., Northcott, D.G.: Ideals defined by matrices and a certain complex associated with them. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **269**(1337), 188–204 (1962)
18. Eisenbud, D.: The geometry of syzygies. Graduate Texts in Mathe (2003)
19. Furue, H., Ikematsu, Y.: A new cryptanalysis against UOV-based variants MAYO, QR-UOV and VOX. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **108**(3), 174–182 (2025). https://doi.org/10.1587/TRANSFUN.2024CIP0003, https://doi.org/10.1587/transfun.2024cip0003
20. Furue, H., Ikematsu, Y.: Key recovery attacks on UOV using p^{ℓ}-truncated polynomial rings. In: Heninger, N., Rosulek, M. (eds.) Advances in Cryptology – CRYPTO 2026. pp. 289–318. Springer Nature Switzerland, Cham (2026)
21. Furue, H., Ikematsu, Y., Hoshino, F., Takagi, T., Kosuge, H., Yamakoshi, K., Akiyama, R., Nakamura, S., Orihara, S., Kinjo, K.: QR-UOV: Specification Document – Round 2. Technical Report Round 2 (March 2025), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/qruov-spec-round2-web.pdf, available at NIST CSRC
22. Furue, H., Ikematsu, Y., Nakamura, S., Akiyamal, R.: Improved cryptanalysis of SNOVA by solving multi-homogeneous systems. In: Advances in Cryptology–ASIACRYPT 2025: 31st International Conference on the Theory and Application of Cryptology and Information Security, Melbourne, VIC, Australia, December 8–12, 2025, Proceedings, Part IV. p. 405. Springer Nature (2026)
23. Hilbert, D.: Ueber die theorie der algebraischen formen. Mathematische annalen **36**(4), 473–534 (1890)
24. Ikematsu, Y., Akiyama, R.: Revisiting the security analysis of SNOVA. IACR Cryptol. ePrint Arch. p. 96 (2024), https://eprint.iacr.org/2024/096
25. Ikematsu, Y., Furue, H.: Extending the applicability of algebraic key recovery attacks on the UOV signature scheme. Cryptology ePrint Archive, Paper 2026/1620 (2026), https://eprint.iacr.org/2026/1620
26. Jaques, S.: Memory adds no cost to lattice sieving for computers in 3 or more spatial dimensions. Cryptology ePrint Archive (2024)
27. Jin, Y., Pan, Y., He, X., Gong, B., Ding, J.: Security analysis on UOV families with odd characteristics: Using symmetric algebra. Cryptology ePrint Archive, Paper 2025/1137 (2025), https://eprint.iacr.org/2025/1137
28. Kaltofen, E.: Analysis of Coppersmith's block Wiedemann algorithm for the parallel solution of sparse linear systems. Mathematics of Computation **64**(210), 777–806 (1995)
29. Kipnis, A., Patarin, J., Goubin, L.: Unbalanced Oil and Vinegar Signature Schemes. In: Stern, J. (ed.) Advances in Cryptology - EUROCRYPT '99, International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, May 2-6, 1999, Proceeding. Lecture Notes in

