Statistical Decoding 2.0: Reducing Decoding to LPN

Kévin Carrier[1], Thomas Debris-Alazard[2], Charles Meyer-Hilfiger[3], and Jean-Pierre Tillich[3]

[1] ETIS Laboratory, CY Cergy-Paris University, kevin.carrier@ensea.fr
[2] Project GRACE, Inria Saclay-Ile de France, thomas.debris@inria.fr
[3] Project COSMIQ, Inria de Paris,
charles.meyer-hilfiger@inria.fr,jean-pierre.tillich@inria.fr

Abstract. The security of code-based cryptography relies primarily on the hardness of generic decoding of linear codes. Essentially, the decoding problem come back to solving a linear system with a constraint on the hamming weight of the solution. The best generic decoding algorithms are all improvements of an old algorithm due to Prange: they are known under the name of information set decoders (ISD). A while ago, a generic decoding algorithm which does not belong to this family was proposed: statistical decoding. It is a randomized algorithm that requires the computation of a large set of parity-checks of moderate weight, and uses some kind of majority voting on these equations to recover the error. This algorithm was long forgotten because even the best variants of it performed poorly when compared to the simplest ISD algorithm. We revisit this old algorithm by using parity-check equations in a more general way. Here the parity-checks are used to get LPN samples with a secret which is part of the error and the LPN noise is related to the weight of the parity-checks we produce. The corresponding LPN problem is then solved by standard Fourier techniques. By properly choosing the method of producing these low weight equations and the size of the LPN problem, we are able to outperform in this way significantly information set decoders at code rates smaller than 0.3. It gives for the first time after 60 years, a better decoding algorithm for a significant range which does not belong to the ISD family.