Computer Science, vol. 1592, pp. 206–222. Springer (1999), https://doi.org/10.1007/3-540-48910-X_15
30. Le, H., Bros, M., Lichtinger, J., Minaud, B., Perlner, R., Smith-Tone, D., Valenzuela, C.: Exploiting SNOVA's structure in the wedge product attack. 6th NIST PQC Standardization Workshop (Sep 2025)
31. Li, P., Ding, J.: Cryptanalysis of the SNOVA signature scheme. In: Saarinen, M.O., Smith-Tone, D. (eds.) Post-Quantum Cryptography - 15th International Workshop, PQCrypto 2024, Oxford, UK, June 12-14, 2024, Proceedings, Part II. Lecture Notes in Computer Science, vol. 14772, pp. 79–91. Springer (2024), https://doi.org/10.1007/978-3-031-62746-0_4
32. Maino, L., Martindale, C.: An attack on SIDH with arbitrary starting curve. IACR Cryptol. ePrint Arch. p. 1026 (2022), https://eprint.iacr.org/2022/1026
33. Nakamura, S., Tani, Y., Furue, H.: Lifting approach against the SNOVA scheme. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences p. 2024EAP1124 (2025)
34. NIST: Call for Additional Digital Signature Schemes for the Post-quantum Cryptography Standardization Process. NIST CSRC (2022), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/call-for-proposals-dig-sig-sept-2022.pdf
35. Patarin, J., Cogliati, B., Faugère, J.C., Fouque, P.A., Goubin, L., Larrieu, R., Macario-Rat, G., Minaud, B.: VOX specification v1.0. Round 1 Additional Signatures, NIST Post-Quantum Cryptography: Digital Signature Schemes (Jun 2023), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/round-1/spec-files/vox-spec-web.pdf, version 1.0
36. Ran, L.: Wedges, oil, and vinegar: An analysis of UOV in the exterior algebra. In: Daemen, J., Thomé, E. (eds.) Advances in Cryptology – EUROCRYPT 2026. Lecture Notes in Computer Science, vol. 16544, pp. 363–388. Springer, Cham (2026)
37. Robert, D.: Breaking SIDH in polynomial time. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. pp. 472–503. Springer (2023)
38. The Sage Developers: SageMath, the Sage Mathematics Software System (Version 11.0). The Sage Development Team (2026), https://www.sagemath.org
39. Tseng, P.E., Wang, L.C., Li, P., Kuan, Y.L.: Investigating the wedge map on SNOVA. Cryptology ePrint Archive, Paper 2026/260 (2026), https://eprint.iacr.org/archive/2026/260/20260225:080243
40. Villard, G.: A study of Coppersmith's block Wiedemann algorithm using matrix polynomials. Citeseer (1997)
41. Wang, L.C., Chou, C.Y., Ding, J., Kuan, Y.L., Leegwater, J.A., Li, M.S., Tseng, B.S., Tseng, P.E., Wang, C.C.: SNOVA: Specification Document – Round 2. Technical Report Round 2 (March 2025), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/round-2/spec-files/snova-spec-round2-web.pdf, available at NIST CSRC
42. Yokonuma, T.: Tensor spaces and exterior algebra. No. 108, American Mathematical Soc. (1992)

A Spurious solutions in unbalanced projection cases

Table 2: Rank computation results under block projection.

n	v	m	ℓ	$o'_{\text{by block}}$	Nonzero minors	Rank$M$	Dif
14	8	4	2	$[0,6]$	3003	3001	2
4	2	1	3	$[1,2,2]$	108	107	1
16	9	4	2	$[0,7]$	11440	11423	$17 = \binom{16}{9-2\times4} + 1$
10	7	3	3	$[0,3,3]$	14400	14399	1
9	4	2	3	$[0,0,5]$	126	124	2

B Cost of the Attack in Quadratic Equations Modeling

Table 3: Estimated parameters of the quadratic modeling when variables are eliminated. Balanced projection is used where $\boldsymbol{o'}$ are taken from Table 1, constant factors are not included, and the dash indicates that no suitable $\boldsymbol{o'}$ was found.

Variants	$\boldsymbol{v}$	$\boldsymbol{m}$	$\boldsymbol{\ell}$	$\boldsymbol{o'}$	$\log_2 \boldsymbol{E_0}$	$\log_2 \boldsymbol{U}$	$\log_2(\boldsymbol{U^\omega})$
SNOVA-I ($q = 16$)	37	17	2	8	27.5	48.6	136
	25	8	3	5	16.4	47.2	132
	24	5	4	5	16.0	62.6	176
SNOVA-III ($q = 16$)	56	25	2	11	40.2	71.4	201
	49	11	3	11	38.2	105.4	296
	37	8	4	7	24.7	93.2	262
	24	5	5	4	13.2	67.4	189
SNOVA-V ($q = 16$)	75	33	2	13	50.0	91.8	258
	66	15	3	14	50.3	138.9	390
	60	10	4	–	–	–	–
	29	6	5	5	17.3	84.0	236

C Experimental Results for Proposition 2 and 3

Table 4: Actual rank of $\Phi_{\text{SNOVA}}^{\langle m \rangle}$ when using m base 2-forms. Only parameters where the number of linear independent equations is less than the number of variables are reported, so there is a meaningful lower bound.

n	v	m	ℓ	Experimental Rank	Difference between experimental rank and Formula 9
			$\ell = 2$		
5	3	1	2	43	0
5	3	2	2	83	0
6	4	2	2	125	0
7	4	1	2	793	0
7	5	2	2	175	0
8	6	2	2	233	0
8	6	3	2	345	0
8	6	4	2	454	0
8	5	1	2	1498	0
8	5	2	2	2700	0
9	5	1	2	12075	0
9	6	1	2	2576	0
9	6	2	2	4773	0
9	6	3	2	6598	0
10	7	1	2	4137	0
10	7	2	2	7802	0
10	7	3	2	11002	0
			$\ell = 3$		
3	2	2	3	18	0
4	3	2	3	24	0
4	3	3	3	36	0
4	3	4	3	48	0
4	3	5	3	60	0
7	5	1	3	4126	0
7	5	2	3	7677	0
8	6	2	3	13366	0

Table 5: Actual rank of $\Phi_{\text{SNOVA}}^{\langle \ell m \rangle}$ when using ℓm scaled 2-forms. Only parameters where the number of linear independent equations is less than the number of variables is reported, so there is a meaningful lower bound.

n	v	m	ℓ	**Experimental Rank**	**Difference between experimental rank and Formula 10**
5	3	1	2	66	0
6	4	1	2	98	0
6	4	2	2	190	0
7	4	1	2	1098	0
7	5	1	2	136	0
7	5	2	2	266	0
7	5	3	2	390	0
8	5	1	2	2102	0
8	6	1	2	180	0
8	6	2	2	354	0
8	6	3	2	522	0
8	6	4	2	684	0
9	5	1	2	15489	0
9	6	1	2	3642	0
9	6	2	2	6534	0
9	7	1	2	230	0
9	7	2	2	454	0
9	7	3	2	672	0
9	7	4	2	884	0
9	7	5	2	1090	0
9	7	6	2	1290	0
10	6	1	2	35331	0
10	7	1	2	5876	0
10	7	2	2	10816	0
10	8	1	2	286	0
10	8	2	2	566	0
10	8	3	2	840	0
10	8	4	2	1108	0
10	8	5	2	1370	0
10	8	6	2	1626	0
10	8	7	2	1876	0
11	8	1	2	8982	0
11	8	2	2	16822	0
11	8	3	2	23540	0
3	2	1	3	18	0
4	3	1	3	24	0
4	3	2	3	48	0
5	4	1	3	30	0
5	4	2	3	60	0
5	4	3	3	90	0
5	4	4	3	120	0
6	4	1	3	3105	0
6	5	1	3	36	0
6	5	2	3	72	0
6	5	3	3	108	0
6	5	4	3	144	0
6	5	5	3	180	0
7	5	1	3	6381	0

D Cost of our Attack in a 2D Memory Model

When evaluating the cost of an attack, so far we have assumed free memory access. In this section, we propose a more refined analysis of the cost incurred by memory accesses, based on a 2D memory model, following [3]. Recall that ℓ is the number of blocks, and for each block, a minor of size v is selected. After projection, block r contains $n_r = v + o'_r$ indices.

A column of the linear system is indexed by a tuple $\mathbf{T} = (T_1, \ldots, T_\ell)$ with $T_r \in \binom{[n_r]}{v}$. Here $\binom{[n_r]}{v}$ means the set of subsets of size v of $[n_r]$. For every block, split the indices into: h_r watched indices and $n_r - h_r$ unwatched indices. A hash address records which watched indices belong to T_r.

Let us now consider one particular hash address $H_r(T_r)$ having weight i_r. That means there are i_r watched indices belong to T_r, because $|T_r| = v$, we still need to choose $v - i_r$ indices from $n_r - h_r$ unwatched indices. Therefore, such a fixed address contains $\binom{n_r - h_r}{v - i_r}$ possible T_r. Note that there are $\binom{h_r}{i_r}$ different addresses of weight i_r.

Since the choices in different blocks are independent, if we consider ℓ blocks instead of one particular hash address, the size of each bucket having weight vector i is $U_{\mathbf{i}} = \prod_{r=1}^{\ell} \binom{n_r - h_r}{v - i_r}$ and the number of such buckets is $N_{\mathbf{i}} = \prod_{r=1}^{\ell} \binom{h_r}{i_r}$. Then every coordinate is allocated exactly once because

$$\sum_{\mathbf{i}} N_{\mathbf{i}} U_{\mathbf{i}} = \prod_{r=1}^{\ell} \sum_{i_r} \binom{h_r}{i_r} \binom{n_r - h_r}{v - i_r} = \prod_{r=1}^{\ell} \binom{n_r}{v} = U.$$

D.1 Memory access cost of one field element within a partition.

We denote $u_r = \binom{n_r}{v}$ for simplicity. The Type 1 and Type 2 equations described in Section 4.6 show that every row has exactly the below structure:

1. Type 1 (block k): $T_k = J_k \setminus \{a, b\} \implies |J_k| = v + 2$
   $\implies$ # support families $S_{1,k} = \sum_k \binom{n_k}{v+2} \prod_{r \neq k} u_r$
2. Type 2 (blocks k, k')
   - $T_k = J_k \setminus \{a\} \implies |J_k| = v + 1$,
   - $T_{k'} = J_{k'} \setminus \{b\} \implies |J_{k'}| = v + 1$.

   $\implies$ # support families $S_{2,k,k'} = \sum_{k<k'} \binom{n_k}{v+1} \binom{n_{k'}}{v+1} \prod_{r \notin \{k,k'\}} u_r$

Then total number of support families is $S = S_1 + S_2$. Then each field element is accessed $\rho = \frac{S}{U}$ times on average per multiplication by a processing cluster. Taking $C_2 = 2^{-12.8}$ [26], the average memory access cost of a field element within some memory partition is given by

$$\psi_1 = C_2 \frac{\sqrt{2L}}{U} \sum_{\mathbf{i}: U_{\mathbf{i}} > 0} N_{\mathbf{i}} \big(L + \log_2(2LU_{\mathbf{i}})\big) U_{\mathbf{i}}^{3/2} \text{ where } L = \log_2(q^\ell).$$

D.2 Additional costs due to the hashing strategy.

For SNOVA system with the density given in Section 4.6, weighted-average row weight is

$$w = \frac{S_1\binom{v+2}{2} + S_2(v+1)^2}{S_1 + S_2}.$$

Note that the additional cost is determined by the number of distinct hash addresses per row. Thus, we suppose that $0 \le h_r \le v$ and by that, we can ensure that a Type 1 support contains at least two unwatched indices and a Type 2 support contains at least one unwatched index.

For Type 1 equations, let J_k be a uniformly chosen $(v+2)$-subset in block k, and let $j = |J_k \cap D_k|$. The number of addresses touched by this row is

$$\binom{j}{2} + j + 1$$

Averaging over all uniformly chosen Type 1 supports J_k, the expected number of distinct addresses touched per Type 1 row is

$$\kappa_{1,k} = \sum_{j=0}^{h_k} \left(\left(\binom{j}{2} + j + 1 \right) \frac{\binom{h_k}{j}\binom{n_k - h_k}{v+2-j}}{\binom{n_k}{v+2}} \right).$$

For Type 2 equations, by the same argument, the expected number of distinct addresses touched by one Type 2 row is

$$\kappa_{2,k,k'} = \sum_{j=0}^{h_k} \sum_{j'=0}^{h_{k'}} \left((j+1)(j'+1) \frac{\binom{h_k}{j}\binom{h_{k'}}{j'}\binom{n_k - h_k}{v+1-j}\binom{n_k - h_{k'}}{v+1-j'}}{\binom{n_k}{v+1}^2} \right).$$

Because a row touching κ addresses sends κ field elements on average, the mean number of transmissions per row is

$$g = \frac{\sum_k S_{1,k}\kappa_{1,k} + \sum_{k<k'} S_{2,k,k'}\kappa_{2,k,k'}}{S_1 + S_2}.$$

Each transmitted partial sum contains L bits and crosses a memory of size $2UL$. Thus the additional communication cost per multiplication is

$$\psi_2 = C_2 \frac{g}{w} L\sqrt{2UL}$$

Optimal h. Following [3], we need to find optimal $\mathbf{h} = (h_1, \ldots, h_\ell)$ such that the following sum is smallest

$$\text{Total Memory Cost Per Multiplication} = \rho\psi_1 + \psi_2$$

where ψ_1 is the memory access cost of one field element within a partition and ψ_2 is the additional overhead costs due to the hashing strategy as computed. We report the optimal $\mathbf{h}$ for each parameter set and the cost in the 2D Memory Model with comparison to the RAM model in Table 6.

Table 6: Optimized hashes and corresponding 2D Memory Model costs in base-2 logarithms using $C_2 = 2^{-12.8}$ as in [26].

Variant	v	m	ℓ	$\mathbf{o}'$	$\mathbf{h}$	$c_{\mathbf{RAM}}$	$c_{\mathbf{2D}}$	Δ
SNOVA-I	37	17	2	$[1, 14]$	$(37, 6)$	110	113	3
SNOVA-I	25	8	3	$[0, 7, 8]$	$(0, 15, 0)$	110	114	4
SNOVA-I	24	5	4	$[5, 5, 5, 5]$	$(0, 0, 0, 22)$	155	169	14
SNOVA-III	56	25	2	$[1, 19]$	$(56, 8)$	148	159	11
SNOVA-III	49	11	3	$[10, 11, 11]$	$(28, 0, 0)$	246	282	36
SNOVA-III	37	8	4	$[3, 8, 8, 8]$	$(38, 0, 0, 4)$	214	241	27
SNOVA-III	24	5	5	$[0, 5, 5, 5, 5]$	$(0, 0, 0, 0, 21)$	155	169	14
SNOVA-V	75	33	2	$[1, 25]$	$(74, 10)$	189	210	21
SNOVA-V	66	15	3	$[10, 15, 15]$	$(37, 0, 0)$	313	365	52
SNOVA-V	60	10	4	-	-	-	-	-
SNOVA-V	29	6	5	$[0, 5, 6, 6, 6]$	$(0, 26, 0, 0, 0)$	181	201	20